Università degli Studi di Camerino

School of Advanced Studies

Doctorate Course in

Life & Health Sciences – One Health

XXXIII Cycle

Improved Alzheimer's disease detection by MRI using multimodal machine learning algorithms

PhD Student	Curriculum Coordinator	Supervisor
Gopi Battineni	**Prof. Anna Maria Eleuteri**	**Prof. Francesco Amenta**

Abstract

Dementia is one of the huge medical problems that have challenged the public health sector around the world. Moreover, it generally occurred in older adults (age > 60). Shockingly, there are no legitimate drugs to fix this sickness, and once in a while it will directly influence individual memory abilities and diminish the human capacity to perform day by day exercises. Many health experts and computing scientists were performing research works on this issue for the most recent twenty years. All things considered, there is an immediate requirement for finding the relative characteristics that can figure out the identification of dementia.

The motive behind the works presented in this thesis is to propose the sophisticated supervised machine learning model in the prediction and classification of AD in elder people. For that, we conducted different experiments on open access brain image information including demographic MRI data of 373 scan sessions of 150 patients. In the first two works, we applied single ML models called support vectors and pruned decision trees for the prediction of dementia on the same dataset. In the first experiment with SVM, we achieved 70% of the prediction accuracy of late-stage dementia. Classification of true dementia subjects (precision) is calculated as 75%. Similarly, in the second experiment with J48 pruned decision trees, the accuracy was improved to the value of 88.73%. Classification of true dementia cases with this model was comprehensively done and achieved 92.4% of precision.

To enhance this work, rather than single modelling we employed multi-modelling approaches. In the comparative analysis of the machine learning study, we applied the feature reduction technique called principal component analysis. This approach identifies the high correlated features in the dataset that are closely associated with dementia type. By doing the simultaneous application of three models such as KNN, LR, and SVM, it has been possible to identify an ideal model for the classification of dementia subjects. When compared with support vectors, KNN and LR models comprehensively classified AD subjects with 97.6% and 98.3% of accuracy respectively. These values are relatively higher than the previous experiments.

However, because of the AD severity in older adults, it should be mandatory to not leave true AD positives. For the classification of true AD subjects among total subjects, we enhanced the model accuracy by introducing three independent experiments. In this work, we incorporated two new models called Naïve Bayes and Artificial Neural Networks along support vectors and KNN. In the first experiment, models were independently developed with manual feature selection. The experimental outcome suggested that KNN

is the optimal model solution because of 91.32% of classification accuracy. In the second experiment, the same models were tested with limited features (with high correlation). SVM was produced a high 96.12% of classification accuracy and NB produced a 98.21% classification rate of true AD subjects. Ultimately, in the third experiment, we mixed these four models and created a new model called hybrid type modelling. Hybrid model performance is validated AU-ROC curve value which is 0.991 (i.e., 99.1% of classification accuracy) has achieved. All these experimental results suggested that the ensemble modelling approach with wrapping is an optimal solution in the classification of AD subjects.

Declarations

Myself Gopi Battineni, hereby declare that the presenting thesis entitled "Comprehensive machine learning algorithms to evaluate the older adult dementia progression" is my original academic work that was done during my PhD period (2017- 2020), under the supervision of Prof. Francesco Amenta, Director of School of medicinal and health products sciences, University of Camerino, Italy. I submit this thesis to the fulfilment of the requirements for the award of Doctor of Philosophy in One health, to the International School of Advanced Studies, University of Camerino, Italy. We have critically reviewed and approved the final draft of the thesis and are responsible for the content and similarity index of it.

Acknowledgements

I wish to express my sincere gratitude to my supervisor Prof. Francesco Amenta for his warm kindness towards me during the period of doctoral studies. I strongly cherish his continuous research feedback, theoretical, and backing up from the first day of my PhD work. With his ultimate guidance, he corrected my mistakes at each step and helped me to grow up as a fruitful researcher. Prof. Francesco Amenta is one of the kind-hearted people which I came across in my life and encourage all of his research team to create novel works that can help the real-time public. I must say thanks for his hard work during my manuscript's corrections, it boosts up my confidence and helped me to work with more enthusiasm.

I appreciate the kind guidance of Prof. Seyed Khosrow Tayebati and Prof. Daniele Tomassoni in enriching my research experience. Besides, I would like to acknowledge the efforts of my colleagues Giulio Nittari, Graziano Pallotta and Simone Baldoni during their research collaboration.

I am very thankful to the director of the International School of Advanced Studies Prof. Roberto Ciccocioppo for providing the PhD scholarship. I would like to say many thanks to Cristina Soave and other SAS staff members for providing me with the right information and helping me in every administrative query. They are very kind heart and highly supportive of me.

I would like to convey my sincere thanks to my wife Nalini Chintalapudi who supported me during experimental works. Besides, I am very thankful to my friend Getu Gamo Sagaro for his extended support during the preparation of manuscripts. Because of their kind of collaboration, it is only possible to publish more than 50 research papers during my tenure. Last but not least, this thesis is dedicated to my mom for her unconditional love and support since my childhood. Without her support, I would never have come across to journey of the doctoral degree.

Table of contents

List of Figures

List of Tables

Chapter 1

1. Introduction

Dementia is a disorder – generally of a progressive or chronic, in which there is weakening in intellectual capacity (for example the ability of thinking process) earlier than what may be normal from normal ageing [1]. It influences memory, thinking, perception, estimation, orientation, capacity of learning, language, and judgment. However, there is no effect on consciousness [2]. The exhaustion in cognitive function normally goes with and once in a while went before, by disintegrating in social behaviour, controlling of emotions, or inspiration.

Dementia results from an assortment of infections and wounds that fundamentally or optionally influence the mind, for example, Alzheimer's Disease (AD) or stroke. Dementia is one of the significant reasons for incapacity and reliance among older adults around the world [3]. It may be uncontrollable, for the individuals who have it, yet also for their careers and families. There is frequently an absence of mindfulness and dementia understanding, which results in stigmatization and difficulties to care for and diagnose [2], [3]. The effect of dementia on society, carers, and family at large can be physical, mental, social, and financial.

A few people with dementia can't control their feelings, and their characters may change [4]. Dementia ranges in seriousness from the mildest stage, when it is simply starting to influence an individual's work, to the most extreme stage and the individuals had completely depended on others for their daily activities [5]. Signs and indications of dementia result when healthy neurons (nerve cells) of the brain suddenly stop functioning, lose connections with other neurons, and demise. While everybody loses a few neurons as they age, individuals with dementia experience far more loss.

While dementia is more normal as individuals become older (equal to half of all people age ≥85 may have some type of dementia), it is not a common phenomenon of ageing [6]. Numerous individuals live into their 90s and beyond with no indications of dementia [5], [7]. Similarly, a dementia-like frontotemporal disorder often happens in middle-aged people than older people.

The reasons for dementia can differ based on the types of brain alterations that can happen. AD is the most well-known reason for dementia in older adults. Other dementia types include Lewy body dementia, vascular dementia, and frontotemporal disorders. It is also common to have mixed dementia, which is a combination of at least two dementia

types [8]. For instance, some people can possess both Alzheimer's and vascular dementia.

Dementia influences every individual in an alternate manner based on the disease influence and the individual's character before getting sick. The signs and symptoms connected to dementia can be perceived in three ways such as early-stage dementia, moderate dementia, and late-stage dementia [9]. In the early-stage dementia is frequently unnoticed because the onset is progressive and symptoms included forgetfulness, getting lost in familiar pleases, and forgetting about the time. In moderate dementia, the symptoms and signs become clearer and more preventive which includes getting forgottenness about people's names, being lost at home, asking help for personal care, behaviour changes including repeated questions [1], [9]. Late-stage dementia represents inactivity and total dependence on others and symptoms include being unconscious of the place and time, experiencing issues perceiving family members and companions, walking difficulties, and behaviour changes like aggression.

1.1. Dementia types and social impact factors

There is a wide range of types of dementia. Alzheimer's is the most well-known type and may add to 60–70% of global cases. Other significant dementia forms include vascular dementia, dementia with Lewy bodies (irregular totals of protein that create inside nerve cells), and disease groups that add to frontotemporal dementia (degeneration of the frontal flap of the mind). The limits between various types of dementia are ambiguous and often mixed forms also coexist.

A survey suggests nearly 50 Million people had dementia and 60% of them are from developing countries. Every year, there is a chance of developing 10 Million cases every year [10]. The assessed proportion of the common public aged more than 60 years and with onset dementia is between 5-8%. The absolute number of individuals with dementia is extended to arrive at 82 million out of 2030 and 152 of every 2050 [11]. Quite a bit of this expansion is owing to the rising quantities of individuals with dementia living in low-and centre pay nations.

The total number of people with dementia is projected to reach 82 million in 2030 and 152 in 2050. Much of this increase is attributable to the rising numbers of people with dementia living in low- and middle-income countries. Dementia has important economic and social implications regarding direct social and medical care costs and the costs of casual care [12]. In 2015, the complete worldwide social expenses of dementia were assessed to be US$ 818 billion, equal to 1.1% of GDP worldwide [13]. The overall

expense as an extent of GDP differed from 0.2% in developing countries to 1.4% in developed countries.

Dementia would be uncontrollable for the patient's families and their careers. Pressures in the forms of emotional, physical, and financial can trouble people's lives, and support is needed from the social, financial, and health systems. Individuals with dementia are much of the time denied the essential rights and opportunities accessible to other people [14]. In some countries, physical and compound restrictions are utilized widely in-home care for older adults and in better care settings, although when guidelines are set up to maintain the privileges of individuals to opportunity and decision [15]. A good and consistent legal support is needed on globally acknowledged human rights are needed to guarantee the highest quality care for individuals with dementia and their careers.

1.2. Dementia types

As mentioned, dementia is caused by harm to or loss of nerve cells and their brain connections. Depending upon the brain region that is affected because of damage, dementia can influence individuals diversely and present various symptoms. Dementias are frequently assembled by what they share practically speaking, for example, the protein or proteins saved in the cerebrum or the brain part which is highly influenced. Some diseases seem to be dementias, for example, those brought about by a response to drugs or nutrient inadequacies, and they may improve with treatment.

1.2.1. Progressive dementias

Some dementias can be progressive and are not adjustable includes

Alzheimer's

Alzheimer's disease is the most well-known reason for dementia. However, not all reasons for AD are known, doctors do realize that a little rate is identified with transformations of three genes that can be accepted from parents to children [16]. While a few distinct genes are most likely engaged with Alzheimer's disease, one significant quality that builds risk is apolipoprotein E4 (APOE).

AD patients have plaques and tangles in their minds. Plaques are groups of a protein called beta-amyloid, and tangles are stringy knots comprised of tau protein [17]. It's an idea that these groups harm healthy neurons and the components associated with them. Other hereditary variables may make it almost certain that people can also develop AD.

Vascular dementia

This second most regular kind of dementia is happened by damage to the vessels that source blood to your cerebrum [18]. Problems caused by blood cells can cause strokes or harm the brain in other possibilities, for example, by the damage of the white matter of the brain. The most well-known indications of vascular dementia incorporate challenges with critical thinking, focusing, slow thinking, and organization. These will in general be more recognizable than cognitive decline.

Lewy body dementia

Lewy bodies are unusual balloons like clusters of protein that have been found in the cerebrums of individuals with Lewy body dementia, Parkinson's disease, and Alzheimer's [19]. This is also a type of progressive dementia. The symptoms and signs include visual hallucinations and issues with concentration and focus. Different signs incorporate slow movement and uncoordinated, inflexibility (parkinsonism), and tumours.

1.2.2. Frontotemporal dementia

This dementia is a collection of diseases that have been characterized by the breakdown of nerve cells and their associations in the frontal and fleeting projections of the cerebrum, the regions for the most part connected with behaviour, personality, and language [20].

1.2.3. Mixed dementia

Many studies on the brains of people over 80 years with dementia demonstrate that many had a combination of different reasons such as AD, Lewy body dementia, and vascular dementia [21]. Some studies are continuing to understand how mixed dementia can affect treatments and symptoms.

1.3. Dementia diagnosis and treatment

There is nobody test to decide whether somebody has dementia or not. Experts analyse Alzheimer's and different kinds of dementia dependent on a cautious clinical history, an actual assessment, tests in labs, and the progressive changes in thinking, daily capacity, and personal behaviour with others [22].

Doctors can assess a dementia person with a high level of certainties. Although, it is hard to decide the exact dementia type as that the side effects and brain changes of various dementias can be overlap. Sometimes, experts can diagnose dementia but do not specify which type it is. In such cases, it is recommended to meet specialists like neurologists or gero-clinician.

There is no precise treatment or cure for dementia, also to estimate its severity [23]. Various new medicines are being explored in different phases of clinical trials. But much can be offered to help and improve the lives of dementia patients and their families and careers. It is recommended to follow some measures to prevent the risk by

- ✓ Early finding to promote primary and ideal management
- ✓ Optimization of cognition, activity, wellbeing, and physical health
- ✓ Recognizing and treating going with physiological and behavioural symptoms
- ✓ Distinguishing and treating testing conduct and mental side effects
- ✓ Giving data and continuing support to carers.

Screening people in danger for AD depending on electronic health records (EHR) in preclinical stages may prompt early recognition of AD and to better remedial techniques for deferring the beginning of AD [24]. Current biomarkers of AD require the specimen's collection such as serum or liquid or image data. Thereafter, EHR data, for example, patient records in clinical settings, or regulatory health data, do not need extra time or effort for data collection. Additionally, with the advances in computer techniques, the measures of such information have gradually increased. Since it is universal, practical, and tremendous, the computational healthcare database might be an important asset for testing adaptable AD predictive models and similar to other diseases [24]. Nevertheless, regardless of its huge likely worth, the smaller scale of knowledge is known for the extent of understanding the use of healthcare data in the prediction of dementia.

For the prediction of dementia risk, earlier models are commonly associated with predefined medical profile factors, for example, sociodemographic (age, gender, education), cognitive profiles, physical activity, and midlife medical risk factors like body mass index (BMI), systolic circulatory strain, and full-body cholesterol level. But an important question is whether those generalized predictive models in clinical settings can effectively evaluate the heterogeneous aetiologies of multifactorial AD with a small number of selected variables. In reality, a meta-examination study shows that multifaceted models are the optimal choice of AD risk prediction, though single-factor models are not performing well [25], acclaiming precise AD risk prediction requires an enormous element space. In this work, we test the degrees to which an information-driven machine learning (ML) model is useful in large-scale health information that includes thousands of patient health records to make a precise prediction of AD risk. In the next sections, a brief discussion on how machine learning models are evolved in the last decades and their applications in the healthcare domain especially in AD risk predictions.

1.4. Machine learning

In this section, the authors like to discuss the basic concepts involved in machine learning (ML). ML is considered an integral part of Artificial Intelligence (AI) which presents a concept of software engineering which attempts to mimic computers like human behaviours [26]. One of the fundamental necessities for any intelligent method is learning. The greater part of the scientists today concurs that there is no knowledge without learning. Hence, AI techniques like ML and deep learning algorithms are important parts of machine intelligence.

Machine learning has been classified into many types like supervised, unsupervised, reinforcement [27]. Because the focus is on the "learning" field, there are many types that we can encounter as an expert. Some forms of learning describe full-fledged educational fields with a variety of algorithms such as "supervised learning." Some suggest powerful methods which can use for our projects, such as "transfer learning".

In supervised machine learning, a class of drawbacks involves employing a model to find out an association between input variables and also the target variable [28]. Models are matched on training information comprised of input and outputs and adapted to build predictions on taking a look at sets wherever solely with the provided inputs and the model outputs are compared to the withdrawn target variables and that are used to evaluate the model performance.

There are two main sorts of supervised learning problems: one is involved with a classification problem that can help to predict a category label and another one is a regression problem that helps to predict a numerical value [29]. In simple terms, we can write to them as

Classification: involved in class label prediction.

Regression: involved in the prediction of the numerical label.

Both regression and classification issues could have one or a lot of input parameters and these variables are also any data types like categorical or numerical [28], [29]. Algorithms are cited as "supervised" as a result of they learn by creating predictions given samples of input data, and also the models are supervised and corrected via a formula to higher predict the expected target outputs within the training dataset. Some supervised algorithms (such as logistic regression) have been specifically designed to do regression (such as linear regression) or classification and a few may be used for each sort of problem with simple changes (such as artificial neural networks).

Whereas unsupervised machine learning defines a class of problems that involve using a model to define or exclude relationships from data [30]. In a comparison of supervised learning, unsupervised machine learning only applies to out-of-product input or changing targets. Also, unsupervised learning no teacher adjusts the model, as is the case with supervised learning [31]. In this, there is no teacher or educator involved, and the algorithm must learn to process data without this guide. There are many unsupervised learning modes, although there are two main problems that a specialist often encounters: they involve the acquisition of data groups (clustering) and the density estimates that include summarizing the data transmission. Simply they can address as

Clustering: A problem involving access to data groups.

Density Estimation: Problem that includes summarizing data transmission.

An example of a grouping algorithm is k-Means where k refers to the number of data retrieval groups [32]. An example of an estimated algorithm calculation is the Kernel Density Estimation which involves using small groups of a closely related data sample to estimate the distribution of new scores in a problem area. Both clustering and densities estimation can be done to study patterns in the data. Additional unsupervised modelling methods can be used, such as visualization that includes graphing or data processing in a variety of ways and prediction that include reducing data size.

- Visibility: Unsupervised learning problem involving creating data sites (scatter plot matrix).

- Prediction: Uncontrolled learning problem that involves creating a representation of low data (principal component analysis).

Besides, reinforcement learning defines the type of issues in which the agent is working on the situation and must learn to use the report [33]. Environmental use means that there is no standard training data, instead of the goal or set of goals to be achieved by the agency, the actions that can be taken, and the report on performance-oriented goals. It is in line with supervised learning because the model has a specific response that the user can learn from, although the report can be delayed and mathematically noisy, posing a challenge to the agent or model for combining cause and effect [34]. An example of a reinforcement problem is playing a game where the agent aims to get high scores and can make a move in the game and get a report in terms of penalties or compensation.

Finally, the lines between supervised and supervised learning are disturbed, and there are many hierarchical patterns from each study section. In this work, we will look more closely at some of the mixed learning (hybrid modelling) concepts: self-supervised, semi-supervised, and multidisciplinary learning.

Deep learning is a subcategory in machine learning algorithms that are learned by artificial neural networks (ANN). These algorithms are applied in self-driving vehicles and automatic voice assistants like SIRI or ALEXA [35]. Deep learning plays a vital role in voice control in consumer goods such as mobiles, TVs, tablets, and Wi-Fi connected speakers. These algorithms conduct feature selection algorithms to detect image or voice characteristics [35], [36]. The association between AI technologies including machine learning and deep learning is presented in Figure 1.1. In clinical practice, deep learning algorithms have the capability of producing high disease prediction accuracy which outperformed human intelligence. ANN models contain many layers and are trained by large samples of labelled data.

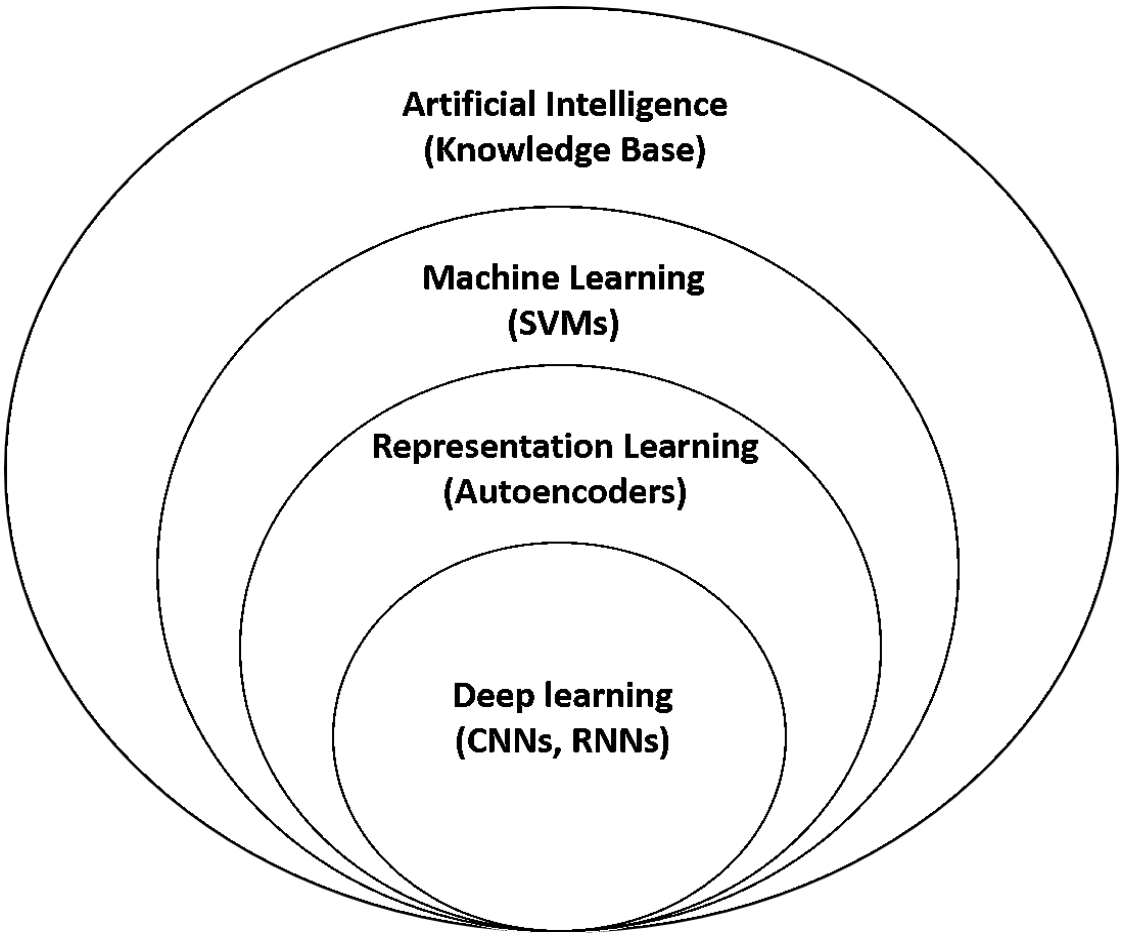

Figure 1.1. Relationship between AI, ML, and deep learning.

1.4.1. ML framework and performance calculations

The simple machine-learning framework in medical diagnosis is explained in Figure 1.2. The framework included seven individual steps to evaluate the final disease diagnosis and each step was further explained in detail.

Data collection: The model accuracy was decided by the quantity and quality of input medical data. The outcome of the data collection step represents the data used for training purposes. The medical datasets that available from the UCI repository, Kaggle, etc. were collected.

Data preparation: Because of the advancements in the IT industry, high volumes of information are collected from different industries. An IT and database research has offered arise of a way to deal with store and control this important data for decision making in future purposes. Data preprocessing or preparation were identified patterns and key information from these large datasets. This step involves the collection of medical data sets to conduct model training and testing. Data cleansing was involved in duplicate removal, normalization, and error correction.

Model selection: Machine learning involves several models that are available to do both regression and classification tasks. In this step, we select a singular model among the group of ML models for dataset training. Especially in medical diagnosis, selecting the correct model is important because every model was designed to perform different tasks.

Model training: In this step, the chosen model is properly trained to make disease predictions with the highest accuracy. For example, in cancer diagnosis, linear regression algorithms are used to retrieve patient type with malignancy or not.

Model evolution: It is necessary to evaluate the machine-learning algorithm before adoption into the medical domain. After feature reduction or data preprocessing, and model development, we need to evaluate whether a particular model is accurately identified a disease. Different performance metrics are available to assess different machine learning algorithms. The accuracy, area under the curve (AUC) are used for disease classification purposes, and parameters precision, sensitivity, and F1 scores are used for sorting purposes. Figure 1.3 presents the confusion matrix example to define the model performance.

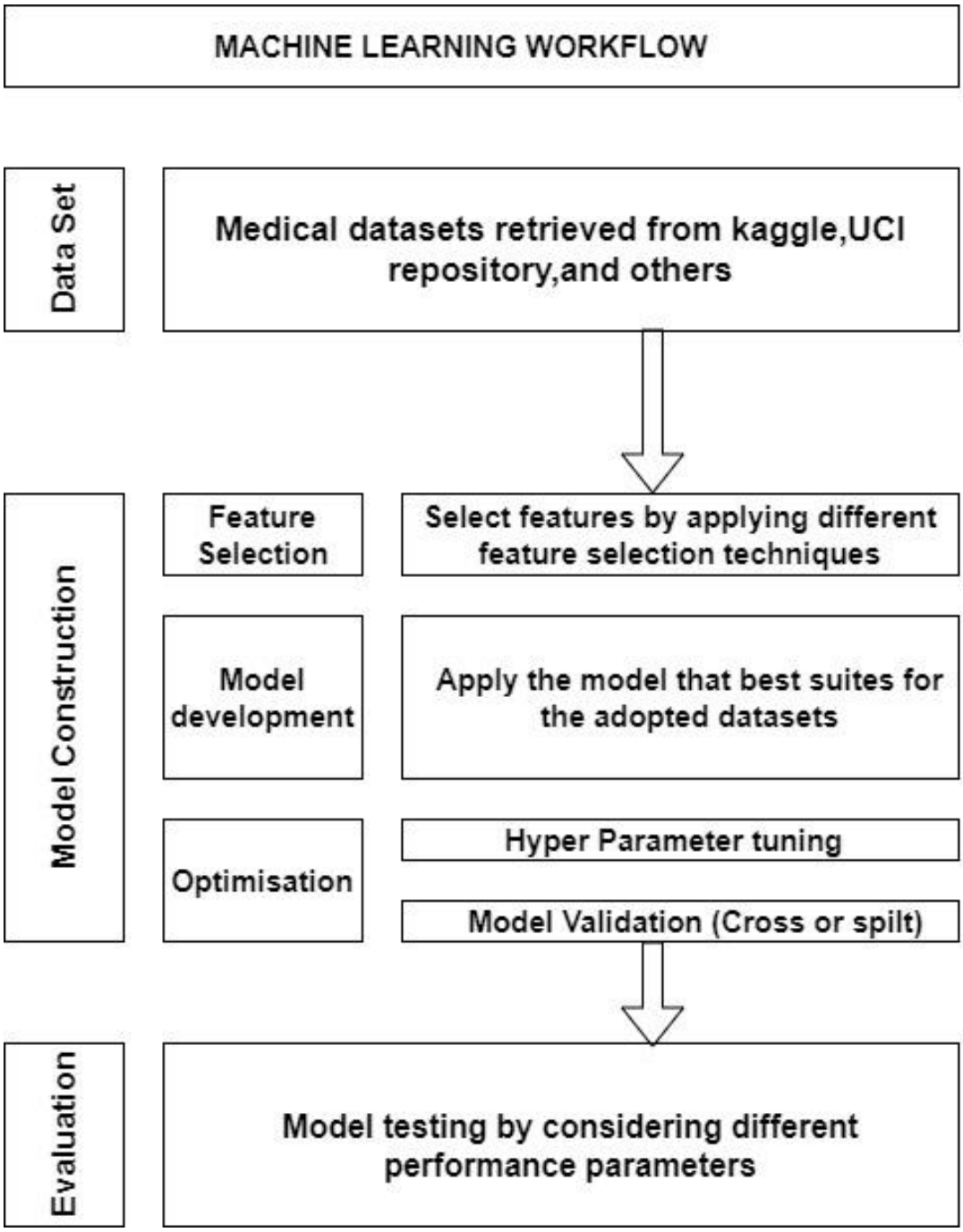

Figure 1.2. Machine learning framework in medical diagnosis.

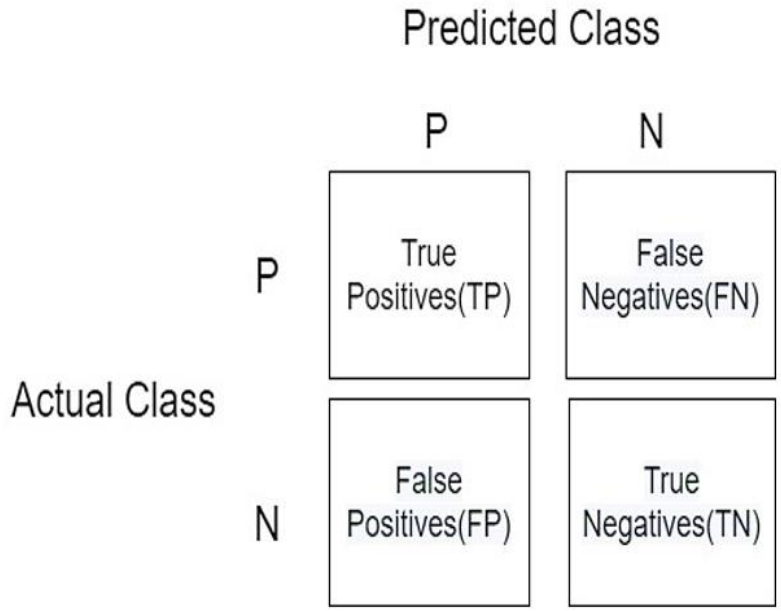

Figure 1.3. Confusion matrix

From the above confusion matrix, we define the following performance metrics.

Accuracy (A): Portion of actual prediction subjects among total subjects;

$$A\ (\%) = \frac{\text{True Positives+True Negatives}}{\text{TP+TN+FP+FN}} * 100$$

Sensitivity (s): True positive subject percentage

$$s\ (\%) = \frac{\text{TP}}{\text{TP+ FN}} * 100$$

Precision (p): Percentage of true positives from total positives

$$p\ (\%) = \frac{\text{TP}}{\text{TP+FP}} * 100$$

F-Score: Weighted harmonic mean for precision and sensitivity

$$\text{F-score} = 2 * \frac{\text{sp}}{\text{s+p}}$$

The area under the curve (AUC): AUC is the visualization tool of multiclass classifier performance and is one of the major evolution parameters to check any classification model performance.

Hyperparameter tuning

Hyperparameter tuning helps to improve model performance. The hyperparameters include distribution and initialization values, training steps, learning rates, etc.

Model validations

Model validation can be done in two ways either spilt or cross-validation (CV). In spilt or hold-out validation is the point at which we split up the given dataset into a 'train' and 'test' set [37]. The preparation set is the thing that the model is prepared on, and the test set is utilized to perceive how well that model performs on unseen information. A typical split when utilizing the hold-out technique is utilizing 80% of the information for preparing and the excess 20% of the information for testing.

CV is another model training technique that can evaluate prediction accuracy. The greatest test in ML is approving the model with trained data. To guarantee the incorporated ML model is producing the noise-free model examples [38], computer researchers utilize CV strategies. When compared to hold-out methods, the CV techniques offer the most straightforwardness in assessing models with low bias and accordingly is one of the most mainstream strategies in ML algorithms. The works presented in this report used plenty of ML modelling approaches including single model or multi modelling techniques. The k-fold CV strategy was utilized to perform model validation. The AD patient datasets were separated into 'k' folds to conduct training with test information, and the leftover 'k-1' folds were joined to frame trained information. Original data were arbitrarily isolated into 'k' folds (k_1, k_2 ..., k_i), and the model testing was performed by 'k' times. For instance, in the main emphasis, on the off chance that subset (k_1) used as test data, at that point the leftover subsets (k_2, ..., k_i) were joined to direct model preparation, and this cycle was repeated for the rest of the 'k' values. Some studies detailed that to overcome issues related to imbalanced data sets, the ideal incentive for 'k' could be 5 or 10. With the most noteworthy (k) values, the distinction in trained and sample datasets would in general secure low values.

Model predictions

By using test data, the model conducts the classification of medical label data, ultimately validates, and better approximates to verify how the model was performed on real-time medical diagnosis data. In the next sections, the authors present the importance of ML, and deep learning techniques in the context of medical diagnosis.

1.4.2. ML algorithms in clinical practice

Health issues caused especially by chronic diseases are the main reason for worldwide medical costs. Individuals that suffered from these diseases required permeant treatments. To do this, ML models were frequently applied in the prediction, classification, and diagnosis of different chronic diseases. In this section, the authors discussed some contribution studies involving ML model applications in the primary diagnosis of some chronic diseases.

The Healthcare industry collects large sources of medical information but is not mined to identify the hidden pattern details to make effective decision-making. Disclosure of hidden patterns and connections regularly goes unexploited. Comprehensive data mining methods can help to overcome this limitation. Data mining targets a set of given information to identify important and possibly useful patterns. Some example techniques like Bayesian models, artificial neural networks, decision trees, genetic algorithms, and associate rule mining are largely utilized to discover patterns or knowledge, which is previously not known.

Initially, deep learning can use unlabeled data during preprocessing; thereafter it is well suited for imbalanced datasets and achieves a knowledge base [39]. At present these are largely involved in all other problems that are not able to address by traditional AI techniques. ANN is the latest deep learning algorithm that is discovered the functionality of different industries. Deep neural networks (DNN) are characterized contributions to profits through a complex composition of layers that presents building blocks including nonlinear functions and transformations. Medical experts feel that deep learning could be a promising solution in disease identification and symptom detection [40]. There is an expectation which these deep learning and DNN can alter the chance of getting medical errors like often getting symptomatic errors.

The deep learning techniques can assist the radiologists who specialized in diagnosing the diseases by MRI, computed tomography (CT) scans, and X-rays [41]. There is also the assumption that deep learning can replace human intelligence within the next five years. Because diagnostic imaging holds in medical diagnosis is usually fit to deep learning models. There are plenty of scenarios that drive the integration of deep learning with several diagnostic practices including radiology like shortage of healthcare workers, hike in medical costs, and large incoming imaging data etc.

At present, these algorithms enhance the workflow of the diagnostic process but do not mean to replace human intelligence.

1.5. Problem statement

The problem statement defined in this work can summarize research motivations in total. This problem statement can be more or like a mission statement or goal, driving all the following efforts rather than typical research issues. The main problem statement of this research is presented as follows

- **Research statement**

In which way we can achieve better AD risk prediction accuracy and identification of dementia stages with comprehensive machine learning algorithms.

To support the above statement, we perform three types of machine learning algorithms such as traditional models (supervised), deep learning models, and ensemble learning models on dementia data subjects.

From the above statement, we have drawn the following research questions (RQ)

- **RQ1**

  How machine learning algorithms can help to identify the risk factors of dementia? Is it possible to assess the prediction accuracy in detecting dementia among old adults?

- **RQ2**

  How far these proposed ML models can validate in real-time dementia studies? Are there any particular tests to calculate the model performance?

- **RQ3**

  How can other ML models like ensemble learning enhance the model prediction accuracy? What type of dementia datasets are imported to execute the machine learning experiments?

- **RQ4**

  How can we justify the performance of these models in subjects with mild AD or younger subjects, also to calculate the prediction accuracy using other biological tests like blood markers or cerebrospinal fluid (CSF)?

1.6. Research objectives and contributors

In this work, we addressed the important applications of machine learning in medicine especially in dementia prediction and classification of different AD subjects. For

supervised machine learning models, we perform the AD predictions with numerical data. For MR brain image data, we applied deep learning type convolutional neural network models (CNN) for dementia type classification. In the end, we also test and compare these models by combining more than two or more models (hybrid or ensemble modelling) to propose the comprehensive machine learning model for identification and classification of dementia and performance calculation has done in terms of ROC values. The main objectives of this work have been divided into three major types.

Research objective -1

- To manage the traditional machine learning models in dementia prediction
- To do the model training with given featured dementia datasets
- To compare the performance of different ML models with dementia subjects
- To find the optimized supervised machine learning model

Research objective - 2

- To apply the deep learning models on MR image brain datasets
- To calculate model performance and compare it with other supervised learning models

Research objective - 3

- To develop models with manual selection of MRI features
- To do modelling done with automatic feature selection, and
- To develop a single model with ensemble learning or hybrid modelling.

RQ1: Traditional machine learning algorithms in dementia prediction

Several models were proposed in the literature for the identification and prediction of dementia in older adults. By exploiting typical ML techniques, it's impossible to analyse this important volume of data due to time consumption and efforts. Therefore, new modelling techniques are developing with numerous rules and programs to avoid these issues. Besides that, the choice of the correct algorithm is not a straightforward task because it depends on multiple factors equivalent to data volume, information type, and outcomes concerning business requirements.

The first contribution of the work is to implement novel modelling techniques using supervised learning algorithms such as support vectors, logistic regression, naïve Bayes, etc. These models are considered to solve the data problems associated with dementia classification and predictive analysis. The outcome models provide a convincing and

adaptable structure for MRI and proposed a classification problem method that has the capability as a system for dementia (AD) classification.

RQ2: Validity of models verified by performance metrics

As mentioned earlier, it's imperative to utilize new information while assessing the presented model to forestall the probability of overfitting the training set. Nonetheless, here and there it's valuable to assess the developed model since we are building it to locate the best boundaries of a model - however, we can't utilize the test set for this assessment, or, more than likely we'll wind up choosing the boundaries that perform best on the test information yet perhaps not the boundaries that sum up best. To assess the model while as yet building and tuning the model, we make a third subset of the data called a validation set. A common train/test/validation split is utilizing 60%, 20%, and 20% of the data for training, testing, and validation respectively. It is also mandatory to mention data shuffling before making these data splits and by that each split can have an accurate notation of the dataset.

RQ3: Ensemble models in dementia detection

Ensemble models or mixed ML models are combining outcomes from numerous models to improve the total model performance. The main cause for occurring errors in any model because of bias, noise and variance. These ensemble models can reduce such types of issues. These are specially designed to amplify the accuracy and stability of ML algorithms. Ensemble modelling can be done in different ways either boosting, bagging or wrapping. The main objective is not only to classify AD diagnosis but also verify the adopting models comprehensively classifying the demented subjects. As mentioned, sensitivity or true positive rate can address how many MR images are accurately associated with the corresponding AD types. It is also important to understand that greater sensitivity of a model did not guarantee better accuracy because usually there is a trade-off between independent performance metrics. To do that selection of ensemble learning type approach is also important.

1.7. Organization of thesis

The overall organization of the thesis is presented in this section.

Chapter 2 presents a brief outline of the literature review on machine learning for different medical problems like cancers, diabetes, hepatic fibrosis, heart attacks etc. It presents the concepts with the ML modelling advantages in the prediction and

classification of diseases. At the end of the chapter, all related literature was framed in Tabular form for better understating of readers.

Chapter 3 demonstrates the AD prediction with single model machine learning techniques. In this work, we attempted to adopt support vector machines (SVM) in the expectation of dementia, and performance validation was done by statistical analysis. Data were explored from the Open Access Series of Imaging Studies (OASIS-2) called longitudinal collection of 150 subjects of 373 MRI information. we consider the attributes like MR delay; CDR, ASF, AGE, and GENDER included with MMSE that corresponds to subject ID. We categorically accept that it is a novel method of inspecting the significance of every parameter during forecasting dementia in elder patients. Regardless of it, this work intends to anticipate dementia in senior people by SVM calculations to achieve promising results.

Chapter 4 also presented the AD prediction with a single supervised ML model called decision trees. Pruned type decision trees (J48) were utilized to do prediction analysis on AD subjects. Validation of the adopted model was done by cross-validation techniques. Model performance was evaluated by parameters like precision, accuracy, and receiver operating characteristic (ROC) curve. Predictions by generated decision tree have been correctly mapped and examined with confidence attributes of dementia status. At last, the high confidence correlated value of attributes can predict dementia in of particular adult, and the referenced model clarifies and forecast the patient's condition by using explicit advantages to help patients by helping them ahead of time.

Chapter 5 advances this study by doing comparative supervised ML techniques by adopting three conventional models named SVM, K-nearest neighbours, and logistic regression. In this work, we introduced a feature reduction technique called principal component analysis to identify the high correlated features to the targeted dementia group attribute. Outcomes approved that the three models are precisely group dementia patients with better accuracy from 96.7-98.3%. We also validate the adopted models with recall, precision and AUC. The AUC of LR and KNN presented an optimal prediction model, with the end goal that these two prescient models were done better classification of the dementia patients.

Chapter 6 presents comprehensive machine learning model development with inclusion of four supervised models namely Naive Bayes (NB), artificial neural networks (ANN), K-nearest neighbour (KNN), and support vector machines (SVM) were presented. The receiver operating characteristic (ROC) curve metric were used to validate the model performance. Each model evaluation was done in three independent experiments. In the

first experiment, a manual feature selection was used for model training. In the second experiment, automatic feature selection was conducted by wrapping methods, and the last experiment consisted of a new approach with mixed modelling called ensemble learning. In this chapter, the significance of joint ML modelling for AD-onset prediction in elderly people has been demonstrated. Besides, hybrid modelling enabled 98% accuracy in predicting AD in older adults. The outcomes suggest that joint modelling, with limited features, is the best practice to assess AD-onset by subject prediction.

Chapter 7 explains the importance of deep learning techniques in the analysis of brain image studies. We considered the Open Access Series of Imaging Studies-3 (OASIS-3) dataset with 2,168 Magnetic Resonance Imaging (MRI) images of patients with very mild to different stages of cognitive decline. We applied deep learning-based convolution neural networks (CNN) which are well-known approaches for diagnosis-based studies. This work presented a deep CNN with 10-fold cross-validation and achieved more than 80% accuracy. While applying computing methods for diagnosis, a small portion of datasets are presented. Therefore, our model maintained a random image selection of train, test, and validation datasets. The proposed model produced promising results in AD image classification. The most notable outcome for this study is the progressions among predictiveness of AD diseases.

Chapter 8 finishes with the conclusion of this thesis work and defines the directions for future work and the possible expansion to the contributions discussed in the thesis work.

Chapter 2

Machine learning in medicine

In this chapter, the background information on the main concepts and ML models involved in the prediction, and classification of chronic diseases. The recent ML evolutions in different medical regions and their advantages, followed by the machine learning framework representation presented in detail. Computer vision technologies that are trending in the medical domain and their different approaches are discussed. Deep learning which is an integral part of the model development is presented with different algorithms in practice. Some supervised ML models and deep learning practices like ANN and CNN were explained which is employed as a prediction of diseases like cancers, diabetes, hepatic fibrosis, etc.

2.1. Introduction

From the earliest starting point, machine learning models are designed to evaluate clinical information. Today these techniques are becoming fundamental tools to do insight analysis of medical data [42]. Particularly, since the last decade, the computerized transformation gave relatively less expensive and available to accumulate medical information. Nowadays, hospitals are largely equipped to collect and monitor data from Information Technology (IT) systems. Moreover, this information is largely gathered from big data frameworks. Given this, machine learning algorithms are better suited to the examination of medical data, and specifically, there is a great deal of work that has been done in clinical analysis especially for diagnostic issues [43].

Medical diagnostics are a classification of medical tests, which intends to identify infectious diseases, conditions, and ailments. These clinical diagnostics fall under the class of in-vitro medical diagnostics (IVD) that can be purchased by end-users or utilized in research Centre settings. Biological samples are disengaged from the human body, for example, blood or tissue to give results. Because of the multiple opportunities for utilization of ML in medical diagnosis, clinical imaging work processes are well on the way to being affected in the near term. ML-driven methods that autonomously process 2D or 3D image scans to recognize clinical signs (like lesions or tumours) or decide possible diagnosis determinations have been distributed and some are advancing through administrative steps thrives the market.

Among ML, deep learning techniques are largely utilized on layered portrayals of different features, such called neural networks. To understand the deep learning techniques powerful image data is required to perform recognition tasks [44]. For instance, if a person entered a dark room and searched for the light switch. From previous experience, the person can figure out how to connect light switches with unsurprising areas inside the design of a room. Numerous computer vision-based picture handling calculations, including deep learning, imitate this behaviour to assess factors, which related to the recognition task that needs to be done. Due to the consideration of the multiple complexities of factors, deep learning has its capacity for image interpretation, especially in clinical practice.

The historical progress associated with ML applications in clinical analyze are shows that easy form and straightforward to manage algorithms, frameworks, and approaches have developed to empower progressed and modern data analysis. Both deep learning and machine learning are largely integrated with data mining techniques [45]. Data mining has significance concerning finding the designs, anticipating, and disclosure of information in various spaces. Data mining algorithms and techniques like clustering, classification makes a difference in finding the patterns to choose what has to come to business structures to develop. It is a wide application area nearly in each industry where the information is produced because data mining is considered as one of the most significant frontiers in database and data frameworks, and one of the most encouraging interdisciplinary improvements in IT management.

A higher collection of medical documents is a valuable resource to retrieve new and valuable knowledge that can be found through data mining. Deep learning and data mining techniques are user-based approaches to identify hidden and novel data patterns. These are highly applicable in identifying key patterns among big datasets. At present, these are highly applied in healthcare systems especially medical diagnoses to predict or classify diseases. Simultaneously, machine learning (ML) can detect and diagnose serious diseases like cancer, dementia, diabetes, etc. Especially deep learning is one application that highly applicable to the healthcare context is digital diagnosis. Besides, it can detect patterns of individual diseases within patient electronic health records (EHR) and produces feedback on anomalies to the doctor. This chapter presented a brief discussion including ML and deep learning approaches in a clinical context, differentiated between structured and unstructured patient data patterns, and provide references to applications of the mentioned methods in medicine. Besides, it also highlights performance measures and evaluations used in the diagnosis prediction and classification process.

2.2. ML algorithms in cancer predictions

Due to the capability of detecting hard to discern patterns from complex and noisy datasets, ML algorithms were often used to detect cancer maladies. Several studies were attempted machine learning techniques for prognostication and risk prediction of different cancers. Alabi R et.al [46] present the uncertainties of relapses in the initial stage of oral tongue squamous cell carcinoma (OTSCC) in decision making of oral lounge cancers. The authors were collected 311 patients' data from five different university hospitals and compared the performance of four-ML algorithms namely Naive Bayes (NB), Boosted Decision Tree (BDT), Support Vector Machine (SVM), and Decision Forest (DF). Preliminary outcomes highlighted that BDT has generated the highest accuracy of 81% and 0.78 value of F-score and SVM generated the lowest accuracy of 68% and 0.63 value of F-score respectively.

Manabu T et.al [47] conducted machine learning algorithms by digital slide images to do the early prediction of colorectal cancer (CRC) metastasis on 397 subjects. A few morphologic boundaries were separated from entire slide pictures of cytokeratin immunohistochemistry images. A random forecast (RF) model was employed by doing data split as a trained dataset of ($n = 277$) images to predict lymph node metastatic also test dataset of ($n = 120$) images. The performance outcomes were further compared to machine learning models and conventional approaches of datasets. Ultimately, lymph node metastatic prediction by ML algorithms was outperformed by other conventional models.

Bikesh Kumar S [48] determines breast cancer biomarkers to conduct predictions by anthropometric and clinical features including ML algorithms. Feature correlation and selection methods were employed to evaluate the correlation between different features. Moreover, famous classifiers, for example, SVM, NB, quadratic discriminant, linear discriminant, K nearest neighbours (KNN), RF, and logistic regression (LR) are introduced for breast cancer predictions. Results highlighted that among glucose, age, and resisting are seen as generally important and viable biomarkers for malignant growth prediction. Further, the KNN classifier accomplishes the highest 92.1% of classification accuracy.

Raghava B et.al [49] did the predictive analysis on total pathological response following Neoadjuvant Chemotherapy to detect breast cancers using ensemble machine learning. Term ensemble learning defines the process of combining multi models. The results are further validated by K-fold cross-validation and generated 99.08% of accuracy. In similar, Leili T et.al [50] did a performance comparison of six ML algorithms to classify survivors of breast cancers and metastasis. Among 550 patients, 85% of them, not experience metastasis, and 83.4% were alive. In a prediction analysis of survival, the

SVM produced the highest 93% accuracy. For the prediction of metastasis, the logistic regression generated the highest 86% of total accuracy.

2.3. Diabetes prediction

Diabetes is a typical continuous medical problem happening when the pancreas has not delivered sufficient insulin. Raised glucose levels are the common consequences of serious diabetes. Thereafter, diabetes will make serious harm to the nerves and veins. Propelled diabetes is problematic by coronary ailment, visual impedance, and kidney disappointment.

As per the World Health Organization (WHO) reports, about 425 million individuals are globally suffered because of diabetes. Some studies are reported that family history, unhealthy diets, hypertension, lack of physical activity, and obesity are risk factors for getting Type 2 diabetes. Women are having a high tendency to get at risk for type 2 diabetes because of a high number of pregnancies, low insulin consumption, and high cholesterol levels [51]. Early recognition of the sickness can offer patients the chance to make the fundamental way of life changes and along these lines can improve their future [52]. To do that computer scientists are recommending cost-effective ML and data mining techniques for the diagnosis of diabetes.

Few investigations lead to expectation examination utilizing ML algorithms to analyze diabetes. Nahla B et.al [53] involved SVM algorithms to detect diabetes mellitus and achieved 94% accuracy. Moreover, Quan Z et.al [54] employed J48 decision trees, random forests, and neural networks. Scholars mentioned RF is an ideal algorithm to produce better accuracy (of 80.4%) in the classification of diabetic patients.

Deepti S et.al [55] proposed a predictive model to estimate the probability of diabetes. Scholars employed NB, J48 decision trees, and SVM algorithms and concluded that NB generated the highest 76.3% of accuracy than others. On the other hand, Battineni G et.al [56] predicted causes for diabetes in Pima Indian female patients by a comparative machine learning study. Plasma glucose concentration was the major cause of diabetes happening in these female groups, which is followed by other risk factors like multiple pregnancies, high insulin release, etc. The study by Chaki J et.al [57] had provided a systematic investigation of AI and ML approaches for self-management of diabetes mellitus and identification.

2.4. Hepatic fibrosis

Hepatic fibrosis addresses the injury fixing reaction to the liver from a wide assortment of etiologies. Cirrhosis is the high exceptional phase of fibrosis, implying more than

fibrosis alone, yet rather a contortion of the liver parenchyma related with septate and knob arrangement, adjusted bloodstream, and the likely advancement of liver disappointment. Real-time tissue elastography (RTE) is one of the contemporary techniques and promising imaging methods since it is both non-obtrusive and gives precise assessments of hepatic fibrosis. It is reported that pattern recognition approaches and machine learning models were largely studied in the early diagnosis of hepatitis diseases especially for clinical figures and biochemical indices.

Yang C et.al [58], utilized four conventional ML models namely NB, KNN, SVM, and random forests to develop the clinical decision system to measure hepatitis B. Eleven RTE image characteristics are retrieved from 513 patients with liver biopsies. The test results indicated that the employed ML models successfully outperformed the LFI (liver fibrosis index) technique and the Random Forest (RF) model produced the most noteworthy normal precision among the four ML algorithms. This outcome recommends that modern ML techniques can be incredible assets for assessing the phase of hepatic fibrosis and produce a guarantee for future medical practices.

Jiang Z et.al [59], was developed a simple model to differentiate patients of clinically significant fibrosis (METAVIR F2-F4) and patients of no or mild fibrosis (METAVIR F0-F1). This study involved 204 community healthcare patients and 34 serum attributes including gender, age, and infection duration that are involved to differentiate fibrosis by the SVM algorithm. Before SVM implementation, feature selection was conducted by a sequential forward floating selection (SFFS) process. Results mentioned that the adopted SVM model to identify patients of clinically substantial fibrosis with 96% accuracy.

Hashem A et.al [60], a study has presented single and multistage classification models to predict the degree of liver fibrosis patients with infection caused by chronic hepatitis C. The studies that are previously reported diagnostic techniques are not successful to predict early-stage fibrosis irrespective of producing higher accuracies. Given this, scholars of this study developed both single-stage and multistage ML model classifiers to predict the degree of liver fibrosis by employing decision trees, neural networks, nearest neighborhood clusters, and logistic regression models. Preliminary results mentioned that the classification accuracy in means of AUC of multi-model ranging from 0.874 to 0.974 represents improved classifier accuracy than other studies.

Chen H et.al [61], presented a sophisticated hybrid model by integrating SVM with local Fisher discriminant analysis (LFDA) for diagnosis of hepatitis. The improved LFDA-SVM algorithm was further compared with the other three conventional methods such as SVM with Fisher discriminant analysis (FDA-SVM), SVM associated principal component

analysis (PCA-SVM), and conventional SVM models. The LFDA-SVM model was outperformed by other models and produced 96.77% of accuracy. It is one of the promising and advanced establishing tools in the diagnosis of hepatitis with great performance.

Stoean R et.al [62] produced a model by support vector machines to determine liver fibrosis in chronic hepatitis C. The model was developed by the comprehension of learning components in SVM and the evolutionary algorithms to do engine optimization. By involving evolutionary techniques, it successfully claims better performance than conventional SVM methods, also confirms the significance of new methodology near to reliable support within the medical diagnosis.

Polat K et.al [63] did predictive analysis by PCA and artificial immune recognize systems. By reducing the feature set to five from 19 with the help of PCA techniques the developed system resulted in 94.12% accuracy. Scholars also mention that this approach can benefit other medical diagnoses and reduce the doctor's mental stress. To sum up the results of the mentioned studies, machine learning models are the best techniques for staging hepatic fibrosis than other statistical calculations.

2.5. Heart attacks

Heart attacks are on the head of the deadly ailments list. They are viewed as a major cause of global deaths. As indicated by the WHO statistics, in 2020, about 17 million deaths were caused because of heart diseases [64]. In America, heart diseases such as sudden strokes, hypertension, and coronary heart diseases are the main causes of death. Only because of coronary heart disease one in seven people lost their lives that resulting in about 366,800 deaths per year in the USA. Approximately, 3% of American adults (7.9 million) are facing the problem of cardiovascular failures. Moreover, a single person dies every 37 seconds from heart attacks [65]. Given this, there is a persevering requirement for an exceptionally precise framework that works as an assessment tool to detect hidden patterns of clinical data of heart diseases and reduce the risk of coronary failures. The ML classification algorithms are recently largely incorporated to diagnose heart diseases.

Desai S et.al [66] was applied logistic regression and back proportion neural networks (BPNN) to predict heart diseases for the Cleveland dataset. The models that were developed were further validated by 10-fold cross-validation. The models are greatly assisting doctors in taking effective decisions in the diagnosis of heart failure. Besides, Ahmad H et.al [67] proposed a tool to find a better ML algorithm that achieves high accuracy for the prediction of heart diseases. The four classification algorithms such as

decision tree, SVM, RF, and logistic regression were employed and experiments were conducted by total features and selective features. Study outcomes mentioned that the RF classifier algorithm was outperformed by others and produced 94.9% of accuracy.

Enriko I et.al [68] develop the heart disease forecasting model by using the K-nearest neighbours (KNN) model with simple patient health features. Authors recommend The KNN with a weighing parameter approach to improve the accuracy in disease diagnosis than others. On the other hand, some authors are proposing hybrid or combined ML algorithms to do a better diagnosis of heart diseases. The hybrid modelling involved two phases. At first, feature selection is conducted and later selected features portion is involved in the development of classification models [69]. Similarly, Maji S et.al [70] were employed these models by integrating decision trees with artificial neural networks (ANN) to produce better performance in heart disease diagnosis.

Also, Nguyen T et.al [71] was applied these algorithms by combining fuzzy models with genetic algorithms for the Cleveland heart disease dataset. The outcomes demonstrated that the genetic algorithms integrated with the fuzzy model are generated better results than other single models like SVM, neural networks, etc. Besides, Manogran et.al [72] also present a system that contains neuro fuzzy and multiple kernel approach inference to do a diagnosis of heart disease. In further, the system was tested by the dataset of metabolic reactions and achieved 98% of sensitivity.

Nazari et.al [73] developed a model by integrating fuzzy inference and the fuzzy analytic hierarchy process (FAHP). Scholars presented a data set of a Tehran city hospital for system training and testing. The fuzzy inference has been used to evaluate the possibility to expose heart problems for individuals and the FAHP for feature weight calculations that contributed to the development of heart attacks.

2.6. Asthma or chronic obstructive pulmonary diseases

Chronic obstructive pulmonary disease (COPD) is a type of lung disease caused by expanding shortness of breath. These diseases are highly involved with morbidity and mortality and are the third driving reason for death overall in the USA, and China [74]. The early diagnosis of COPD patients with future high costs could reduce the medical expenses by exacerbation events, and reduce disease evolution. Some studies addressed the ML algorithm to diagnose future high-cost COPD individuals.

For example, Shaochong L et.al [75] incorporated smooth Bayesian network (SBN) algorithms to predict the COPD patients involved with future high costs. The developed SBN model aims not only to obtain high prediction accuracy also sophisticated

generalizability than other benchmark ML algorithms. In similar, Peter J et.al [76] presents ML characterization of COPD subtypes by insight analysis from gene study. The longitudinal characterization of COPD gene subjects has provided the relationship between lung image characteristics, COPD progression, and molecular markers.

Moreover, Tadahiro G et.al [77] compared the performance measures of different ML algorithms to predict hospitalization and critical care among emergency patients with COPD exacerbation. Outcomes ML algorithms enhanced the prediction ability of patients diagnosed with Asthma or COPD exacerbation. Alternatively, Maarten L et.al [78] present a tool to identify functional respiratory imaging features related to COPD and disease forecasting by ML algorithms in early understanding and quantification of disease progression. The FRI features such as total specific image-based airway resistance and volume integrated with SVM algorithms generated 80.65% accuracy and 82.35% of sensitivity.

Sandeep B et.al [79] proposed an ML-based framework to evaluate COPD severity. High correlated features were selected by linear forward feature selection and KNN was used as a classification algorithm. Results mentioned that the biomechanical feature set outperforms with 0.81 of AUC than density (AUC =0.71) and texture (AUC =0.73) based feature sets. This study provides evidence of the effectiveness of biomechanical features in the severity and presence of COPD. Likewise, Jianfei Z et.al [80] also applied a feature weighted survival ML model for the prediction of COPD failures.

2.7. Kidney injuries

Intense kidney injury is a typical clinical disorder emphatically connected with an abundance of dismalness and mortality. Patients who develop it are at constant risk for delayed and increasingly costly hospitalization, chronic kidney ailment and dialysis, vital adverse cardiovascular situations, and death. Currently, diagnosing kidney failures is made as per the rise of serum creatinine (sCr) focus or decrease in urinary yield, both are of aberrant markers of renal capacity and maybe longer days behind the beginning of injury and practical decline. These limitations add to underdiagnosis and identify patients with high risks for kidney injury, especially in an emergency condition. Besides, integrating early identification and risk of kidney injury stratification by really engaged support for clinical decisions.

Evaluating the glomerular filtration rate (GFR) is a key parameter to identify initial resistance of kidney functionality, assessing dynamic kidney disintegration and intricacies, altering the measurements of medications, and controlling the risks for chronic kidney diseases. To improve the precision of GFR, ensemble learning methods

were applied [81]. Ensemble learning is a type of ML algorithm that results by combining individual mathematical models to generate better outcomes. Liu X et.al [81] employed ensemble learning models and conduct experiments on 1419 individuals. The independent evaluation of GFR was conducted from sex, age, and serum creatinine with help of SVM, ANN, regression modelling, and ensemble learning. Results mention that the precision of ensemble learning dominates the normal regression models.

ML algorithms were also employed to forecast severe kidney injuries after aortic arch surgeries. For example, Guiyu L et.al [82] compared different ML algorithms with conventional logistic regression to predict acute kidney injury (AKI) after arch operations. Results mentioned that the Gradient boosting algorithm was comparatively produced a better performance than SVM, logistic regression, and random forest algorithms. Therefore, in this section, we discussed how ML algorithms were applied in the studies of kidney injuries which particularly focused on understanding the association between phenotype and genotype.

2.8. ML and deep learning in the AD prediction

Machine learning is an ideal decision of investigation for examining a huge scope of administrative medical data containing a large number of descriptors from countless people. Studies show fruitful uses of AI and ML models to the enormous scope of regulatory information in disease forecasting other than AD (cancers, diabetes, metabolic disorder, hepatic fibrosis, heart attacks, etc). Given the ongoing fast development among ML modelling, the use of AI technologies in medical forecasting modelling is probably going to deeply affect medicine [83], [84].

Despite this, Based on the global Alzheimer's disease (AD) report [85], an individual exposed to dementia will be born every three seconds around the globe. At present, nearly 55 million global population is suffering from this disease. These numbers could be tripled by 2050 and reach 152 million population will be suffered from dementia. AD largely occurs in older people and has a great impact on their daily lives. By early identification of this issue will save medical costs and provide a healthy individual lifestyle. Luckily, because increasing numbers in ML algorithms are well positioned to control this disease at an early phase.

For instance, by collecting speech data from older dementia patients it is possible to identify AD by incorporating ML methods. For this logistic regression with cross-validation, algorithms are the optimal solution to predict early dementia [86]. A comprehensive ML algorithm was developed in [87] by combining four ML algorithms (SVM, NB, KNN, and ANN) to achieve higher accuracy for the early diagnosis of AD. The

outcomes of the developed algorithm produced 99.1% accuracy during the identification of true AD subjects. In similar, Battineni G et.al [88], [89] successfully presented a study on the prediction of dementia and performance validation by exploring the use of SVM and decision tree algorithms.

All the above-mentioned studies involved model development by medical data retrieved from IT systems and can predict only a single outcome. In contrast, Charles K et.al [90] adopted an unsupervised ML approach namely a Conditional Restricted Boltzmann Machine (CRBM) to evaluate AD development. The authors presented 18-month projections on 44 clinical characteristics after 1909 AD patients. The presented unsupervised model was accurately predicted the alternations of the AD cognitive scale and identified subcomponents with better sensitivity. Javier D et.al [91] presented a sophisticated approach to predicting late-onset AD and experimental models perform about 72% of classification accuracy.

Besides, Ramon C et.al [92] were analyzed the possibility of assessing an anatomical index which is known as AD risk factors. The neuroimaging databases were collected, further applied high dimensional ML algorithms and results identified cognitive status, age, and cognitive function are the main factors to support the risk of getting AD. In this manner, there are plenty of other studies like [93], [94] were explained the importance of ML models to produce promising results to estimate the genetic risk of AD. However, a survey on ML algorithms for longitudinal analysis of brain image data of AD was well presented in [95].

As per conducted research, there are several potential indicators for dementia prediction, which can commonly be arranged dependent on the accompanying types of models namely health-based models, neuropsychological models, wellbeing-based models, genetic risk scores, and multifactorial models [96]. The appropriateness of these models spreads in numerous ways [97], [98]. The coupling of multiple neural networks with magnetic resonance imaging (MRI) has been utilized to isolate healthy cerebrums from progressive Mild Cognitive Impairment (p-MCI), because of the primary decay of the cerebrum due to Alzheimer's disease [99]. Routine fundamental consideration patient records in the UK have been and are right now used to build up a risk score for the reasons of assessing how in danger an individual might be of creating dementia, by utilizing ordinary measurable techniques and current AI calculations [100].

Some researchers employed random forest (RF) algorithms for the identification of patients with p-MCI and sustainable age-related MCI through the regional examination of the protein amyloid-β and positron emission tomography (PET) [101]. In an ongoing EMIF-AD study, an ML model on extreme gradient boosting (XG-Boost), deep learning,

and RF algorithm, has been implemented for Alzheimer's based dementia determination utilizing metabolites in the blood which were demonstrated by the investigation to be as precise indicators as the generally acknowledged yet intrusive to quantify cerebrospinal fluid (CSF) biomarkers [97].

Besides, machine learning recognizes compound patterns in high dimensional information, which are then used to make clinical forecasts in new datasets. In dementia, the focal point of ML models has been cognitive centric and neuroimaging [102]. Generated imaging biomarkers of AD incorporate hippocampal and entorhinal cortex volume decrease, and, basal forebrain cores misfortune. Furthermore, examples of decay in regions including the paralimbic zones, parietal affiliation zones, lateral temporal, temporoparietal, and frontal cortices have been recognized as imaging markers of AD-related neurodegeneration that can be seen in people earlier to the beginning of clinical signs [103]. In more precisely, utilized ML to distinguish morphological anomalies in the hippocampus, entorhinal cortex, basal ganglia, praecuneus, and the cerebellum as significant in pre-clinical periods of AD. Progressively, non-psychological signs and side effects are viewed as likely early markers of neurodegenerative sickness [104].

On other hand, the published studies that define AD risk prediction through machine learning models have shown limited prediction accuracy in terms of area under curves ranging from 0.68 to 0.78 [105]–[107]. These studies were used medical data for model evaluation which limits generalizability in other scenarios. Besides, previous studies did not address a primary assumption of supervised machine learning approaches -are the data labels, right? For AD, we realize that the principles used to do disease labels are mistake inclined due to under coding, supplier variety, and different components. This brings up various issues in machine learning usage.

This turns up new thoughts of machine learning to improve accuracy even doing non-ideal symptomatic practices if it utilizes the raw labels. Because of this, we extended our previous research works for large-scale datasets as well as more powerful machine learning prediction models.

2.9. Others

This section presents the discussion on machine learning intervention for other chronic diseases. The study conducted by Finkelstein J et.al [108] defines the ML algorithms to personalize an early diagnosis of asthma exacerbations. Patient telemonitoring brings about an accumulation of huge measures of data about patient illness direction. The authors of this study present comprehensive approaches in the utilization of

telemonitoring information for building machine learning algorithms that can forecast asthma exacerbations before they happen. Experiments were conducted by using adaptive Bayesian networks, Bayesian classifier, and SVM algorithms, and performance in terms of accuracy was achieved 77%, 100%, and 80% respectively. It proves that ML models have the capacity to developing personalized clinical decision support systems.

At the same time, depression is among the main sources of mental illness in developed countries. To adequately target intercessions for patients in danger for a more awful long-haul clinical result, there is a need to distinguish indicators of chronicity and remission at early stages. Because of this, Dinga R et.al [109] presented predictive values on a wide range of clinical, biological, and psychological factors in predicting depression causes. They adopted a penalized logistic regression algorithm and archives 66% of prediction accuracy while the diagnosis of course depression.

Besides, muscle pain-related diseases called Fibromyalgia (FM) are viewed as a constant, musculoskeletal agony state of clinical unpredictability that presumably emerges from dysfunction for central pain preparing pathways. It is based caused by sleep disturbances, depression or anxiety, and fatigue. But FM diagnosing remains challenging for medical experts because neither lab tests nor imaging techniques are not available which can medically confirm or identify the FM diagnosis. To do this, Fred D et.al [110] analyses to characterize and differentiate the FM patient's classes from chronic pain patients. In oppose to other established studies that were associated with classification, rather this study included clustering techniques to categorize the pain and symptom severity.

Lastly, Periodontitis is an oral disease type driven by deregulated aggravation initiated by polymicrobial networks that structure on subgingival tooth sites. The periodontal pocket and gingival sulcus form unique natural specialities for microbial colonization and the subgingival microbiota drives the provocative procedure that prompts periodontal tissue destruction. These infection-related diseases were differentiated between chronic and aggressive periodontitis of microbial profiles conducted by support vectors was well explained in [111]. The authors highlight the use of SVM algorithms in the prediction and diagnosis of periodontitis. To sum up, it is important to present the precise ML algorithms or methods is the most important part to make precise decisions in medical diagnosis [112].

Chapter 3

Performance calculation of dementia prediction by support vector machines (SVM)

ML is considered as one of the contemporary methodologies in foreseeing, distinguishing, and making choices without having a human association. It is rapidly advancing in the clinical business going from finding to the perception of diseases and the investigation of illness transmission. ML models were created to recognize the issues in MRI settings. In this chapter, we presented the utilization of support vector machine (SVM) in the expectation of dementia, and performance validation was done by statistical analysis. Data were explored from the Open Access Series of Imaging Studies (OASIS-2) called longitudinal collection of 150 subjects of 373 MRI information.

3.1. Introduction

These days, ML models are logically used in neuroimaging examines like forecasting of AD from an auxiliary MRI. Additionally, numerous investigations attempted distinctive ML procedures in foreseeing AD and its causes [113], [114]. In the investigation of AD forecast and recovery, a multistage classifier using ML, including the Naive Bayes classifier, SVM, and K-closest neighbour (KNN) was utilized to assemble Alzheimer's sickness more adequately and successfully [115]. Likewise, an investigation from Ref. [116], inferring that the usage of locally linear embedding (LLE) sort of unsupervised learning was used to arrange AD dependent on central MRI information. Furthermore, some primary investigations with ML procedures presumed that these strategies are legitimate and achieve with high accuracy (up to 98%) in diagnosing clinical occasions with an examination of patient clinical records [117].

Regardless of it, AD is one of common in dementia and related generally with elder people [22]. In this chapter, we disclose how to anticipate dementia and compute execution by utilizing support vectors. Commonly, SVM's are considered supervised type ML models, which addresses the information issues identified with regression and classification problems [118]. An SVMs give a convincing and versatile structure for MRI, and that the proposed classifier insight strategy has potential as a framework for the appraisal of characterization solutions [119]. Also, this is used to sort dementia subjects and is like the exploration that utilization a uniform calculation to separate three Primary progressive aphasia (PPA) subtypes in anticipating PPA [120]. Recognizing early

morphological changes in the brain and making introductory findings is efficient for dementia.

High-intensity MRI information can be used to support forecast and finding dementia disease [120]. To do this, we propose to locate an ideal arrangement by exploring different radial basis function (RBF) kernels in the SVM. The proposed technique for estimation is inspired by another methodology of utilizing an ensemble SVM for dementia grouping [121], utilizing MRI information and mini-mental state examination (MMSE).

In opposite, we consider the attributes like MR delay; CDR, ASF, AGE, and GENDER included with MMSE that corresponds to subject ID. We categorically accept that it is a novel method of inspecting the significance of every parameter during forecasting dementia in elder patients. Regardless of it, this work intends to anticipate dementia in senior people by SVM calculations to achieve promising results.

3.2. ML model and formulas

Support vector machines

SVM is a discriminative classifier formally characterized by an isolated hyperplane. The output of this algorithm is an optimal hyperplane that classifies new examples and cases, which supports hyperplane called support vectors [122]. In two-dimensional (2D) space, this hyperplane is a line isolating into two sections wherein each class lay on either side. Moreover, these algorithms are considered as part of supervised machine learning which is helpful in classification and regression analysis. For instance, data classification has been done based on the multiple lines that have ready to propose an optimal solution. However, choosing optimized hyperplane is a hard job since it should not be noise sensitive and generalize them correctly. Therefore SVM algorithm depends on the forecasting of optimized hyperplane (Figure 3.1) that gives the most considerable minimum distance to the trained data set and also helps to achieve a good margin value (i.e. Here margin is a separate line to the nearest class points) [123].

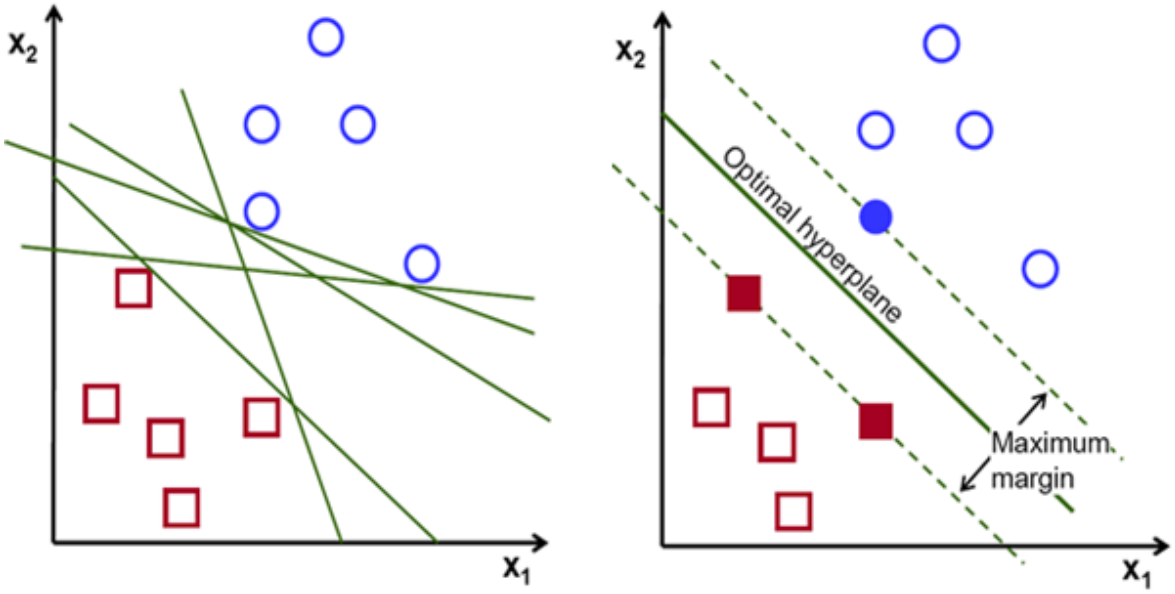

Figure 3.1. data classification using multiple lines (left) and data classification: optimal hyperplane (right)

Tuning parameters

To understand the working of SVM, it is essential to know about some key metrics named as the kernel, regularization, and gamma.

Kernel

In machine learning, "kernel" is a technique used to solve the non-linear problem with the use of the linear classifier. It is involved in changing linearly non-separable data into linearly separable data [124]. The idea behind the Kernel is linearly non-separated data in <n>dimensional space may linearly spread in high dimensional space.

In a mathematical notion, a kernel is defined as K (a, b) = <f (a), f (b)>, Where K is a kernel function, and a, b are n-dimensional inputs, and 'f' is mapping from n-dimensional to m-dimensional space (i.e., generally m should be greater than n).

Regularization

The regularization parameter (represented by 'C' in the python library) explains the SVM optimization, how much we want to escape the misclassifying of every trained data set. For higher values of 'C,' the hyperplane will classify all the training data correctly; similarly, for low 'C,' optimizer looks for greater margin separating hyperplane though it will misclassify the more data points [125]. In precise,

1. If C is high, optimization tends to better value and cover all the characteristics, and
2. If C is low, the optimizer tends large margin, and it would not include more points.

Gamma

Gamma describes the impact a particular training data has reached [122], [124], [125]. By having high gamma values points close to a possible separation line are found in the calculation and for low gamma values (Figure 3.2), points far away from feasible separation line will also taking into consideration in the calculation of separation line [126].

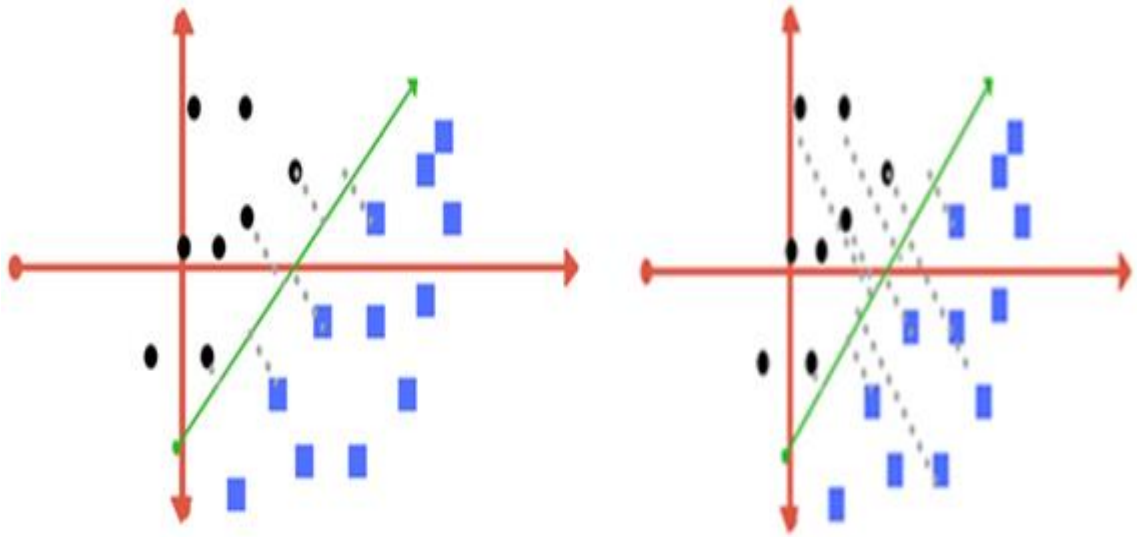

Figure 3.2. High Gamma Close points will be considered (left) and Low Gamma Far away points are found (right)

3.3. OASIS-2 Dataset

Subjects

The dataset consists of a longitudinal collection of 373 MR image sessions performed over 150 subjects. All subjects are right-handed includes both men and women (age range- 60 to 96 years) [127]. 72 subjects are characterized as non-demented (ND) throughout the study, 64 subjects are characterized as demented (D) including 51 are with mild to moderate AD. The remaining 14 subjects are categorized as non-demented at the first scan and were subsequently transferred to demented on later visits called converted (C) type.

Each subject had undergone a full screening of complete clinical assessment conducted at Washington University Alzheimer's disease research Centre (ADRC). Subjects with age typical brain alterations like leukoaraiosis, mild atrophy, and primary dementia cause of AD are included. Collection of MR images conducted within one year before or succeeding clinical assessment (Range- 0 to 352 days, Mean-111 days). Twelve AD confirmed subjects were scanned after a longer delay (Range- 374 to 924 days, Mean-653 days) but were included because of having CDR scores higher than zero in some old clinical assessments. Subsequently, two ND subjects were scanned somehow longer than one year before clinical assessment (Range- 392 to 431 days) also included because their successive clinical assessments continue to have no symptoms of dementia. In this manner, each subject had gone through two or more individual scanning occasions with a mean delay of 719 days (Range: 183-1707 days) between visits. The summary of the demographic status of subjects is presented in Table 3.1. Diagnostic subject characteristics of different age groups on the first clinical visit are explained in Table 3.2.

Table 3.1. Summary of demographic status of subjects

Subject count	78 D	72 ND
Male	40 D	22 ND
Female	38 D	50 ND
Age range (in years)	60-96	
Median	77.0	
Mean±SD	77.01±7.3	

Table 3.2. Age and diagnostic subject characteristics on the first clinical visit [127]

Subjects		**Non Demented**					**Demented**				
Age	N	n	Mean	Male	Female	Convert	n	Mean	Male	Female	CDR 0.5/1
60s	34	23	65.71	6	17	3	11	65.67	8	3	8/3
70s	71	35	74.91	11	24	4	36	73.97	20	16	29/7
80s	41	26	84.30	9	17	7	15	82.33	7	8	13/2
90s	4	2	92.50	0	2	0	2	93.00	1	1	1/1
Total	**150**	**86**	**75.82**	**26**	**59**	**14**	**64**	**74.95**	**36**	**29**	**52/13**

MRI acquisition methods

For individual subjects, three or four separated T1-weighted MRI scans were acquired with a 1.5T Siemens Vision MRI scanner. A high-resolution Magnetization Prepared Rapid Acquired Gradient Echo (MP-RAGE) was managed to handle the classification of given subject scans. For each subject, separate scan files were converted from Siemens proprietary IMA to 16-bit NiFTI1 format by traditional conversion program. The MR images were corrected for inter-scan head rotation and wrapped spatially into atlas space. However, the transformation outcome places the brains in a correlated coordinate system and bounding box as the actual atlas. Therefore, the outcome is unique, high contrast, averaged MP-RAGE image in atlas space was been achieved. The comprehensive description of image acquisition and post-processing steps is well described in [127].

The estimated total intracranial volume (e-TIV) was manually estimated intracranial volume of an atlas. Normalized Whole-brain Volume (n-WBV) was computed with the FAST program in the FSL software suite. Image segmentation was done to classify brain tissue as spinal fluid, white, or grey matter. This segmentation process iteratively assigned voxels to tissue classes based on high probability estimates of the hidden Markov random field model. Ultimately, n-WBV is measured as a proportion of accumulated voxels across the brain mask classified as tissue. The normalized volume is expressed in the percentage of total grey and white matter voxels across e-TIV [128]. The calculation of atrophy rates was estimated as the slope of the line connects to n-WBV. The details of MRI acquisition are mentioned in Table 3.3.

Table 3.3. MRI acquisition details [127]

MR characteristics	Values
Sequence	MP-Rage
TR (repetition time)	9.7 msec
TE (echo time)	4.0 msec
Flip angle	10°
TI	20 msec
TD	200 msec
Orientation	Sagittal
Thickness	1.25 mm
Gap	0 mm
Slice number	128
Resolution	256 × 256 (1 × 1 mm)

Feature description

The dataset includes 373 MRI information and 15 features (attributes). The detailed description of the features is listed in Table 3.4. In the dataset, the subject 'Group' attribute that specifies subject dementia status is considered as a target variable of binary classifier. In this study, we used distinct scoring rules such as CDR, mini-mental state examination (MMSE), SES to determine brain state (Table 3.5). All subjects underwent similar analysis procedures and tests including MMSE.

Table 3.4. Dataset feature description

Features	Description
Subject ID	Subject identification number
MRI ID	Image identification number of an individual subject
Visit	Number of subject visits
Gender	Male/Female
Hand	Right/Left-handed
EDUC	Subject education level (in years)
SES	Socio-economic status
MMSE	Mini-mental state examination score
CDR	Clinical dementia ratio score
e-TIV	Estimated total intracranial volume result
n-WBV	normalized whole brain volume result
ASF	Atlas scaling factor
Age	Subject age while scanning
Group	Demented/Non-demented/Converted
MR delay	Magnetic Resonance (MR) delay is the delay time that is before the image procurement is performed in real.

Table 3.5. Scoring rules

Features	Range	Condition
CDR	0-3	None-0, Very mild-0.5, Mild-1, Moderate-2, Extreme -3
MMSE	1-30	Extreme impairment (<10) Moderate dementia (10-19) Early-stage Alzheimer's aliment (19-24) Normal (>25)
Visit	0 or 1	Low status - 0 High status - 1

3.4. Methods

In the current study, we consider a longitudinal collection of MR image data from the OASIS training data set [127] containing demented and non-demented cases with ages from 60 to 96. Sample size (N=150) subjects are attended scanning sessions for two or more visits, and sessions were separated for at least one year with a total of 373 MR sessions. Figure 3.3. presents MRI image classification based on clinical dementia ratio (CDR) score that ranges from 0-2 and a total of 353 sessions were distributed based on three classifiers demented (146), non-demented (190), and converted (37) are assessed.

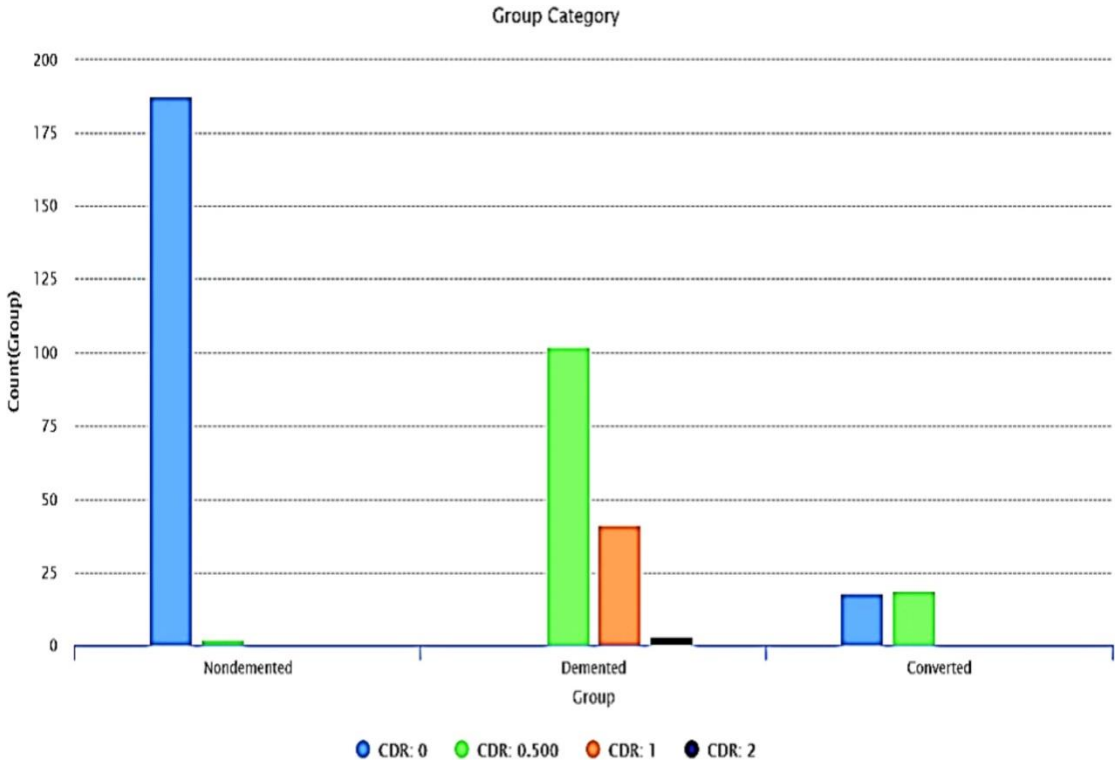

Figure 3.3. Categorization of dementia sessions based clinical dementia ratio (CDR).

We have applied the SVM algorithm to the given dataset using the rapid miner tool, and grouping has been done based on present dementia status. The attributes included in the OASIS longitudinal training data set were Subject ID, MR Image ID, Group, Visit number, MR delay, SEX, age, Mini-Mental State Examination (MMSE) [129], and clinical dementia ratio (CDR) [130]. The proposed method layout is shown in Figure 3.4.

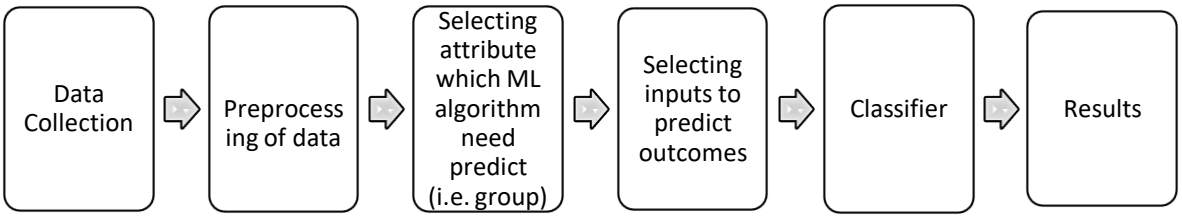

Figure 3.4. Methodology Layout.

Data collection

The trained data set was collected from the Open Access Series of Imaging Studies (OASIS) including longitudinal MRI data of 150 subjects.

Data preprocessing

The real-world data is available more likely incomplete with missing of some entries. So, data preprocessing is a data mining technique to address these issues. Missing entries were filled-up by averaging particular attribute values.

Attribute selection

It is mandatory to choose a particular characteristic to predict the outcome by mapping corresponding input values. In this study status of the dementia category selected as the target variable

Input matching

To avoid vagueness, it is essential to select valuable input attributes to match similar outcomes. Consequently, we choose particular attributes (CDR, MMSE, and MR Delay) that are closely related to deciding the dementia outcome.

Classifier

Three classifiers are named demented, non-demented, and converted. In this particular training data set, some subjects considered as demented at the initial visit later transformed into the non-demented case, and vice versa were called as a converted type.

Eventually, the classification accuracy is achieved and analyzed. Accuracy is defined as the percentage of correctly predicted outcomes divided by the total number of samples.

$$\text{Accuracy} = \frac{\text{True predicted an outcomes}}{\text{Total number of samples}} \times 100$$

3.5. Results

When the mapping has done by input features (independent features) with a target variable, the machine will run the SVM calculation consequently.

Kernel

The outcome kernel model with 150 support vectors has been generated, and three different categories of training data set are observed. As discussed earlier, kernel equations are written as

$$\mathbf{K(ND, D, C)} = < \mathbf{f(79)}, \mathbf{f(50)}, \mathbf{f(21)} >$$

Where K is kernel function with three-dimensional input values such as non-demented (ND), Demented (D), and converted (C) with corresponding mapping values 79, 50, and 21(Table 1) respectively. And bias value is equal to -0.267.

*Bias (offset) generally used for compensating feature vectors that are not centred around the zero (Refer Table 3.6)

Table 3.6. Kernel Model with SVM algorithm

Total number of Support Vectors: 155; Bias (offset): -0.267	
Number of support vectors for class Non-demented (ND)	79
number of support vectors for class Demented (D)	50
number of support vectors for class Converted	21

Gamma vs C

Gamma and regularization (C) values are considered as essential elements to decide optimal hyperplane, and radial basis function or RBF kernel is a different kernel approach function in SVM algorithms. Figure 3.5 represents the spatial distribution of Gamma (RBF) and C values, and Table 3.7 is presenting the total possible SVM optimal parameters along with the performance value. As discussed earlier, low gamma values cover far points$(1 \times 10\hat{} - 4)$, and if C is large, (C = 100) optimization should be better and cover all the aspects. Therefore, SVM anticipated $\text{Optimal Parameters}(\text{Gamma}, \text{C}) = (1.0\text{E} - 4, 100)$ is producing high-performance values and is reflected by the brown circle.

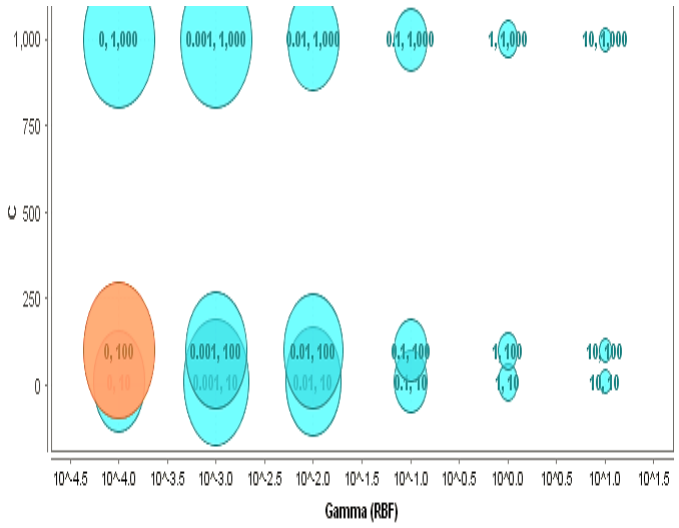

Figure 3.5. Spatial distribution of Gamma (RBF) Vs C values

Table 3.7. Performance parameters

Gamma (RBF)	C	Performance
1.0E-4	10.0	0.6294642857142857
1.0E-4	100.0	0.6875
1.0E-4	1000.0	0.6875
0.0010000000000000002	10.0	0.6696428571428571
0.0010000000000000002	100.0	0.65625
0.0010000000000000002	1000.0	0.6875
0.010000000000000004	10.0	0.6428571428571429
0.010000000000000004	100.0	0.6517857142857143
0.010000000000000004	1000.0	0.6294642857142857
0.09999999999999998	10.0	0.5714285714285714
0.09999999999999998	100.0	0.5714285714285714
0.09999999999999998	1000.0	0.5714285714285714
1.0000000000000007	10.0	0.53125
1.0000000000000007	100.0	0.53125
1.0000000000000007	1000.0	0.53125
10.0	10.0	0.5089285714285714
10.0	100.0	0.5089285714285714
10.0	1000.0	0.5089285714285714

An outcome performance of the SVM is calculated as

$$\text{Accuracy} = \frac{\text{True Predicted Subjects}}{\text{Total Subjects}} * 100 = \frac{105}{150} * 100 = 70\%$$

which is almost equal to the SVM generated values. However, the optimal performance of the system is achieved as 68.75% by utilizing Gamma (RBF) and C values (Refer to Table 2). Precision (i.e., percentage of positive prediction) values for each category are calculated by

$$\text{Precision} = \frac{\text{Positively predicted subjects}}{\text{Total subjects of a particular category}} \times 100$$

For example, precision for Non-Demented subjects is achieved as 64.18% and eventually, demented precision is calculated 75%; surprisingly there are no true predicted values for converted category subjects (Table 3.8.). SVM algorithm predicted two subjects as a converted category, but actually, it belongs to the non-demented type. In conclusion, no ML algorithm produces 100% accurate results since each one has its pros and cons.

Table 3.8. SVM performance matrix

Type	True Non-demented	True Demented	True Converted	Precision
Prediction of Non-demented	43	14	10	64.18%
Prediction of demented	8	27	1	75.00%
Prediction of converted	2	0	0	0.00%

Subject group classifications are in line with the study designed for investigating diagnostic agreements [131]. However, it is not feasible to predict dementia disease with a single attribute or parameter. Thus, we examine other key parameters such as MMSE, AGE, n WBV, and MR delay that matched with the targeted group column. At the same time, we tried to exclude other demographic values like Gender, SES, EDU, and ASF since these parameters are not good enough in dementia prediction, also by considering many attributes performance may get low [132]. Besides, outcomes mentioned that 100 subjects (refer to Figure 3.6.) are predicted non-demented (actually these distributed as 63ND, 24D, and 13C types), and 47 subjects predicted as demented (but these distributed as 11ND, 34D, and 2C types). Finally, 3 non-demented subjects were forecasted as converted types.

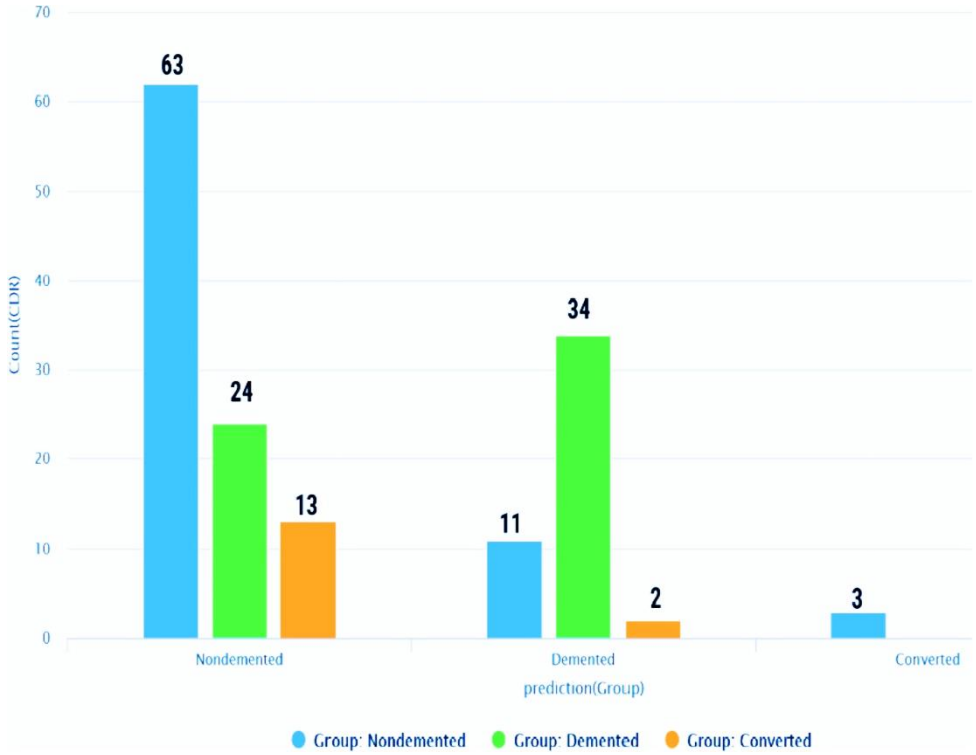

Figure 3.6. Subject Classification between Predictions subject groups Vs Actual Subject Groups.

3.6. Chapter summary

Dementia is one of the huge medical problems that has tested medical specialists around the world. Moreover, it generally occurred in more adults with age more than 60. Tragically, there are no appropriate medications to fix this sickness, and in some cases, it will straightforwardly influence an individual's memory aptitudes and lessen the human capacity to perform everyday exercises. Numerous medical services experts and computing researchers were performing research exercises on this issue throughout the previous twenty years. In any case, there is an extraordinary requirement for the distinguishing proof of significant qualities that can gauge the identification of dementia. In this chapter, single ML modelling was presented with help of SVM for grouping and prediction of AD.

The results summarized in this chapter were fully published in **Battineni G, Chintalapudi N, Amenta F. Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). Informatics Med Unlocked. 2019;16:100200. doi:10.1016/J.IMU.2019.100200.**

Chapter 4

Late-Life AD detection using pruned decision trees

As presented in Chapter 3, we did a similar type of work using single machine learning modelling with decision trees. Pruned type decision trees (J48) were utilized to do prediction analysis on AD subjects. Validation of the adopted model was done by cross-validation techniques. Model performance was evaluated by parameters like precision, accuracy, and receiver operating characteristic (ROC) curve.

4.1. Introduction

There are some settled plans and propositions for clinical practice on some outside assessments and hard-coded into their product. But these projects are limited to information accuracy because they are created from various individuals and conditions. Besides, dementia is one of the worldwide clinical issues that were highly popular. The majority of the examinations are identified with dementia causes clarifying the prevention of risk, early medication, and quick disease diagnosis in more established grown-ups. Hence, it is obligatory to direct some high-level examinations managing these sicknesses.

As a rule, subjects with Mild Cognitive Impairment (MCI) are significant gatherings for the AD cure as they are at the prodromal stages and a higher danger of disease. AD and various types of dementia were turning into a worldwide test and keeping an eye on the death of one in three senior people groups in the USA. While the explanations behind these infections have not yet been comprehended, they can successfully influence talk, memory, and other basic mental capacities.

The present chapter concentrates on AD detection among older adults through managing MRI demographic information and AD forecasting evaluated with given features. Pruned decision trees (J48) model was employed to conduct this analysis.

4.2. Decision Tress

Decision trees are the traditional ML modelling approaches and produce results with higher accuracy when contrasted with others. It is an algorithmic modelling approach that information parting was finished by unmistakable conditions [9]. Numerous investigations were viewed as choice trees as an extraordinary way to deal with directing a prescient examination. In AD expectation, we start from the tree root highlight and look at this include with other tree hub highlights. Because of the connection, we seek after the branch identifying with that worth and leap to the following hub [10]. It is imperative to keep distinctive AD gatherings and another tree inside hubs until we accomplish a leaf hub with a predicated class. To understand the decision tree algorithm is better, the impurity concept has to dig into more in detail (refer to Figure 4.1.).

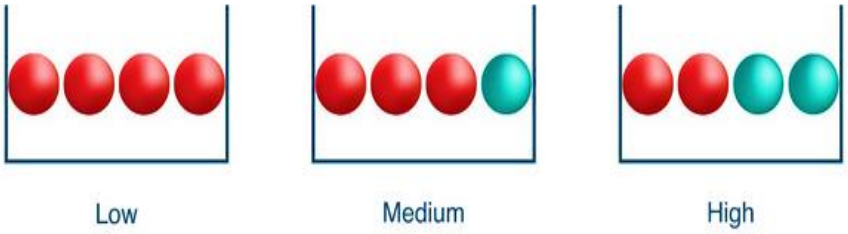

Figure 4.1. Colour ball concept to understand the purity

In the first bowl, the probability of getting a red ball is 100% as the bowl fills with pure red colour ones. Similarly, the central bowl needed more information than the left, and more data is required to understand the purity level and the right bowl needs maximum information since both colours are distributed equally. Therefore, we can conclude that the left bowl is pure, the middle is less impure and the right is completely impure. To measure the impurity of any sample, two parameters called entropy and Gini index will calculate it.

Entropy is the amount of data that needs to describe some sample. In a homogenous sample, entropy is '0' and for a heterogeneous sample, entropy is '1'. Mathematically, entropy is written as

$$\text{Entropy} = - \sum_{i=1}^{n} pi * \log (pi)$$

Gini index (GI) or Gini impurity is a measure of sample inequality ranging from 0 to 1. If GI=0, the sample is perfectly homogeneous and for GI=1, the sample having maximum inequality in the sample. It is the sum square of individual class probabilities of each class. It can be illustrated as

$$\text{Gini index} = 1 - \sum_{i=1}^{n} P_i^2$$

4.3. Results

The same dataset presented in Chapter 3 was applied with single ML modelling by decision trees. Given the AD grouping, all the independent features were supplied as input to the J48 decision trees model. K-fold validation (CV) methods were utilized for model validation. The CV is a resampling procedure with a unique feature 'k', that is utilized in model assessment on a restricted data set. Given the 'k' value dementia dataset has spilt into test and train groups. The CV was conducted with k = 5 to overcome the fitting problems, which presents for five data folds (or subgroups) for testing, and k-5 folds for preparing purposes had utilized. For producing the pruned decision tree, we included limited key features like CDR, MMSE, n-WBV, gender, and MR delay since these are mostly correlated with the AD groupings.

The performance of the model was assessed by accuracy, precision, and area under the receiver operating characteristic curve (AU-ROC). Dataset pre-processing was done by a selection of highly correlated features coupled with AD subjects. Training of model was held between the outcome AD group and the remaining features can help in the model operation of AD prediction. Figure 4.2. presents the experimental results of the decision tree outcome.

From Figure 4.2, it is clear that 331 images were accurately classified out of 373 with 88.7% of accuracy. A weighted average on a prediction of true AD cases (i.e., precision) of 86.7% was recorded. The precision true AD subjects are assessed as 91.3%.

$$\text{True AD prediction} = \frac{\text{True AD subjects}}{\text{Total of all true positives}} = \frac{188}{188+18} * 100 = 91.3\%$$

The J48 pruned decision tree with CDR as a central node can be observed in Figure 4.3. If the branch CDR ≤ 0, classification of MRI session was done as AD_{non} with 92% of accuracy. The second branch CDR > 0, parting into two parts of MR delay is the central node. It developed an accuracy of AD subjects is 98.2%, alongside another branch with an MMSE central node. This tree follows the base node with an AD group classification.

Predictions by generated decision tree have correctly mapped and examined with confidence attributes of dementia status. At last, the high confidence correlated value of attributes can predict dementia in of particular adult, and the referenced model clarifies and forecast the patient's condition by using explicit advantages to help patients by helping them ahead of time.

Correctly Classified Instances	331	88.7399%
Incorrectly Classified Instances	42	11.2601%
Kappa statistic	0.7992	
Mean absolute error	0.1085	
Root mean squared error	0.2609	
Relative absolute error	28.1314%	
Root relative squared error	59.4574%	
Total Number of Instances	373	

--------- Detailed Accuracy By Class ----------

	Precision	ROC Area	Class
	0.913	0.937	Nondemented
	0.924	0.962	Demented
	0.409	0.650	Concerted
Weighted Avg.	0.867	0.918	

---------- Confusion Matrix ----------

a	b	c	< -- classified as
188	1	1|	a = Nondemented
0	134	12|	b = Demented
18	10	9|	c = Converted

Figure 4.2. Performance of J48 model

Initially, the subjects with AD were mapped with the independent features which are highly correlated with dependent AD group category. The CDR rating assessed late-life AD forecasting. Irrespective of age, if CDR ≤ 0, at that point subjects were delegated as AD_{non}, and CDR > 0 high percent of subjects were named as AD, and rest were as AD_{con}. The resulted decision tree was created with various sub-branches and left a choice toward the end, considered as a leaf of the comparing branch. Eventually, results proposing that pruned decision tree models are perhaps the best methodologies with a precision of 88.7%.

As mentioned, the ROC curve metric was assessed as principal examination in the clinical analysis [133], and basically, it's a plot with the true positive rate on the y-axis and false positive rate on the x-axis. Figure 4.4 presents experimental results of ROC curve. According to ref [134], in the classification of disease diagnosis, ROC close to one that implies it has successfully classified disease type with GI equal to zero. On the off chance that it close to zero said to have a most noticeably worst measure of distinctness. In this analysis, we got the ROC of AD grouping is 0.962, which implies that a complete characterization of AD patients was finished.

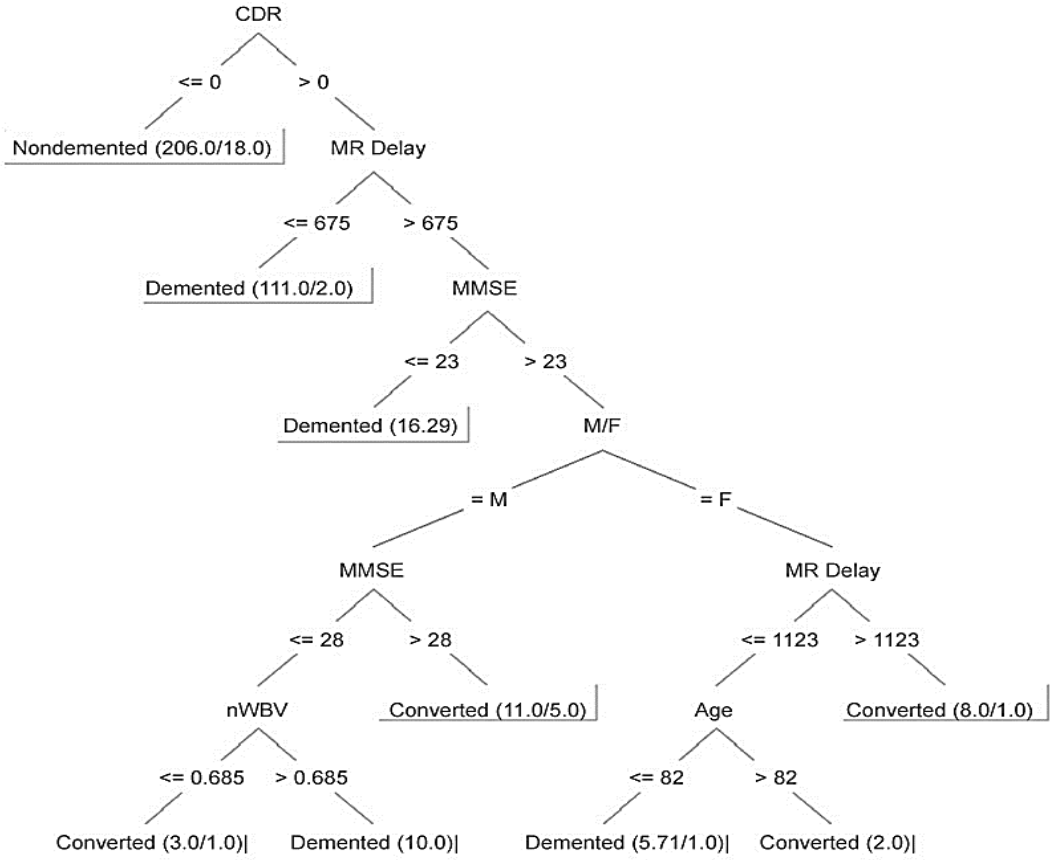

Figure 4.3. Pruned decision tree (J48)

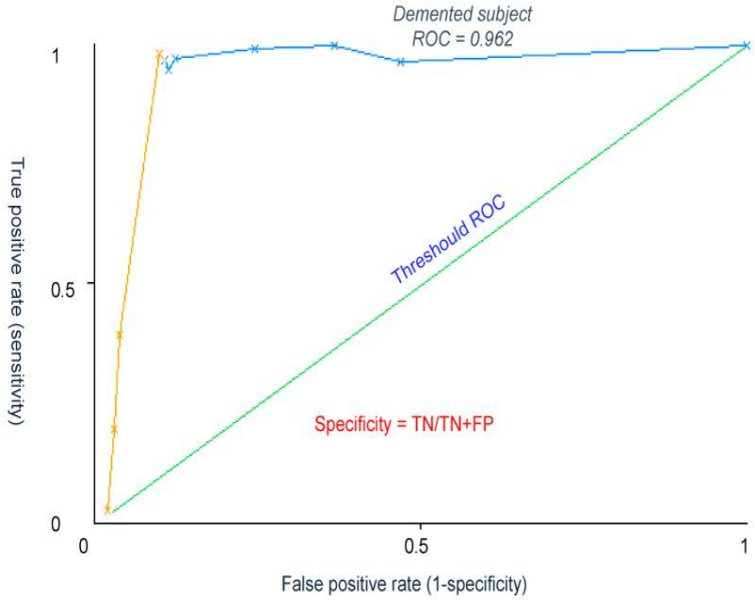

Figure 4.4. ROC curve of AD subjects.

4.4. Chapter summary

In the diagnosis of AD in older adults by MCI studies, MRI demographic data alongside different attributes profoundly significant in AD prediction. In this chapter, we presented the building of an ML model with limited features (pruning) strategy to improve classification accuracy. Particular clinical diagnostics have created with the association of ML usage. However, few studies only associated with the classification of AD subjects.

Diagnosis of AD is not an easy task and it requires sophisticated data analysis since it may require test data, physical examination, psychological testing, research office studies, and MR images. As of this, we consider explicit features, for example, CDR, MR deferral, MMSE, and n-WBV.

Single modelling associated with support vectors presented only the prediction accuracy of 68.75% (Refer chapter 3) which is a relatively low accuracy value in brain-related studies. To overcome this problem, we adopted J48 decision tree modelling with pruning technique and achieved 96.2% of classification accuracy. Chapter 3 & 4 associated with single modelling methods, and we extend our research in the next chapters by adopting multi modelling approaches.

The results summarized in this chapter were fully published in *Gopi B, Nalini C, Francesco A. Late-Life Alzheimer's Disease (AD) Detection Using Pruned Decision Trees. Int J Brain Disord Treat. 2020;6(1). doi:10.23937/2469-5866/1410033*

Chapter 5

Comparative machine learning approach in dementia patient classification using principal component analysis (PCA)

In this chapter, we present another study designed to create multi-modelling ML method with a dimensional reduction of AD dataset known principal component analysis (PCA). For that, we built up a feature extraction strategy with the association of three supervised ML models, such as SVM, K-closest neighbour (KNN), and logistic regression (LR). The working background of SVM models is already presented in section 3.2.

5.1. K-nearest neighbors

KNN algorithm is one of the more basic techniques utilized in machine learning. It is a strategy favoured by numerous individuals in the business due to its convenience and low count time. KNN is a model that groups any information depending on the samples with similar characteristics. It utilizes test data to make an "informed speculation" on what an unclassified point should be classified. KNN is frequently utilized in simple proposal frameworks, picture acknowledgement innovation, and dynamic models.

KNN works on account of the profoundly established numerical hypotheses its uses. While actualizing KNN, the initial step is to change information focuses into featured vectors, or their numerical value. The calculation at that point works by finding the distance between the numerical estimations of these features. The most well-known approach to discover this distance is the Euclidean distance, as demonstrated as follows.

$$d(\mathbf{p},\mathbf{q}) = d(\mathbf{q},\mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}.$$

KNN runs on this equation to process the distance between every data point and the test data. It then finds the likelihood of these points similar to test data and orders is dependent on which focuses share the high probabilities. The visualization of the above formula is presented in Figure 5.1.

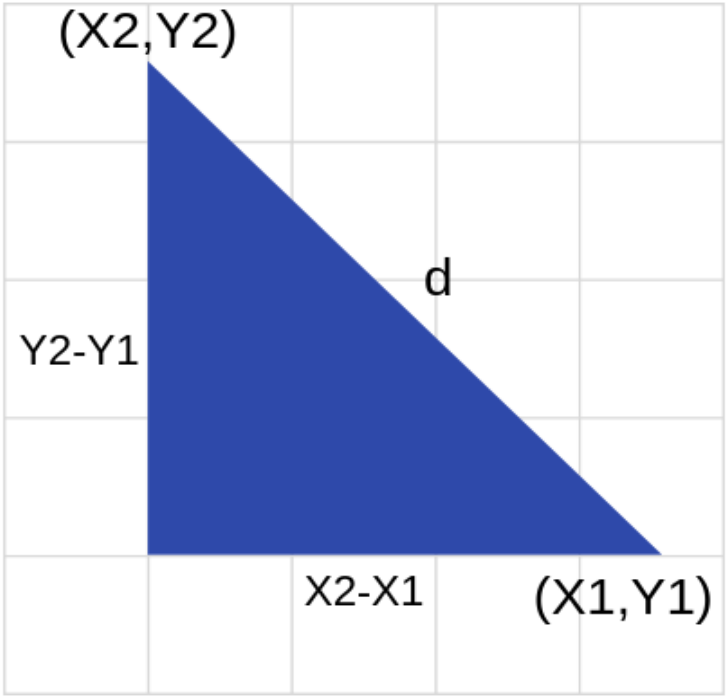

Figure 5.1. Visualisation of Euclidian distance formula

5.2. Logistic regression (LR)

Logistic regression (LR) is another classification algorithm used to distribute perceptions into the discrete arrangement of classes. It is characterized into the twofold, multi, and ordinary level types. LR doesn't demonstrate a connection between non-continuous variables but permits the forecasting of the discrete attributes [135]. It is extremely simple to execute and very effective during model training.

Mathematically, LR is written as multiple linear regression with the equation by

$$\text{Logit}\,(\text{P}) = \left(\frac{\text{m}(\text{x} = 1)}{1-(\text{p} = 1)}\right) = \beta + \beta 1.\,\text{x}1 + \beta 2.\,\text{x}2 - -\beta\text{i}.\,\text{xm for i} = 1 \ldots .\,\text{N}$$

In the given dataset, subjects classified as either demented or non-demented that presents a simple logistic binary function. As discussed, two target demented groups (with dementia- '1' or without dementia- '0') have been validated.

Hypothesis W = AX+B

$$H(x) = \text{sig}(W)$$

If 'W' touches positive infinity, then the positive prediction could happen, and if 'W' reaches to negative infinity, then the negative prediction could happen as mentioned in Figure 5.2.

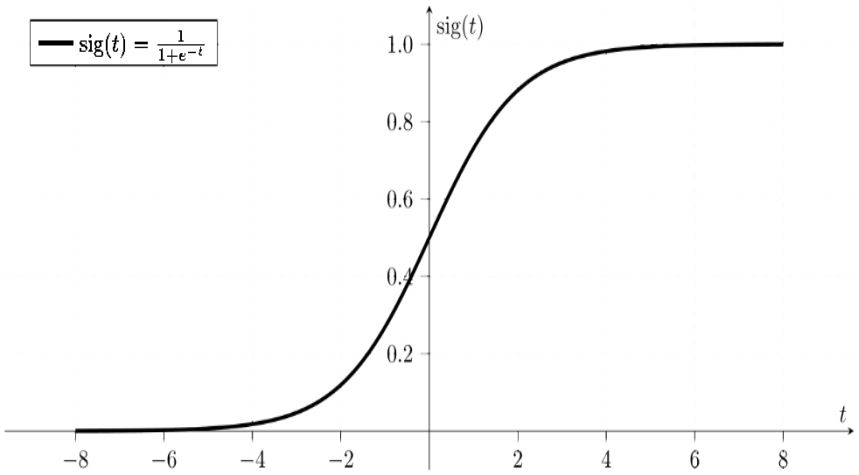

Figure 5.2. Simple binary logistic regression representation (where sig (t) sigmoid activation function).

In the present study, a dataset of 150 patients' information (trained data) contains the relationship between "14 different features (independent value)" and "one group attribute (dependent value)". As of this, in this chapter, we aimed to design a model that can predict a patient group based on other features. A regression line was obtained (with minimum error) by using trained data. Thus, if trained data exposed to the feature extraction technique, the model should predict the patient group with less or no error.

5.3. Principal component analysis (PCA)

This section provides the motivation behind selection of principal component analysis (PCA) for performing the feature extraction technique, also answer how it functions bit by bit, so everybody can get it and utilize it, without fundamentally having a solid mathematical background.

PCA is a broadly covered technique on the web, and there are some extraordinary articles about it, yet just not many of them go directly to the point and clarify how it functions without jumping a lot into the details and the 'why' of things.

PCA step by step explanation

❖ Standardization

The point of this standardization is to normalize the scope of the constant introductory factors with the goal that every single one of them contributes similarly to the investigation. In particular, the motivation behind why it is basic to perform standardization preceding PCA is that the last is quite sensitive to the differences in the underlying factors. That is if there are huge contrasts between the scopes of starting factors, those factors with bigger reaches will rule over those with little ranges (For instance, a variable that ranges from 0 and 100 will rule over a variable that ranges somewhere in the range of 0 and 1), that can lead to one-sided results. Along these lines, changing the information to similar scales can prevent this issue.

Mathematically, this should be possible by deducting the mean and dividing by the standard deviation (SD) for each estimation of every factor.

$$Z = \frac{value - mean}{SD}$$

Once the standardization has been done, all variables can be transformed into the same scale.

❖ Covariance matrix computation

The covariance matrix is a p × p symmetric matrix (where p is the number of measurements) that has as sections the covariances related with all potential sets of the underlying factors. For instance, for a 3-dimensional dataset with 3 factors x, y, and z, the covariance matrix is a 3×3 matrix of this from

$$\begin{bmatrix} Con(x,x) & Con(x,y) & Con(x,z) \\ Con(y,x) & Con(y,y) & Con(y,z) \\ Con(z,x) & Con(z,y) & Con(z,z) \end{bmatrix}$$

❖ How PCA can construct the principal components

For instance, if we consider 3-dimensional data set, there are 3 groups, therefore there are 3 eigenvectors with 3 corresponding eigenvalues. Without further, it is eigenvectors and eigenvalues behind, because the eigenvectors of the Covariance matrix are actual directions there is the most difference (most data) and that we call Principal Components. Also, eigenvalues are essentially the coefficients connected to eigenvectors, which give the measure of change conveyed in every Principal Component.

❖ **Feature vector**

The feature vector is a matrix that has as columns present components of the eigenvectors that we choose to keep. This makes it the initial move towards dimensionality reduction since, in such a case that we decide to keep just p eigenvectors out of n, the set of final data will have just p measurements.

❖ **Allocate the data along with the principal components**

From the previous steps, aside from normalization, we don't roll out any improvements on the information, and simply select the principal components and develop the feature vector, yet the input data collection remains consistently as far as the actual axis (i.e., regarding the initial variables). The point aims to utilize the feature vector shaped utilizing the eigenvectors of the covariance matrix, to reform the information from the original axes to the ones presented by principal components (i.e., that's why it names as Principal Components Analysis). This should be possible by increasing the transpose of the actual data by the translate of the feature vector.

$$\text{Final dataset} = \text{Feature Vector}^{T} * \text{Standardized Original Dataset}^{T}$$

5.4. Model outcome

A comparison of the three ML model's performance has been done. At first, OASIS longitudinal dataset presented to the R platform that presented in Figure 5.3 and model testing directed with two datasets: actual OASIS-2 dataset and a dataset after PCA. Pre-processing associated with the forecasting of missing values by the attribution of K-NN. Feature extraction was performed with the assistance of the PCA procedure. High correlated features were chosen for better results. Every ML classifier was independently by CV methods (with k=10).

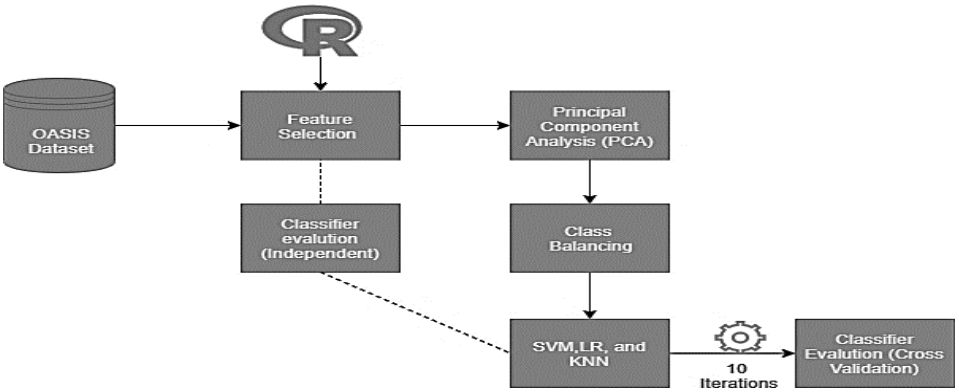

Figure 5.3. Experimental setup

5.5. Results

To forecast explicit patient-related with AD or not, a predictive model has to be accurately characterized by the examples. SVM, LR, and KNN models were utilized to create predictive models (Table 5.a.). The three ML model performance models were analysed with parameters such as recall, precision, and area under the curve (AUC) [136]. Among the given models, LR has produced 98.3% of accuracy that is followed by KNN and SVM with 97.6%, and 96.7% of accuracy respectively. Three models were producing comparative accuracy rates. In often, model accuracy isn't simply enough to pass judgment on the model performance. Thus, the examination of other performance parameters like recall, precision, and AUC is obligatory to characterize model validation.

We discovered a similar accuracy of two models (LR and KNN) about 98± 0.04%. When compared and the other two models, SVM was producing a low positive prediction rate of 97.1%. Simultaneously, sensitivity (recall) can characterize true positives from actual total positives. Both recall and precision depend on the comprehension of the significance of positive results. From Table 5.1, the sensitivity for LR model forecasting registered as 97.4%. On the other hand, KNN produced 98.3% of high sensitivity, and SVM with the low sensitivity of 96.6% can found.

Table 5.1. Performance metrics of different predictive models

Model	Accuracy	Precision	Recall	AUC
SVM	0.967	0.971	0.966	0.983
LR	0.983	0.986	0.974	0.997
KNN	0.976	0.982	0.983	0.996

However, in ML, AUC can help to avoid classification issues. It is one of the key presentation instruments for model execution checks. The AUC was ranging in between 0 to 1 that presented in Figure 5.4. By definition, if AUC ≈ 1, at that point the model was effectively recognizing the target group (dementia). The AUC values of LR, KNN, and SVM were 99.7%, 99.6%, and 98.3% respectively.

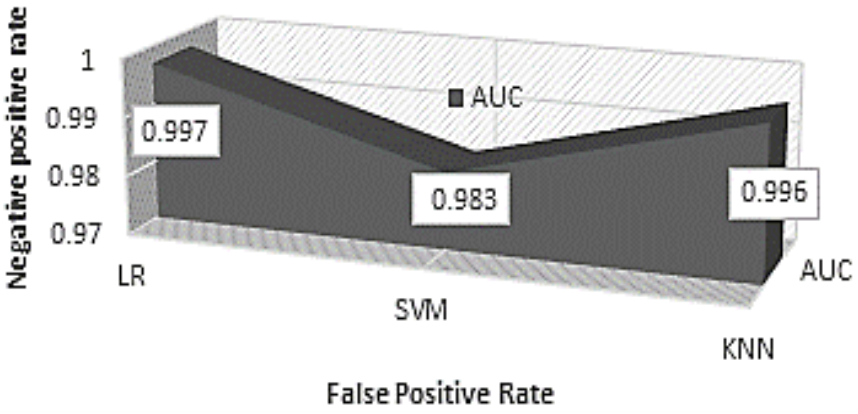

Figure 5.4 Graphical representation of AUC values

5.6. Chapter summary

This chapter described the comparative machine learning approach with three ML models such as SVM, LR, and KNN that were characterized by dementia group patients. Feature extraction was performed by PCA techniques that performed through R statistics. Different performance parameters were described in the model performance. Outcomes approved that the three models are precisely grouping dementia patients with better accuracy from 96.7-98.3%. We also validate the adopted models with recall, precision and AUC. The AUC of LR and KNN presented optimal prediction model, with the end goal that these two prescient models were done better classification of the dementia patients.

The results summarized in this chapter were fully published in *Battineni G, Chintalapudi N, Amenta F. Comparative machine learning approach in dementia patient classification using principal component analysis. In: ICAART 2020 - Proceedings of the 12th International Conference on Agents and Artificial Intelligence. ; 2020. doi:10.5220/0009096907800784*

Chapter 6
A novel Machine-Learning Model Applied to Magnetic Resonance Images (MRI) in AD Prediction

In this chapter, Four ML models, such as Naive Bayes (NB), artificial neural networks (ANN), K-nearest neighbour (KNN), and support vector machines (SVM) were presented. The receiver operating characteristic (ROC) curve metric were used to validate the model performance. Each model evaluation was done in three independent experiments. In the first experiment, a manual feature selection was used for model training. In the second experiment, automatic feature selection was conducted by wrapping methods, and the last experiment consisted of a new approach with mixed modelling called ensemble learning.

6.1. Introduction

Neuroimaging and fundamental MRI provide basic information to AD dementia forecasting and grouping [137], [138]. ML models, combined with MRI data, can give high analytic precision on age-related cognitive decline (ARCD) in dementia subjects [139]. It has been predicted that supervised ML models can produce the knowledge important features that can be correlated with AD sample data [140]. Similarly, it is reported that LR models coupled with CV techniques can improve the prediction accuracy of AD by speech amalgamation [86]. Besides, SVM models, alongside feature reduction techniques can also classify the dementia subjects with 70% accuracy (Refer chapter 3).

In this work, we adopted four models called KNN, Naïve Bayes, ANN, and SVM to detect AD based on MRI images. Three individual experiments were designed to test the model, and model performance was separately evaluated with given MRI characteristic information. The experiments that were done included

1. Models with manual selection of MRI features,
2. Models with automatic feature selection, and
3. A single model with ensemble learning or hybrid modelling.

6.2. Material and methods

Selection of features

In this step, the machine performed an autonomous selection of input features that correlates to the subject group [141]. Selection techniques are largely used and standardized to reduce unnecessary features and to enhance model accuracy [142]. Moreover, this approach measures the relationship between independent variables and the target outcome. Feature selection can be conducted by three approaches, namely, filtering, regularization, and wrapping [142], [143]. In this study, the wrapping technique was used because it amplifies model performance with limited features.

Feature importance

This method results in a "feature score" assigned to independent characteristics and a defined score to each characteristic that is highly correlated with the subject "group". The correlation between each characteristic-associated group variable is shown in Figure 6.1. The CDR rating was excluded during model development because it did not have the highest relevance, but it helps in subject groupings.

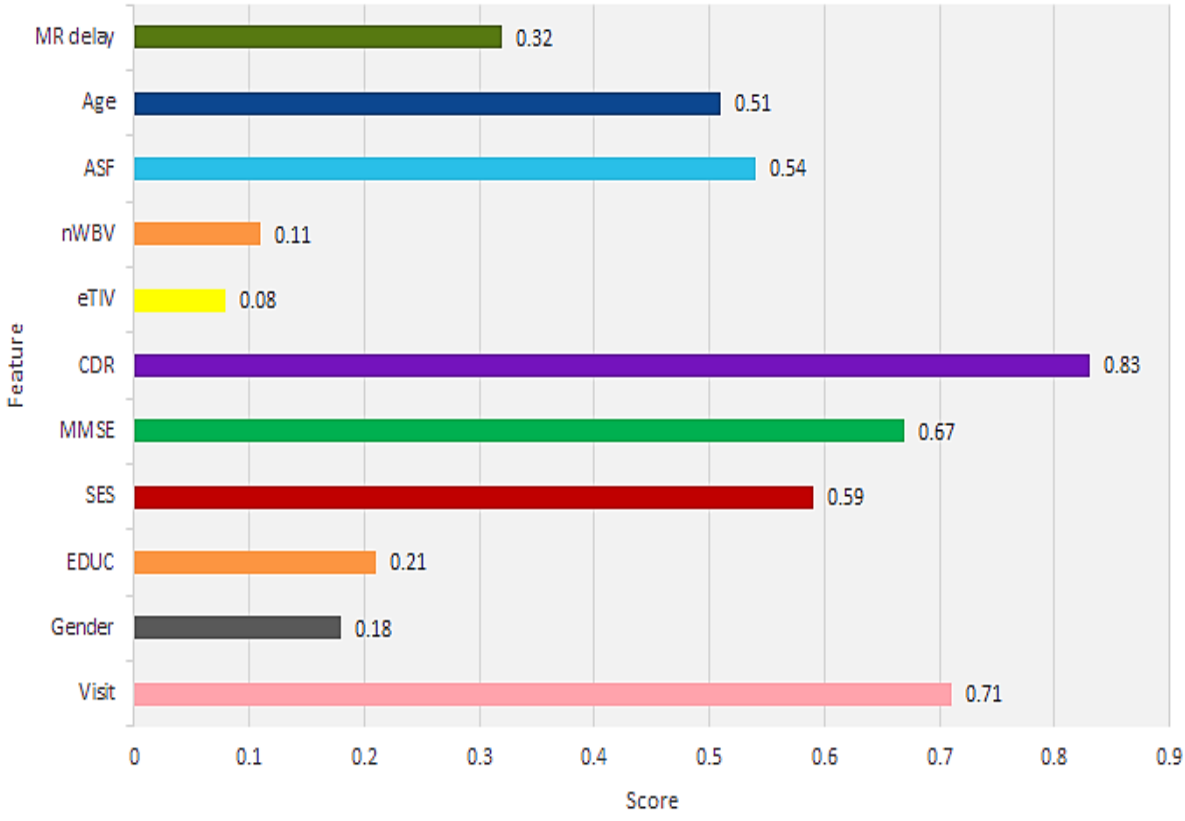

Figure 6.1. Individual feature scores.

Feature Selection with wrapping

In the wrapping method, feature search represents a big challenge in calculating model accuracy [144]. Feature selection can be made as either step backwards or forward, and exhaustive. Feature search helps the identification of primary features in the enhancement of model performance. The MRI characteristics with a correlation of at least 0.5 can automatically help to develop a model. Figure 6.2 shows the scatter plot of feature results following the wrapping method.

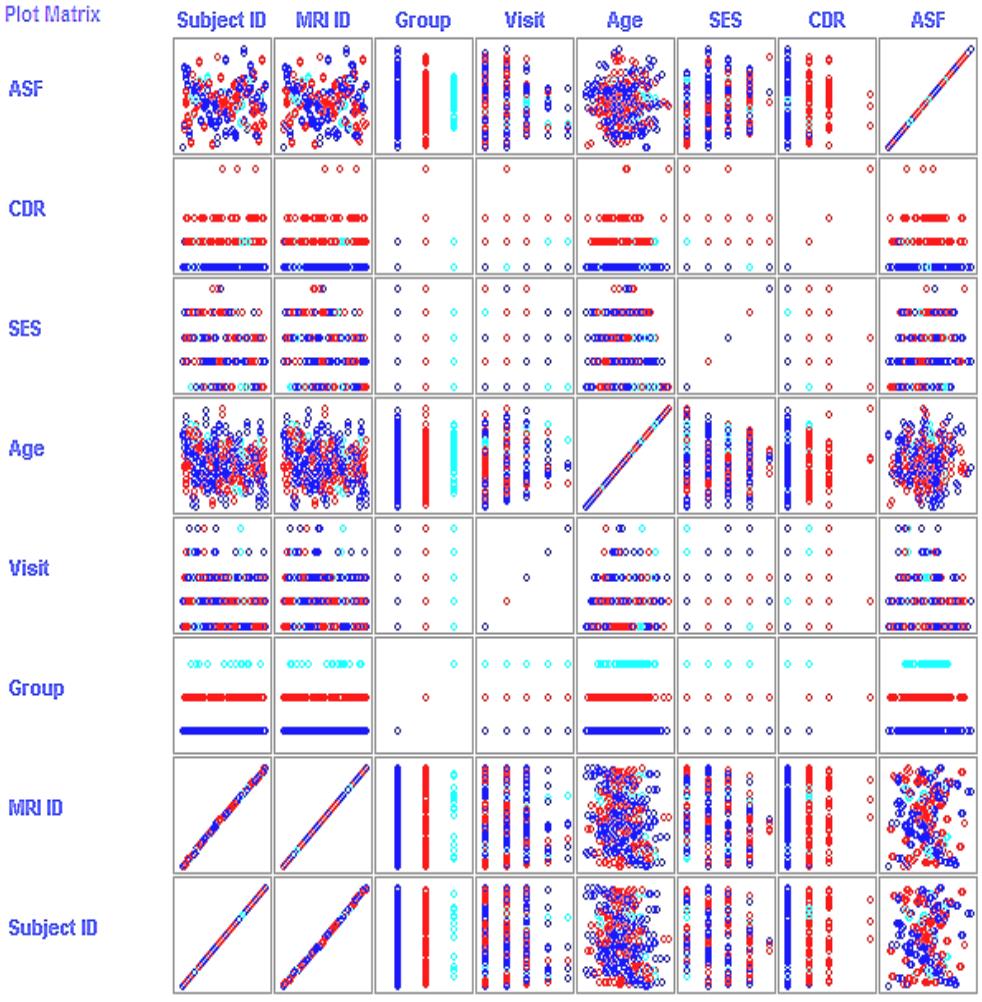

Figure 6.2. Scatter plot of selective features. Blue dots (ND), red dots (D), light blue dots (C).

Model classifiers

Naïve Bayes (NB)

A Naive Bayes (NB) classifier is a probabilistic ML model that is utilized for the classification task [145]. The basic idea of the NB classifier depends on the Bayes hypothesis. The Bayes theorem is defined as

$$P(A/B) = \frac{P\,(B/A)\,P(A)}{P(B)}$$

With the above theorem, we can discover the likelihood of 'A' occurrence, given that 'B' has happened. Here, 'A' is the hypothesis and 'B' is the evidence. The theory made here is that the features or predictors can be independent. That is the presence of one specific component doesn't influence the other. Therefore, it is named it as naïve Bayes classification.

NB is a probabilistic model that forecasts output based on Bayes' equation. Due to its simplicity during target prediction, it has become popular in classification and multiclass predictions [146] like AD classification.

Artificial neural networks (ANN)

In machine learning, artificial neural network (ANN) is one of the important tools. They are biologically inspired systems which are aimed to imitate the approach that humans learn through the brain and thus the name "neural network". Artificial neural networks (ANN), also called connectionism, start to be a standard method to model a phenomenon from examples with good performance [147]. ANN operates through layers which is part of input and output, additionally a hidden layer which includes units that renovate the received input into an analysed prediction through some training which can be obtained by the output layer. Mostly ANN's are an exceptional tool for discovering patterns which are too dense or abundant for a human scientist to obtain and provide training for the machine to recognize.

The ability of ANN in engineering fields has been proven that it is powerful in simulating and predict most cases in various physical phenomena behaviour. ANN was initially adapted from the capability of the biological human brain in which it used a list of neurons, also called nodes, interconnected each other to simulate the biological nerve systems [148]. ANNs are mainly used for pattern recognition, classification, and

prediction and can approximate almost any functional relation. The interrelationship between the input variables and the output parameters is not approximated using a traceable functional coherence. ANNs are more robust and less sensitive to outliers and chaotic components compared to other methods [149].

Artificial neural networks are also referred to as "perceptron" have become an integral part of artificial intelligence in recent decades. This is predicted to have happened due to the establishment of a technique called "backpropagation", which enables the neural network to alter their hidden layers of neurons in circumstances where the result would not match with the actual predicted output, like for example, the network, which was trained to identify say dementia subjects, instead predicts it to be a non-demented case.

An additional vital advancement is the arrival of deep learning neural networks, which facilitates multilayer network feature extraction through different layers to recognize the exact object through training [150].

Model validation and framework

As mentioned, model validation can be done by either holdout (spilt) or cross-validation (CV) techniques. During this study, we adopted the CV technique because of its popularity in target prediction, with low bias. Simultaneously, it also applies a resampling method with limited features during model validation [151]. The model framework used during simulation is represented in Figure 6.3.

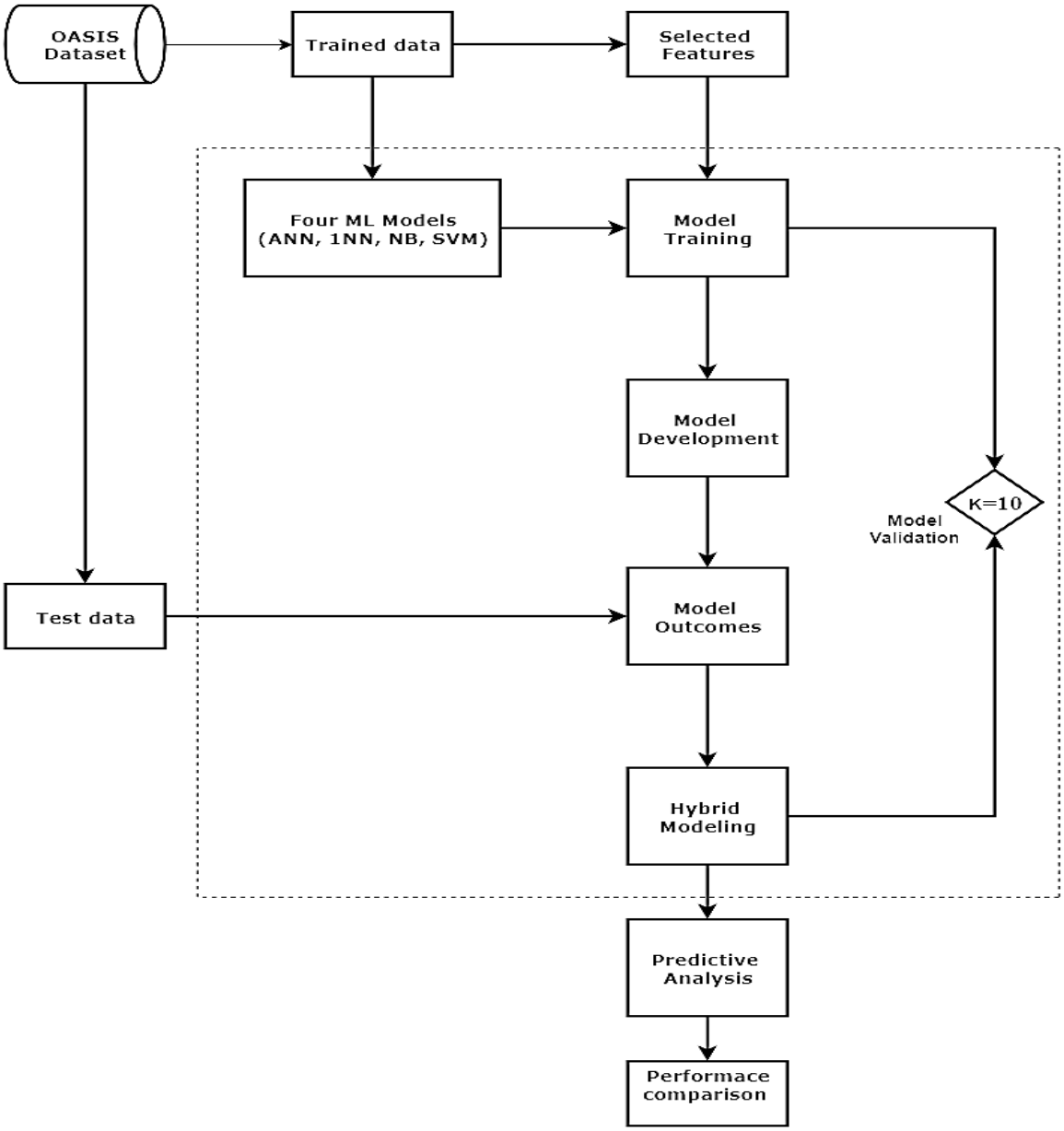

Figure 7.3. Model framework for evaluating predictive classifications.

Experiments design

A large number of MRIs for a low number of subjects could generate bias in dementia detection. Therefore, we considered final MRI scans that define the status of each subject. Three experiments were conducted, including manual and automatic feature selection techniques.

In the first experiment, model training was done using the original dataset with manual feature selection. In ANN, the number of layers (N) is used as a search parameter during model evaluation. In KNN, k is tuned to one (i.e., 1NN). In SVM, the linear kernel coupled to regularization parameter "C" and a standard deviation of radial basis function "r" are implemented in model tuning. Finally, model validation was done with a 10-fold CV to avoid data fitting issues [152]. The model performance was, therefore, assessed by the above parameters.

In the second experiment, limited features that occurred as the result of wrapping were considered for conducting model training. For NB and KNN, an exhaustive search was used to calculate model accuracy with potential feature alliance to select the best of them [153]. In SVM, genetic algorithms (GAs) were used for the feature search. GAs is frequently applied in bioinformatics to generated models with high accuracy [154]. For ANN, the feature search was excluded, and the search consisted of the identification of the hidden neuron layers. Model tuning was adjusted by maintaining batch size as 100 in NB, (C, gamma) as (1.0, 1.0 × 10−12) in SVM, and k = 1 in KNN. MRI characteristics that were highly correlated (≥0.5) with subject groups were selected.

In the third experiment, the four models were combined to develop an ensemble or hybrid model. By doing this, there is the advantage of getting a high prediction accuracy of the adopted dataset. Moreover, combining several models can enable noise reduction (bagging), low bias (boosting), and better predictions (voting). We used a voting technique in this experiment because of the capability to create standalone models from trained data [155].

6.3. Results

Experiment 1: Handling of the feature set before autonomous feature selection

Table 6.1 summarizes the performance outcomes of the four models in manual feature selection. The CDR rating was excluded as it represents a dementia factor that can affect model accuracy. From the performance comparison matrix, it can be seen that the 1NN model offers better performance compared to the other tested models in terms of accuracy, sensitivity, and specificity. As already mentioned, the ROC curve plays a relevant role in diagnostic assessments to differentiate the true state subjects and to find optimal cutoff values. Moreover, a higher ROC offers better dementia prediction in given subjects [40]. Given this, the ANN model correctly discriminates against the true demented subjects, with a ROC of 0.812. The ROC of NB, 1NN, and SVM models produced ROCs of 0.753, 0.787, and 0.796, respectively.

Table 6.1. Performance comparison matrix (4 * 4) of four classifiers.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	ROC
NB	88.76	82.43	85.72	0.753
ANN	83.56	89.92	88.84	0.812
1NN	91.32	89.92	89.56	0.787
SVM	89.67	89.24	89.45	0.796

Experiment 2: Automatic Feature Selection with Wrapping

Table 6.2. shows the model performance outcomes obtained with automatic feature selection. With this approach, progress in terms of accuracy and ROC compared to manual feature selection was noticeable. SVM resulted in high accuracy (96.12%), and 1NN, NB, and ANN produced an accuracy of 95.92%, 93.44%, and 83.56%, respectively. About ROC, NB was a better diagnosis predictor, with 0.942, followed by 1NN, SVM, and ANN, with 0.916, 0.834, and 0.817, respectively.

The results of the present experiment, in which performance results were better than those obtained in the previous one, stimulated the identification of other approaches for maximizing prediction accuracy. We, therefore, extended our work to explore the outcomes of joint modelling with limited features.

Table 6.2**.** Model performance evaluation after feature selection (with selective features).

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	ROC
NB	93.44	98.21	97.32	0.942
ANN	83.56	89.92	88.84	0.817
1NN	95.92	94.92	97.36	0.916
SVM	96.12	94.94	98.23	0.834

Experiment 3: AD Predictions with Hybrid Modeling

To check if a model correctly predicted the target variable (occurrence of dementia), a confusion matrix was used. In this analysis, vertical labelling presents actual subjects, and horizontal labelling presents predicted subjects. As shown in Figure 6.4., 76 subjects

were correctly predicted as AD among 78 subjects, and 71 subjects were correctly predicted as non-AD among 72. Collectively, 147 subjects were properly predicted out of 150 subjects. This results in 98% accuracy. For reaching these conclusions, a hybrid-modelling technique, combining the four adopted models, was introduced.

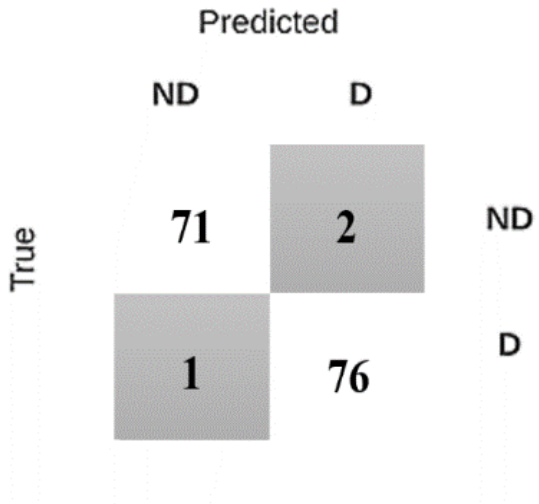

Figure 6.4. Confusion matrix outcome of the hybrid model (D: Demented; ND: Non-demented).

The performance of the individual subject group is presented in Table 7.3. Nondemented and demented subjects were correctly diagnosed with 98.6% and 97.4% accuracy, respectively. The weighted average ROC curve of both subjects nearly touches one. Hence, maximum AD subject predictions have been made without bias because of hybrid modelling. The sensitivity and specificity rates produced were 98.05% and 98%, respectively. The ROC curve of the hybrid model is shown in Figure 6.5. Based on the evaluation of performance differences in the above three experiments, the intervention of hybrid modelling with limited features resulted in being good practice in AD-related studies.

Table 6.3. Performance statistics of hybrid modelling.

Accuracy (%)	**Sensitivity (%)**	**Specificity (%)**	**ROC**	**Class**
98.6	98.7	98.6	0.992	ND
97.4	97.4	97.4	0.989	D
98.0	98.05	98.0	0.991	Weighted average

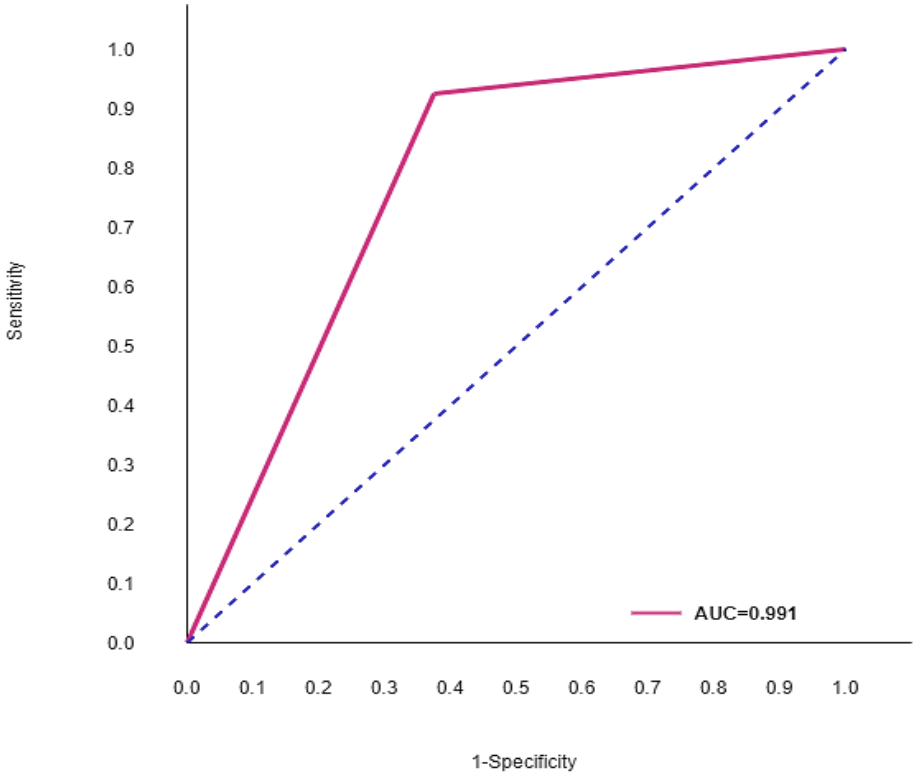

Figure 6.5. Receiver operating characteristic (ROC) curve of the hybrid model.

6.4. Discussion

ML models are highly acknowledged in real-time clinical practice and also in diagnosis and AD treatment selection [112]. Several MRI works have been integrated into ML models to make AD predictions [86], [137], but there has been no comprehensive model to amplify model accuracy. Because of this, we introduced a hybrid model to enhance the precise detection of AD based on the analysis of MRIs.

In this chapter, the significance of joint ML modelling for AD-onset prediction in elderly people has been demonstrated. Three different experiments were conducted, including manual and automatic feature selection techniques. Fourteen independent MRI features were used to identify the AD group using standard diagnostic approaches. Four supervised predictive models (NB, ANN, KNN, and SVM) were used, and the obtained results indicate the prediction accuracy of each model, constantly increasing between experiments. Figure 6.6 compares the prediction accuracy of the three experiments. 1NN generated 91.32% accuracy by manual feature selection; SVM had a high 96.12% accuracy by automatic feature selection, whereas joint or hybrid modelling enabled 98% accuracy in predicting AD in older adults. The outcomes suggest that joint modelling, with limited features, is the best practice to assess AD-onset by subject prediction.

In the first experiment, all the designed classifiers revealed enough performance values in terms of true-positive rates (sensitivity). ANN and 1NN produced the highest sensitivity (89.92%), followed by SVM (89.24%) and NB (82.43%). As mentioned, ROC curve values between 0.5 and 0.7 indicate low prediction accuracy, between 0.7 and 0.9 indicate moderate prediction accuracy, and between 0.9 and 1 indicate high prediction accuracy [156]. From Table 6.1, it is obvious that the four adopted models produce moderate prediction accuracy when checking with manual feature selection.

To amplify model performances, the second experiment was conducted with selective features after wrapping. This resulted in NB of 98.21% sensitivity, followed in descending order by SVM (94.94%), ANN (94.92%), and 1NN (89.92%). Both NB and 1NN predict subject class in a comparatively better manner, with ROC of 0.942 and 0.916, respectively. However, we argued that there could be other possibilities for enhancing prediction accuracy to values higher than those identified in the above two experiments. To support this claim, a hybrid model was developed by combining the four investigated models. A simulation of four recruited models was then performed, and thanks to this approach, the sensitivity of the model attained the highest predicted value of 97.4%, and its ROC was nearly equal to one (Figure 6.7).

The developed model produced better accuracy than other conventional models, but the present study has some limitations. First, the limited number of subjects investigated could hamper the final dementia subject prediction to the overall AD subjects; second, the outcome of the integration of three experiments may have influenced the results. The use of external MRI information does not guarantee data quality and can affect the significance of the study as a whole.

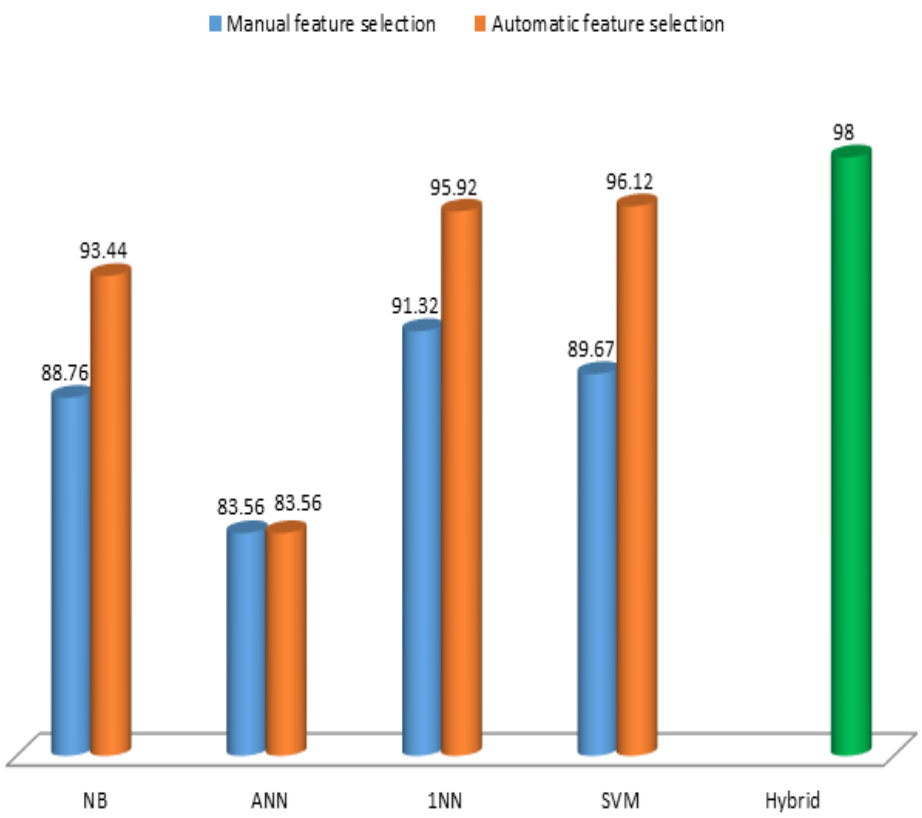

Figure 6.6. Prediction accuracy (in %) comparisons of three experiments.

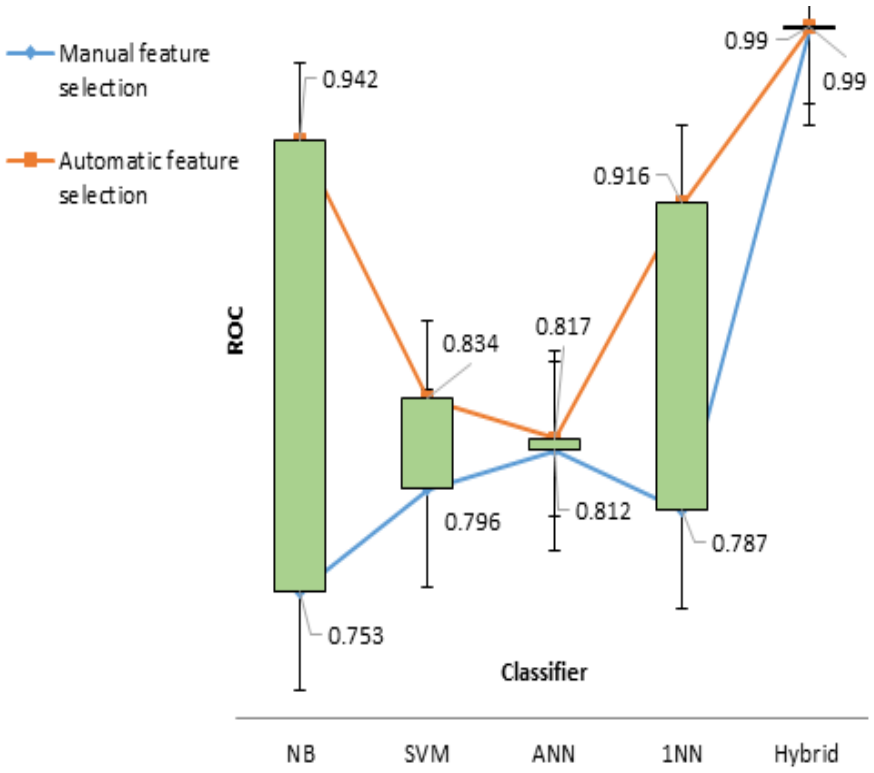

Figure 6.7. ROC comparison of hybrid modeling with other experiments.

Brain studies corroborated with artificial intelligence analysis may offer relatively faster investigation methods to modern neurological research. However, it would be preferable to avoid data limitations and, therefore, to enlarge as much as possible the size of the sample investigated in future studies. At the same time, it is also recommended to apply hybrid modelling to younger subjects or subjects with mild AD and to anticipate prediction accuracy with other biological tests like cerebrospinal fluid (CSF) or blood markers.

6.5. Chapter summary

Adult-onset dementia disorders are serious brain pathologies caused by the loss of neuron functions and to progressive atrophy. AD is the most common of these pathologies. It affects primarily elderly people and has a tremendous impact on the lives of people suffering from it. Given the long time passing between brain lesions bringing about dementia and the onset of clinical symptomatology, early identification of the preclinical and prodromal forms of the disease represents a challenge for medicine. This will reduce medical costs and could contribute to undertaking therapeutic approaches for delaying the conversion of the disease into overt dementia.

Unfortunately, the identification of AD at very early stages is extremely difficult, and there are no tools for its simple detection. We have developed different ML models to predict dementia in the elderly based on MRI findings. The hybrid model with selective features was found to enhance the accuracy of dementia prediction. Experiments with manual feature selection before automatic feature selection with 1NN produced 91.32% of accuracy, and the experiment of automatic feature selection generated 96.12% accuracy by SVM. This value significantly increased using multi modelling and produced 98% of accuracy. The predictive models developed in this study forecast early AD diagnosis and the associated risk of developing dementia. Although it is difficult to develop longitudinal projection models in older adults as compared to the younger population, future research in the field should consider addressing both genetic and nongenetic features of multifactorial hazards.

The results summarized in this chapter were published in Battineni G, Chintalapudi N, Amenta F, Traini E. *A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects. J Clin Med. 2020;9(7):2146. doi:10.3390/jcm9072146*

Chapter 7

Deep learning type CNN model architecture for AD detection through MRI

In this chapter, we considered the Open Access Series of Imaging Studies-3 (OASIS-3) dataset with 2,168 Magnetic Resonance Imaging (MRI) images of patients with very mild to different stages of cognitive decline. We applied deep learning-based convolution neural networks (CNN) which are well-known approaches for diagnosis-based studies.

7.1. Introduction

Alzheimer's Disease (AD) is the most well-known and largely diffused neurodegenerative disorder occurring in the elderly. AD negatively affects patients' everyday lives, causing an advanced decline of cognitive capabilities such as memory, language, behaviour and critical thinking [157]. Changes in cognitive impairment of AD patients start slowly and evolve rapidly over the long run.

Similar to other body parts, brains change as people get older. Some people lost thinking and incidental issues with recollecting certain things. Excessive cognitive decline, and other significant changes in the manner in which brain function is impaired [158]. The first symptoms of AD are trouble recalling recently learned data because Alzheimer's progressions regularly start in the brain areas involved in learning and memory. As Alzheimer's progresses progressively severe symptoms like confusion, mood changes, disorientation, unwarranted doubts about family and companions, and trouble talking appear. Individuals with cognitive decline or other potential indications of AD may think that it’s difficult to remember they have an issue.

AD is a type of dementia with several implications on the cognitive domain, affecting primarily thinking and memory. Specialists and different parental figures screen the movement of AD in patients by assessing the level of decrease in the patients' psychological capacities that are often classified into three stages: very mild (normal cognitive), mild cognitive impairment (MCI), and demented [17]. Figure 7.1 presents the magnetic resonance image (MRI) images of different AD conditions. Although the MCI and dementia patients both are experiencing a reduction of cognitive abilities, dementia patients would suffer from more pronounced difficulties with thinking or hampered judgment.

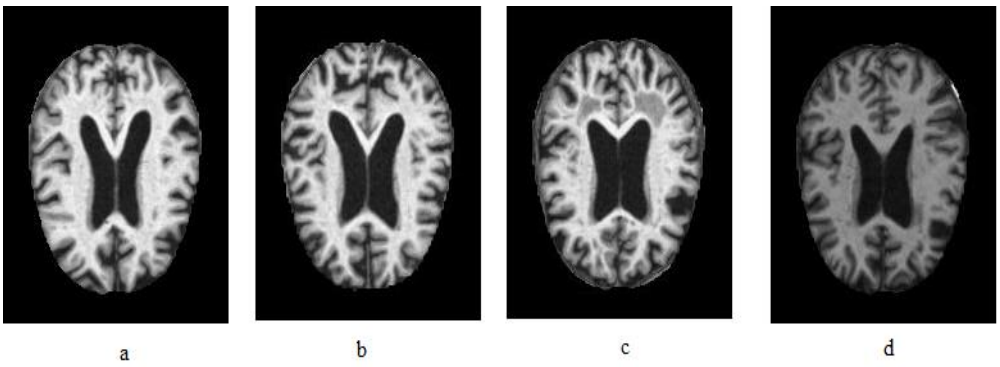

Figure 7.1. Different states of AD presented by MRI images (a) mild dementia; (b) moderate demented; (c) nondemented; and (d) very mild demented.

In clinical practice, the capacity to accurately forecast the patient diagnosis can help by adding appropriate medical decisions on treatment approaches. Recently, machine learning (ML) algorithms are largely applying to forecast and predict diseases and helping in quick decision making [112]. Pattern-related approaches like logistic regression [159], support vector machines [88], and linear discriminant analysis [102] are giving promising results in the prediction of AD development and early AD detection.

Deep learning models were used unlabeled data during preprocessing. These are well suited for imbalanced datasets and achieve a knowledge base [39]. At present these are largely involved in all other problems that are not able to be addressed by traditional artificial intelligence (AI) techniques. Neural networks are the latest deep learning algorithms that have discovered the functionality of different situations. Deep neural networks (DNN) are characterized contributions to profits through a complex composition of layers that presents building blocks including nonlinear functions and transformations. Medical experts feel that deep learning could be a promising solution in AD identification and stage detection [40]. An effective and comprehensive deep learning model can help to identify early AD prediction and ultimately provide timely treatment to the suffered patients.

In this work, we proposed convolutional neural networks (CNN) model of deep learning type for detection of early-stage AD and successfully classify the MRI images on four different dementia stages presented in Figure 7.2.

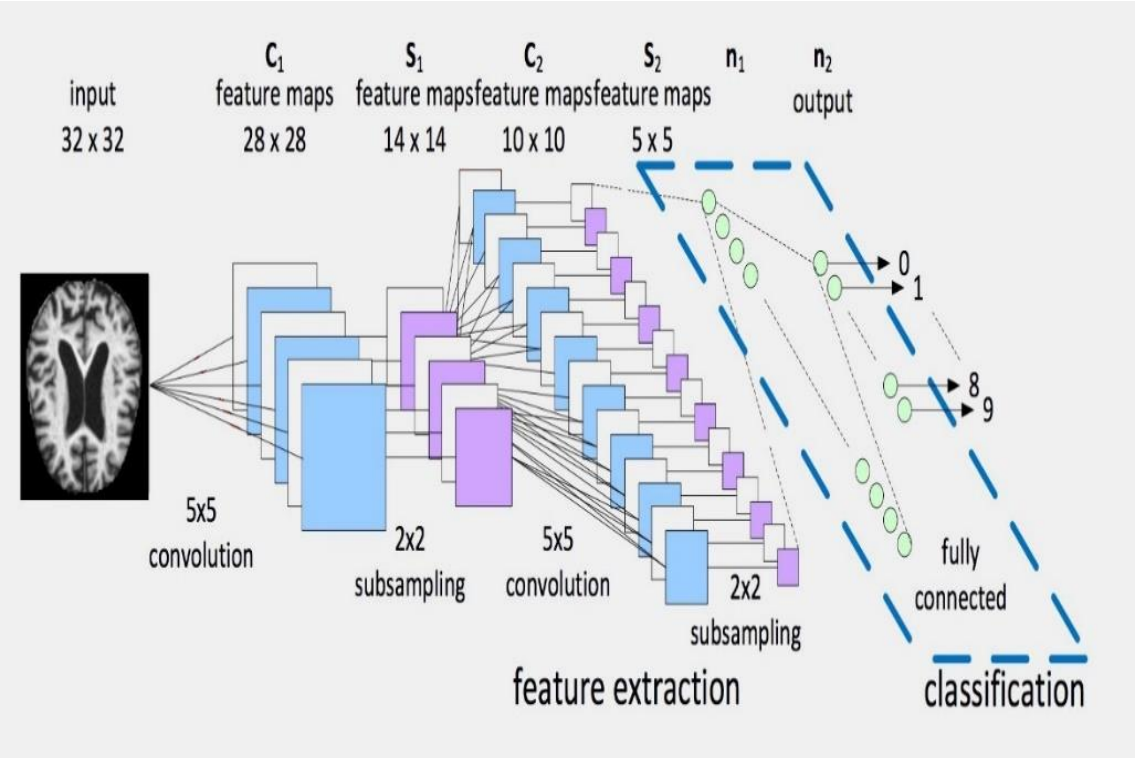

Figure 7.2. Brain image classification with the CNN model framework

Experiments were conducted on longitudinal neuroimages of the OASIS-3 database that include MR scans of T1-weighted, T2 weighted, ASL, SWI, DTI sequences, FLAIR, time of flight, and resting-state BOLD [10].

7.2. Methods

Dataset

The Open Access Series of Imaging Studies (OASIS) contains MR scanning information that is openly accessible to scientific communities. They released OASIS-1 (cross-sectional) and OASIS-2 (longitudinal) MRI datasets among different subjects and these datasets are widely used in many studies [160], [161]. OASIS-3 is the extension of previous datasets. It includes 1,098 patients ageing from 42 to 95 years. Among participants, 609 are associated with normal cognitive decline (very mild), and 489 were associated with different cognitive decline stages. OASIS-3 dataset incorporated both functional and structural features of more than 2,000 MRI images. The dataset outcome of four categories of MR images has presented in Figure 7.3.

CNN model architecture

A convolutional neural network (ConvNet) is deep learning type algorithms that take images as input, assign features based on their importance (biases and learnable weights) to different image objects, and also be able to separate one from the other [162]. When compared with other classification models, ConvNet possesses low complex pre-processing steps. In CNN, each input image is gone through sequence convolution layers namely pooling layers, filtering layers (kernels), and fully connected layers (FCs). To make the proposed model easier for understanding, we created a dense layer block and convolution block. The architecture of the CNN model is inspired by the article [163]. We built the CNN model by using five convolutional slabs covered with convolution layers, feature engineering, max pooling, and classification. We have used cross-entropy as a loss function and Adam as an optimizer. SoftMax has been used to classify the multiclass AD stages since it is associated with a mutually exclusive relationship. The feature representation (f_k) works as an input to the SoftMax layer and interprets output brain stages. A probability score P (k) for each class as defined as

$$P_k = \frac{\exp(fk)}{\sum_{k=1}^{K} \exp(fk)}$$ where fi feature representation; and

Cross entropy loss function (L) = $-\sum_{k=1}^{K} tk \log(pk)$; where t_k ground truth of MR image then

$$\frac{\partial L}{\partial fk} = P_k - t_k$$

Figure 7.3 presents the most relevant procedures followed to construct the feature data of brain images and extraction of AD images developed in this paper. After pre-processing steps, the given image dataset has been divided into training and validation files with standard (80:20) division.

The procedures indicated red line are MR images that fed to the CNN model for training purposes. The model extracts the input image features of trained images under present parameters and supplies them to the SoftMax classifier for testing. The SoftMax function calculates the loss and model accuracy. For avoiding high loss, network parameters are adjusted by the back-propagation algorithm. After applying several iterations (epochs) the better-trained parameters have been achieved. The model visualization metrics like loss and receiver operating characteristic area under the curve (ROC AUC) has been taken as the performance parameter for AD classification since it has been considered one of key metrics in multi-image classification techniques [15]. The experimental setup

and AD detection and classification have been done through TensorFlow [16] and python language.

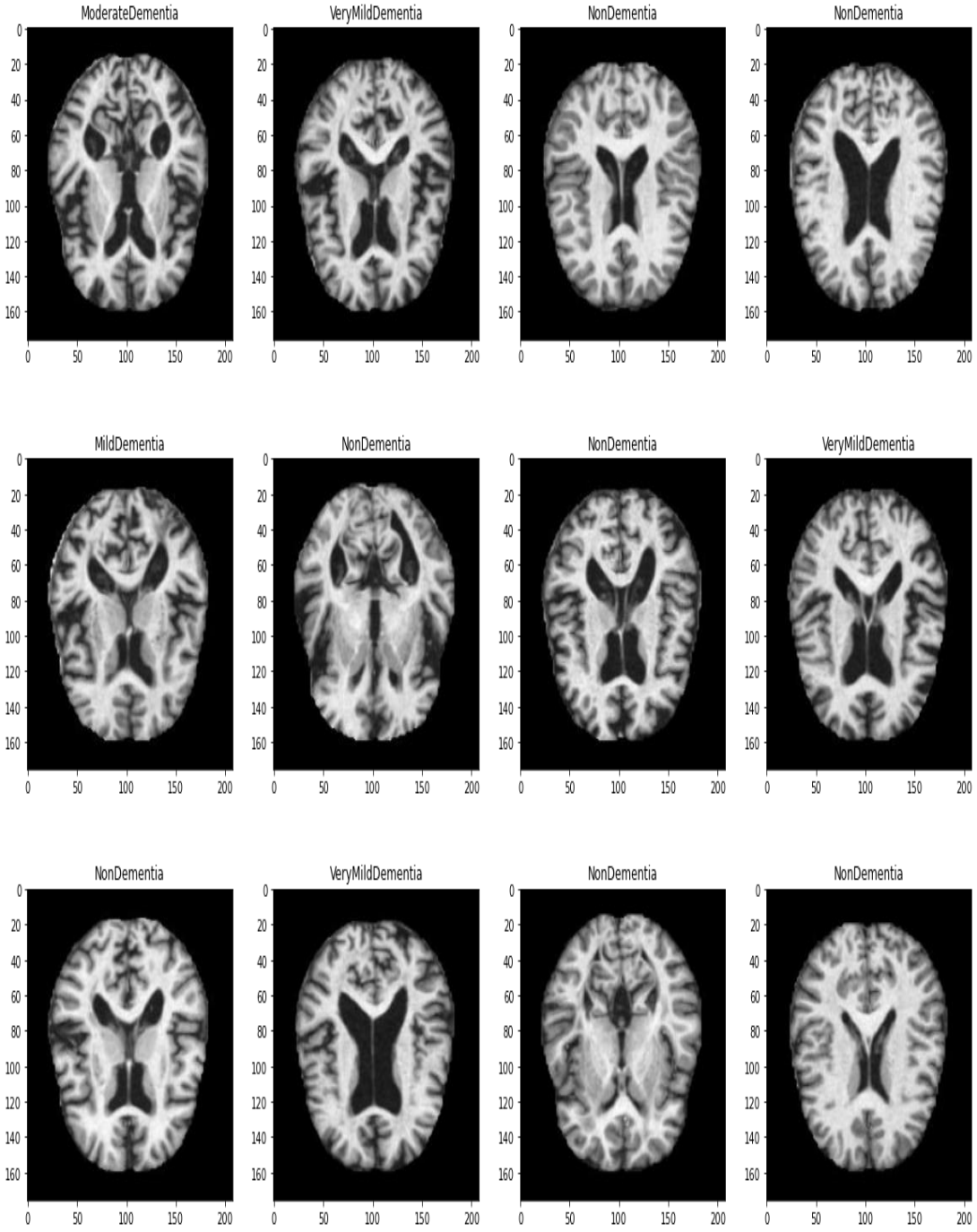

Figure 7.3. Dataset outcome of different dementia stages (3*4 image matrix)

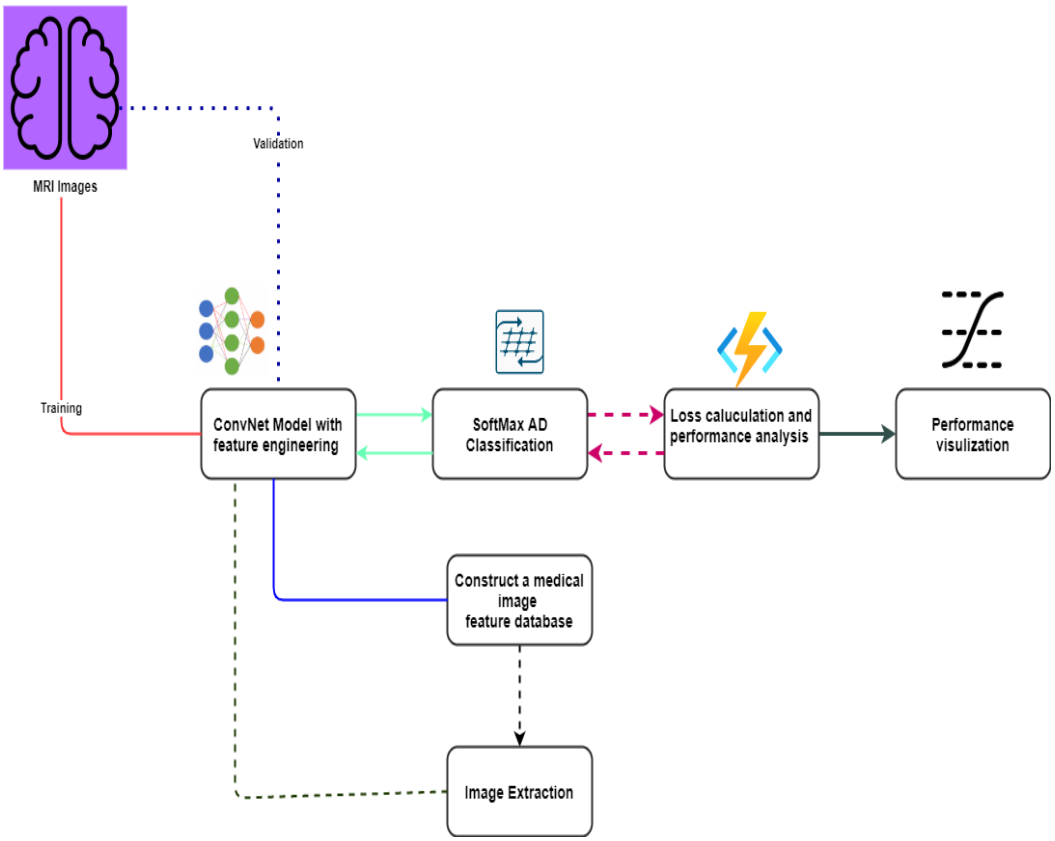

Figure 7.4. Experimental setup of the work

7.3. Results

To do efficient training on our CNN model, a back-propagation algorithm is set to adjust the rate of learning and stop the model automatically once it reaches maximum accuracy. Since the learning rate is one of the hyperparameters that decides model accuracy and time to process the model. OASIS-3 dataset consisted of 2168 independent MRI scanners. Among the given images, 1,734 are used for training and 434 were used for validation purposes. Because of the large image dataset, 10-fold cross-validation has been used and we have used each fold 70% as training, 10% as validation and 20% images are used testing. The distribution of the dataset is presented in Table 7.1.

Table 7.1. Total image distribution

Total Images: 2168	
Type	**Percentage**
Trained images	1517 (70%)
Testing images	434 (20%)
Validation images	217 (10%)

The model-fitting has to be done on a sample of 100 epochs and to prevent model overfitting we stop the model early at the 80th iteration. The model took a run time of 138 min to process the trained images. Figure 7.5 presents a graphical representation of ROC AUC and loss metrics after each iteration on both training and validation image data. Though the model evaluation has been done on the validation dataset, we also perform the experiments on the testing dataset. The testing dataset model AUC curve outcome has presented in Figure 7.6 and the model achieved a ROC of 83.3% which is considered as an optimal classifier for AD image detection and this value is significantly higher than traditional ML approaches [88], [164].

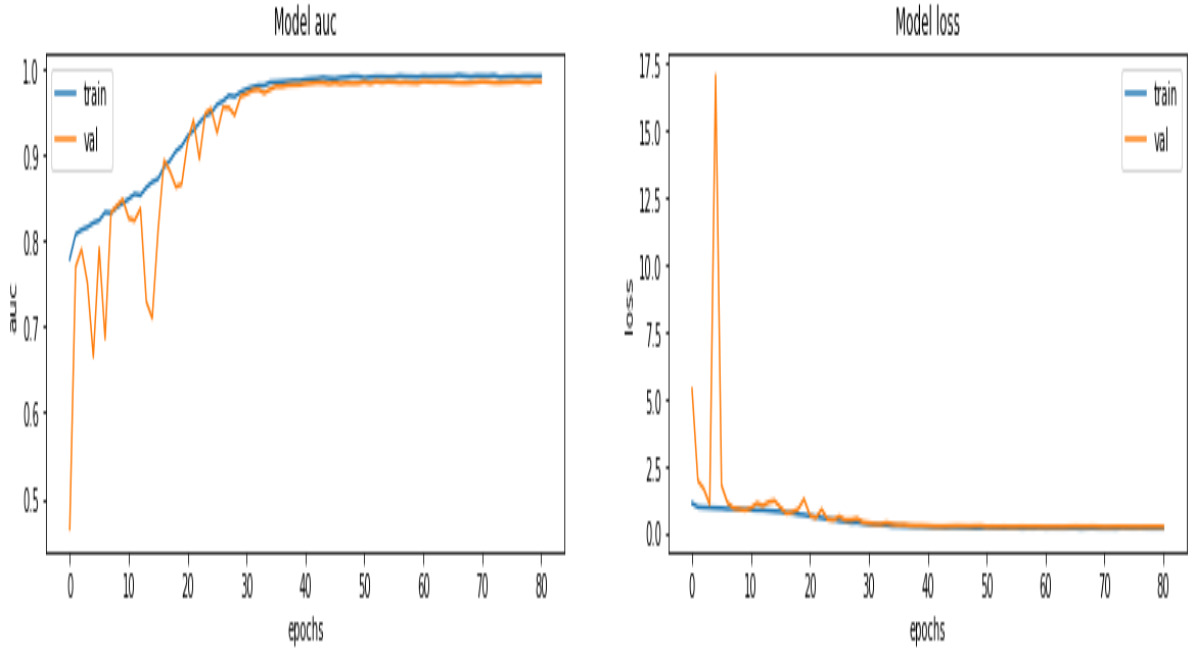

Figure 7.5. Model AUC and loss metric outcomes

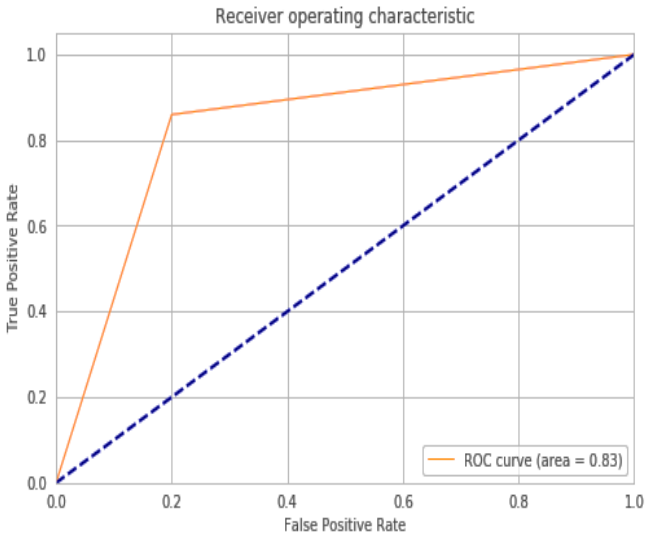

Figure 7.6. The ROC curves outcome for test data.

7.4. Discussion

In this work, we presented a novel deep learning type CNN model for the classification of AD subjects. As mentioned above, AD is the most common adult-onset dementia and contributes about 60-70% of worldwide dementia cases [164]. Unfortunately, there is no proper medication or cure for AD, and advancements in AD cure have been getting slow. Screening among people of AD risk given electronic health records (EHR) in preclinical stages may prompt early identification of AD pathology and to suggest better approaches for complying with the AD beginning. Current biomarkers of AD have required specimen collection (like serum or liquid), MRI image data or more sophisticated markers that at the present can be identified just in highly specialized centres.

On the other hand, the EHRs for example medical records in clinical settings, or administrative health information don't require extra time or effort for data collection. Likewise, with the coming of digitalization, the measures of such information have drastically increased [165]. Since it is omnipresent, enormous, and cost-effective, the digitized medical database might be a significant asset for testing different AD predictive models. Nonetheless, despite its enormous possible value, somehow thought about the degrees to which the enormous scope of EHR data can help in risk of AD prediction [165], [166]. The possible prediction of future AD progression is incredibly significant in clinical practice also, in healthcare research. Advanced neuroimaging techniques like MRI, positron emission tomography (PET) is developed and presented to identify AD-related molecular and structural biomarkers.

Computer scientists are recommending applying sophisticated computing techniques like machine learning and deep learning. For example, Battineni et al (2020) have achieved 99.1% accuracy through the application of ensemble learning models for late-life AD detection among 150 patients [167]. AD prediction among 123 subjects with Pre-MCI and MCI was done by clinically transmittable ML algorithms and results reported the whole sample accuracy of 96.2% [168]. However, most of the outcomes proposed by these algorithms are based on demographic magnetic resonance image (MRI) information. Because of this, researchers believed that deep learning algorithms are the best approaches if brain images were included [169]. Most of the works associated with Machine learning in the early prediction of AD occurred with high success. For instance, it is reported that 94.1% of accuracy by 3D convolutional neural networks (CNN) [170].

This work presented a deep CNN with 10-fold cross-validation and achieved more than 80% accuracy. While applying computing methods for diagnosis, a small portion of datasets are presented. Therefore, our model maintained a random image selection of train, test, and validation datasets. The proposed model produced promising results in AD image classification. The most notable outcome for this study is the progressions among predictiveness of AD diseases.

7.5. Chapter summary

An autonomous AD detection classifier based deep ConvNet framework is presented. We adopted the latest release of the OASIS-3 dataset that contains different categories of AD datasets. For training, more than 1,500 images model took a bit longer process than expected, but it is faster than mankind process. Deep ConvNets do not need any handcrafted feature selection approach because of having autonomous feature tuning. The main limitation of the study is to adopt only a single classifier for the brain MRI data classification and there are other possibilities to do better improvements in the proposed model architecture. Although attained results of higher 80% accuracy while compared over traditional ML classifiers, many advancements are proposed to enhance the model quality.

The results summarized in this chapter were published in Battineni G, Chintalapudi N, Amenta F, Traini E. *Deep learning type convolution neural network architecture for multiclass classification of Alzheimer's disease. BIOIMAGING 2021 - 8th Int Conf Bioimaging; Part 14th Int Jt Conf Biomed Eng Syst Technol BIOSTEC 2021. Published online 2021:209-215. doi:10.5220/0010378602090215*

Chapter 8
Conclusions

The role of computer techniques such as ML and deep learning algorithms in prediction of AD among older adults are successfully presented in this thesis.

We conducted five different works with two dementia datasets; one with demographic information of AD patients (OASIS-2) and other with brain MRI data (OASIS-3). Classification of AD patients with high risk was separated from adopting single ML modelling, multi modelling and hybrid modelling. For identification of true AD subjects with single ML modelling was done by support vector machines and decision tree algorithms. With SVM modelling proposed in Chapter 3, we achieved 75% of classification accuracy. This value has been improved by incorporating other popular supervised model named as decision trees. We achieved 88.73% of accuracy with J48 pruned decision trees.

In multi-modelling studies, we conducted model evaluation by involving different supervised algorithms such as Naïve Bayes, KNN, neural networks, and LR along with SVM model. In chapter 5, we presented comparative ML techniques with PCA algorithms to classify the positive AD subjects. KNN, and LR were significantly produced the higher accuracy values of 97.6% and 98.3% respectively. Besides, these numbers are relatively higher than studies with single modeling. However, the mentioned studies produced significant performance in AD subject classification, getting 100% of accuracy is ultimately important in prediction of serious diseases like dementia.

For that, in the chapter 6 we conducted three different experiments such as multi-modelling with manual feature selection, with automatic feature selection and hybrid modelling. Experiments with manual feature selection before automatic feature selection with 1NN produced 91.32% of accuracy, and the experiment of automatic feature selection generated 96.12% accuracy by SVM. This value significantly increased using multi modelling and produced 98% of accuracy. As mentioned, in disease classifications, computer scientists considered AU-ROC curve as the primary metric to evaluate the model performance. The ROC curve produced 99.91% of classification accuracy in prediction of true demented subjects. As of this, we would like to conclude that ensemble learning models with automatic feature selection are the optimal solutions for early AD prediction in older adults.

On the other hand, deep learning models like neural networks largely involved in image related studies. In chapter 8, we evaluated real time sample of 2168 MRI patient images with very mild to different stages of cognitive decline. We applied deep learning-based model called CNN to identify the dementia progression. Model training was done with 70% of images and achieved 83.3% which is considered as an optimal classifier for AD image detection and this value is significantly higher than traditional ML approaches.

Presented works having some limitations. Firstly, lowest number AD patient's data may hamper the hypothesis of results to the overall population of dementia patients. When compared with existed literature on AD predictions, we accomplish better model performance by involving different experimental strategies and methods. Arrangement and standardization of subject grouping are not precise much of the time, and it may keep an eye on underestimation of dementia in older patients that bring about getting low accuracies through SVM modelling. Moving towards ideal supervised models like decision trees, KNN, and logistic regression models, we attempted to improve the model performance by a selection of feature reduction techniques like pruning and PCA. Finally, it is important also to present the studies on real time brain image data that can produce clear picture of dementia progression in AD patients. This problem will plan to be addressed in future research works and also validate use of deep learning models in real time.

References

[1] C. Holmes en J. Amin, "Dementia", Medicine (United Kingdom). 2020, doi: 10.1016/j.mpmed.2020.08.014.

[2] S. A. Gale, D. Acar, en K. R. Daffner, "Dementia", American Journal of Medicine. 2018, doi: 10.1016/j.amjmed.2018.01.022.

[3] N. Sapkota en S. Subedi, "Dementia as a Public Health Priority", J. Psychiatr. Assoc. Nepal, 2019, doi: 10.3126/jpan.v8i2.28016.

[4] H. Brunnström, L. Gustafson, U. Passant, en E. Englund, "Prevalence of dementia subtypes: A 30-year retrospective survey of neuropathological reports", Arch. Gerontol. Geriatr., 2009, doi: 10.1016/j.archger.2008.06.005.

[5] J. Cerejeira, L. Lagarto, en E. B. Mukaetova-Ladinska, "Behavioral and psychological symptoms of dementia", Front. Neurol., 2012, doi: 10.3389/fneur.2012.00073.

[6] C. Miranda-Castillo, B. Woods, en M. Orrell, "The needs of people with dementia living at home from user, caregiver and professional perspectives: A cross-sectional survey", BMC Health Services Research. 2013, doi: 10.1186/1472-6963-13-43.

[7] H. C. Kales, L. N. Gitlin, en C. G. Lyketsos, "Assessment and management of behavioral and psychological symptoms of dementia", BMJ (Online). 2015, doi: 10.1136/bmj.h369.

[8] I. McKeith et al., "Dementia with Lewy bodies", Lancet Neurology. 2004, doi: 10.1016/S1474-4422(03)00619-7.

[9] C. G. Lyketsos et al., "Neuropsychiatric symptoms in Alzheimer's disease", Alzheimer's and Dementia. 2011, doi: 10.1016/j.jalz.2011.05.2410.

[10] J. Cohen-Mansfield, M. S. Marx, K. Thein, en M. Dakheel-Ali, "The impact of stimuli on affect in persons with dementia", J. Clin. Psychiatry, 2011, doi: 10.4088/JCP.09m05694oli.

[11] M. Wortmann, "Dementia: A global health priority - Highlights from an ADI and World Health Organization report", Alzheimer's Research and Therapy. 2012, doi: 10.1186/alzrt143.

[12] L. Razmerita, K. Kirchner, en P. Nielsen, "What factors influence knowledge sharing in organizations? A social dilemma perspective of social media communication", J. Knowl. Manag., 2016, doi: 10.1108/JKM-03-2016-0112.

[13] C. M. Connell, M. R. Janevic, en M. P. Gallant, "The costs of caring: Impact of dementia on family caregivers", J. Geriatr. Psychiatry Neurol., 2001, doi: 10.1177/089198870101400403.

[14] L. Y. Chien et al., "Caregiver support groups in patients with dementia: A meta-analysis", Int. J. Geriatr. Psychiatry, 2011, doi: 10.1002/gps.2660.

[15] V. Bellou, L. Belbasis, I. Tzoulaki, L. T. Middleton, J. P. A. Ioannidis, en E. Evangelou, "Systematic evaluation of the associations between environmental risk factors and dementia: An umbrella review of systematic reviews and meta-analyses", Alzheimer's Dement., 2017, doi: 10.1016/j.jalz.2016.07.152.

[16] M. Prince, A. Wimo, M. Guerchet, A. Gemma-Claire, Y.-T. Wu, en M. Prina, "World Alzheimer Report 2015: The Global Impact of Dementia - An analysis of prevalence, incidence, cost and trends", Alzheimer's Dis. Int., 2015, doi: 10.1111/j.0963-7214.2004.00293.x.

[17] J. Gaugler, B. James, T. Johnson, K. Scholz, en J. Weuve, "2016 Alzheimer's disease facts and figures", Alzheimer's Dement., 2016, doi: 10.1016/j.jalz.2016.03.001.

[18] J. T. O'Brien en A. Thomas, "Vascular dementia", The Lancet. 2015, doi: 10.1016/S0140-6736(15)00463-8.

[19] Z. Walker, K. L. Possin, B. F. Boeve, en D. Aarsland, "Lewy body dementias", The Lancet. 2015, doi: 10.1016/S0140-6736(15)00462-6.

[20] J. Bang, S. Spina, en B. L. Miller, "Frontotemporal dementia", The Lancet. 2015, doi: 10.1016/S0140-6736(15)00461-4.

[21] J. De Reuck, "Mixed dementia", in Diagnosis and Management in Dementia, 2020.

[22] G. M. McKhann et al., "The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease", Alzheimer's Dement., 2011, doi: 10.1016/j.jalz.2011.03.005.

[23] M. N. Rossor, N. C. Fox, C. J. Mummery, J. M. Schott, en J. D. Warren, "The diagnosis of young-onset dementia", The Lancet Neurology. 2010, doi: 10.1016/S1474-4422(10)70159-9.

[24] D. E. Barnes et al., "Development and Validation of eRADAR: A Tool Using EHR Data to Detect Unrecognized Dementia", J. Am. Geriatr. Soc., 2020, doi: 10.1111/jgs.16182.

[25] B. C. M. Stephan, T. Kurth, F. E. Matthews, C. Brayne, en C. Dufouil, "Dementia risk prediction in the population: Are screening models accurate?", Nature Reviews Neurology. 2010, doi: 10.1038/nrneurol.2010.54.

[26] Y. Tu, "Machine learning", in EEG Signal Processing and Feature Extraction, 2019.

[27] M. Kubat, An Introduction to Machine Learning. 2017.

[28] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques", Informatica (Ljubljana). 2007, doi: 10.1115/1.1559160.

[29] MathWorks, S. M. Chelly, C. Denis, en MathWorks, "Applying Supervised Learning", What is Mach. Learn., 2016.

[30] A. Coates, H. Lee, en A. Y. Ng, "An analysis of single-layer networks in unsupervised feature learning", 2011.

[31] T. Hofmann, "Unsupervised learning by probabilistic Latent Semantic Analysis", Mach. Learn., 2001, doi: 10.1023/A:1007617005950.

[32] A. K. Jain, "Data clustering: 50 years beyond K-means", Pattern Recognit. Lett., 2010, doi: 10.1016/j.patrec.2009.09.011.

[33] S. Bhatnagar, H. Prasad, en L. Prashanth, "Reinforcement learning", in Lecture Notes in Control and Information Sciences, 2013.

[34] L. P. Kaelbling, M. L. Littman, en A. W. Moore, "Reinforcement learning: A survey", J. Artif. Intell. Res., 1996, doi: 10.1613/jair.301.

[35] A. Voulodimos, N. Doulamis, A. Doulamis, en E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review", Computational Intelligence and Neuroscience. 2018, doi: 10.1155/2018/7068349.

[36] Y. Bengio en A. Courville, "Deep Learning of Representations", Intelligent Systems Reference Library. 2013, doi: 10.1007/978-3-642-36657-4_1.

[37] S. Raschka, "Model evaluation, model selection, and algorithm selection in machine learning", arXiv. 2018.

[38] C. Bergmeir en J. M. Benítez, "On the use of cross-validation for time series predictor evaluation", Inf. Sci. (Ny)., 2012, doi: 10.1016/j.ins.2011.12.028.

[39] M. Mittal, L. M. Goyal, S. Kaur, I. Kaur, A. Verma, en D. Jude Hemanth, "Deep learning based enhanced tumor segmentation approach for MR brain images", Appl. Soft Comput. J., 2019, doi: 10.1016/j.asoc.2019.02.036.

[40] M. A. Khan et al., "Gastrointestinal diseases segmentation and classification based on duo-deep architectures", Pattern Recognit. Lett., 2020, doi: 10.1016/j.patrec.2019.12.024.

[41] A. Mittal et al., "Detecting pneumonia using convolutions and dynamic capsule routing for chest X-ray images", Sensors (Switzerland), 2020, doi: 10.3390/s20041068.

[42] S. Wang en R. M. Summers, "Machine learning and radiology", Medical Image Analysis. 2012, doi: 10.1016/j.media.2012.02.005.

[43] B. Kaur, M. Sharma, M. Mittal, A. Verma, L. M. Goyal, en D. J. Hemanth, "An improved salient object detection algorithm combining background and foreground connectivity for brain image analysis", Comput. Electr. Eng., 2018, doi: 10.1016/j.compeleceng.2018.08.018.

[44] M. Mittal et al., "An efficient edge detection approach to provide better edge

connectivity for image analysis", IEEE Access, 2019, doi: 10.1109/ACCESS.2019.2902579.

[45] F. Coenen, "Data mining: Past, present and future", Knowl. Eng. Rev., 2011, doi: 10.1017/S0269888910000378.

[46] R. O. Alabi et al., "Comparison of supervised machine learning classification techniques in prediction of locoregional recurrences in early oral tongue cancer", Int. J. Med. Inform., 2020, doi: 10.1016/j.ijmedinf.2019.104068.

[47] M. Takamatsu et al., "Prediction of early colorectal cancer metastasis by machine learning using digital slide images", Comput. Methods Programs Biomed., 2019, doi: 10.1016/j.cmpb.2019.06.022.

[48] B. K. Singh, "Determining relevant biomarkers for prediction of breast cancer using anthropometric and clinical features: A comparative investigation in machine learning paradigm", Biocybern. Biomed. Eng., vol 39, no 2, bll 393–409, 2019, doi: 10.1016/j.bbe.2019.03.001.

[49] R. Bhardwaj en N. Hooda, "Prediction of Pathological Complete Response after Neoadjuvant Chemotherapy for breast cancer using ensemble machine learning", Informatics Med. Unlocked, vol 16, no May, bl 100219, 2019, doi: 10.1016/j.imu.2019.100219.

[50] L. Tapak, N. Shirmohammadi-Khorram, P. Amini, B. Alafchi, O. Hamidi, en J. Poorolajal, "Prediction of survival and metastasis in breast cancer patients using machine learning classifiers", Clin. Epidemiol. Glob. Heal., no September, bll 1–7, 2018, doi: 10.1016/j.cegh.2018.10.003.

[51] N. G. Forouhi, A. Misra, V. Mohan, R. Taylor, en W. Yancy, "Dietary and nutritional approaches for prevention and management of type 2 diabetes", BMJ, 2018, doi: 10.1136/bmj.k2234.

[52] P. Sajda, "MACHINE LEARNING FOR DETECTION AND DIAGNOSIS OF DISEASE", Annu. Rev. Biomed. Eng., 2006, doi: 10.1146/annurev.bioeng.8.061505.095802.

[53] N. Barakat, A. P. Bradley, en M. N. H. Barakat, "Intelligible support vector machines for diagnosis of diabetes mellitus", IEEE Trans. Inf. Technol. Biomed., 2010, doi: 10.1109/TITB.2009.2039485.

[54] Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, en H. Tang, "Predicting Diabetes Mellitus With Machine Learning Techniques", Front. Genet., 2018, doi: 10.3389/fgene.2018.00515.

[55] D. Sisodia en D. S. Sisodia, "Prediction of Diabetes using Classification Algorithms", 2018, doi: 10.1016/j.procs.2018.05.122.

[56] G. Battineni, G. G. Sagaro, C. Nalini, F. Amenta, en S. K. Tayebati, "Comparative machine-learning approach: A follow-up study on type 2 diabetes predictions by

cross-validation methods", Machines, vol 7, no 4, bll 1–11, 2019, doi: 10.3390/machines7040074.

[57] J. Chaki, S. Thillai Ganesh, S. K. Cidham, en S. Ananda Theertan, "Machine learning and artificial intelligence based Diabetes Mellitus detection and self-management: A systematic review", Journal of King Saud University - Computer and Information Sciences. King Saud bin Abdulaziz University, Jul 04, 2020, doi: 10.1016/j.jksuci.2020.06.013.

[58] Y. Chen et al., "Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B", Comput. Biol. Med., 2017, doi: 10.1016/j.compbiomed.2017.07.012.

[59] Z. Jiang et al., "Support vector machine-based feature selection for classification of liver fibrosis grade in chronic hepatitis C", J. Med. Syst., 2006, doi: 10.1007/s10916-006-9023-2.

[60] A. M. Hashem, M. E. M. Rasmy, K. M. Wahba, en O. G. Shaker, "Single stage and multistage classification models for the prediction of liver fibrosis degree in patients with chronic hepatitis C infection", Comput. Methods Programs Biomed., 2012, doi: 10.1016/j.cmpb.2011.10.005.

[61] H. L. Chen, D. Y. Liu, B. Yang, J. Liu, en G. Wang, "A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis", Expert Syst. Appl., 2011, doi: 10.1016/j.eswa.2011.03.066.

[62] R. Stoean, C. Stoean, M. Lupsor, H. Stefanescu, en R. Badea, "Evolutionary-driven support vector machines for determining the degree of liver fibrosis in chronic hepatitis C", Artif. Intell. Med., 2011, doi: 10.1016/j.artmed.2010.06.002.

[63] K. Polat en S. Güneş, "Prediction of hepatitis disease based on principal component analysis and artificial immune recognition system", Appl. Math. Comput., 2007, doi: 10.1016/j.amc.2006.12.010.

[64] "WHO | The Atlas of Heart Disease and Stroke". https://www.who.int/cardiovascular_diseases/resources/atlas/en/ (toegang verkry Jul 27, 2020).

[65] S. S. Virani et al., "Heart disease and stroke statistics—2020 update: A report from the American Heart Association", Circulation. Lippincott Williams and Wilkins, bll E139–E596, 2020, doi: 10.1161/CIR.0000000000000757.

[66] S. D. Desai, S. Giraddi, P. Narayankar, N. R. Pudakalakatti, en S. Sulegaon, "Back-propagation neural network versus logistic regression in heart disease classification", 2019, doi: 10.1007/978-981-13-0680-8_13.

[67] H. Ahmed, E. M. G. Younis, A. Hendawi, en A. A. Ali, "Heart disease identification from patients' social posts, machine learning solution on Spark", Futur. Gener.

Comput. Syst., 2019, doi: 10.1016/j.future.2019.09.056.

[68] I. K. A. Enriko, M. Suryanegara, en D. Gunawan, "Heart disease prediction system using k-Nearest neighbor algorithm with simplified patient's health parameters", J. Telecommun. Electron. Comput. Eng., 2016.

[69] A. L. Chau, X. Li, en W. Yu, "Support vector machine classification for large datasets using decision tree and Fisher linear discriminant", Futur. Gener. Comput. Syst., 2014, doi: 10.1016/j.future.2013.06.021.

[70] S. Maji en S. Arora, "Decision Tree Algorithms for Prediction of Heart Disease", in Lecture Notes in Networks and Systems, 2019.

[71] T. Nguyen, A. Khosravi, D. Creighton, en S. Nahavandi, "Classification of healthcare data using genetic fuzzy logic system and wavelets", Expert Syst. Appl., 2015, doi: 10.1016/j.eswa.2014.10.027.

[72] G. Manogaran, R. Varatharajan, en M. K. Priyan, "Hybrid Recommendation System for Heart Disease Diagnosis based on Multiple Kernel Learning with Adaptive Neuro-Fuzzy Inference System", Multimed. Tools Appl., 2018, doi: 10.1007/s11042-017-5515-y.

[73] S. Nazari, M. Fallah, H. Kazemipoor, en A. Salehipour, "A fuzzy inference- fuzzy analytic hierarchy process-based clinical decision support system for diagnosis of heart diseases", Expert Syst. Appl., 2018, doi: 10.1016/j.eswa.2017.11.001.

[74] C. Wang et al., "Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study", Lancet, 2018, doi: 10.1016/S0140-6736(18)30841-9.

[75] S. Lin, Q. Zhang, F. Chen, L. Luo, L. Chen, en W. Zhang, "Smooth Bayesian network model for the prediction of future high-cost patients with COPD", Int. J. Med. Inform., 2019, doi: 10.1016/j.ijmedinf.2019.03.017.

[76] P. J. Castaldi et al., "Machine Learning Characterization of COPD Subtypes: Insights From the COPDGene Study", Chest. 2020, doi: 10.1016/j.chest.2019.11.039.

[77] T. Goto, C. A. Camargo, M. K. Faridi, B. J. Yun, en K. Hasegawa, "Machine learning approaches for predicting disposition of asthma and COPD exacerbations in the ED", Am. J. Emerg. Med., 2018, doi: 10.1016/j.ajem.2018.06.062.

[78] M. Lanclus et al., "Machine Learning Algorithms Utilizing Functional Respiratory Imaging May Predict COPD Exacerbations", Acad. Radiol., 2019, doi: 10.1016/j.acra.2018.10.022.

[79] S. Bodduluri, J. D. Newell, E. A. Hoffman, en J. M. Reinhardt, "Registration-based lung mechanical analysis of chronic obstructive pulmonary disease (COPD) using a supervised machine learning framework", Acad. Radiol., 2013, doi:

10.1016/j.acra.2013.01.019.

[80] J. Zhang, S. Wang, J. Courteau, L. Chen, G. Guo, en A. Vanasse, "Feature-weighted survival learning machine for COPD failure prediction", Artif. Intell. Med., 2019, doi: 10.1016/j.artmed.2019.01.003.

[81] X. Liu et al., "Improving precision of glomerular filtration rate estimating model by ensemble learning", J. Transl. Med., vol 15, no 1, bll 1–5, 2017, doi: 10.1186/s12967-017-1337-y.

[82] G. Lei, G. Wang, C. Zhang, Y. Chen, en X. Yang, "Using machine learning to predict acute kidney injury after aortic arch surgery", J. Cardiothorac. Vasc. Anesth., 2020, doi: 10.1053/j.jvca.2020.06.007.

[83] Z. Obermeyer en E. J. Emanuel, "Predicting the Future — Big Data, Machine Learning, and Clinical Medicine", N. Engl. J. Med., 2016, doi: 10.1056/nejmp1606181.

[84] G. Hinton, "Deep learning-a technology with the potential to transform health care", JAMA - Journal of the American Medical Association. 2018, doi: 10.1001/jama.2018.11100.

[85] C. Patterson, "World Alzheimer Report 2018 - The state of the art of dementia research: New frontiers", Alzheimer's Dis. Int. London, UK, 2018, doi: 10.1103/PhysRevLett.78.4414.

[86] L. Liu, S. Zhao, H. Chen, en A. Wang, "A new machine learning method for identifying Alzheimer's disease", Simul. Model. Pract. Theory, 2020, doi: 10.1016/j.simpat.2019.102023.

[87] G. Battineni, N. Chintalapudi, F. Amenta, en E. Traini, "A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects", J. Clin. Med., vol 9, no 7, bl 2146, Jul 2020, doi: 10.3390/jcm9072146.

[88] G. Battineni, N. Chintalapudi, en F. Amenta, "Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM)", Informatics Med. Unlocked, 2019, doi: 10.1016/j.imu.2019.100200.

[89] B. Gopi, C. Nalini, en A. Francesco, "Late-Life Alzheimer's Disease (AD) Detection Using Pruned Decision Trees", Int. J. Brain Disord. Treat., 2020, doi: 10.23937/2469-5866/1410033.

[90] C. K. Fisher et al., "Machine learning for comprehensive forecasting of Alzheimer's Disease progression", Sci. Rep., 2019, doi: 10.1038/s41598-019-49656-2.

[91] J. De Velasco Oriol, E. E. Vallejo, K. Estrada, J. G. Taméz Peña, en T. A. s. Disease Neuroimaging Initiative, "Benchmarking machine learning models for late-onset Alzheimer's disease prediction from genomic data", BMC Bioinformatics, 2019,

doi: 10.1186/s12859-019-3158-x.

[92] R. Casanova et al., "Using high-dimensional machine learning methods to estimate an anatomical risk factor for Alzheimer's disease across imaging databases", Neuroimage, 2018, doi: 10.1016/j.neuroimage.2018.08.040.

[93] E. Moradi, A. Pepe, C. Gaser, H. Huttunen, en J. Tohka, "Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects", Neuroimage, 2015, doi: 10.1016/j.neuroimage.2014.10.002.

[94] S. Lahmiri en A. Shmuel, "Performance of machine learning methods applied to structural MRI and ADAS cognitive scores in diagnosing Alzheimer's disease", Biomed. Signal Process. Control, 2019, doi: 10.1016/j.bspc.2018.08.009.

[95] G. Martí-Juan, G. Sanroma-Guell, en G. Piella, "A survey on machine and statistical learning for longitudinal analysis of neuroimaging data in Alzheimer's disease", Computer Methods and Programs in Biomedicine. 2020, doi: 10.1016/j.cmpb.2020.105348.

[96] World Health Organization, Risk reduction of cognitive decline and dementia: WHO guidelines. 2019.

[97] D. Stamate et al., "A metabolite-based machine learning approach to diagnose Alzheimer-type dementia in blood: Results from the European Medical Information Framework for Alzheimer disease biomarker discovery cohort", Alzheimer's Dement. Transl. Res. Clin. Interv., 2019, doi: 10.1016/j.trci.2019.11.001.

[98] J. Zhou et al., "Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer's disease", Brain, 2010, doi: 10.1093/brain/awq075.

[99] C. Sorg et al., "Selective changes of resting-state networks in individuals at risk for Alzheimer's disease", Proc. Natl. Acad. Sci. U. S. A., 2007, doi: 10.1073/pnas.0708803104.

[100] K. Walters et al., "Predicting dementia risk in primary care: Development and validation of the Dementia Risk Score using routinely collected data", BMC Med., 2016, doi: 10.1186/s12916-016-0549-y.

[101] "New Tool May Predict Dementia Risk Up to 10 Years Later". https://www.medscape.com/viewarticle/873566 (toegang verkry Des 15, 2020).

[102] S. Rathore, M. Habes, M. A. Iftikhar, A. Shacklett, en C. Davatzikos, "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages", NeuroImage. 2017, doi: 10.1016/j.neuroimage.2017.03.057.

[103] M. Ewers, R. A. Sperling, W. E. Klunk, M. W. Weiner, en H. Hampel, "Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia",

Trends in Neurosciences. 2011, doi: 10.1016/j.tins.2011.05.005.

[104] J. D. Davis, R. A. Stern, en L. A. Flashman, "Cognitive and neuropsychiatric aspects of subclinical hypothyroidism: Significance in the elderly", Current Psychiatry Reports. 2003, doi: 10.1007/s11920-003-0073-6.

[105] D. E. Barnes et al., "Development and validation of a brief dementia screening indicator for primary care", Alzheimer's Dement., 2014, doi: 10.1016/j.jalz.2013.11.006.

[106] L. G. Exalto, C. P. Quesenberry, D. Barnes, M. Kivipelto, G. J. Biessels, en R. A. Whitmer, "Midlife risk score for the prediction of dementia four decades later", Alzheimer's Dement., 2014, doi: 10.1016/j.jalz.2013.05.1772.

[107] V. S. Nori, C. A. Hane, D. C. Martin, A. D. Kravetz, en D. M. Sanghavi, "Identifying incident dementia by applying machine learning to a very large administrative claims dataset", PLoS One, 2019, doi: 10.1371/journal.pone.0203246.

[108] J. Finkelstein en I. cheol Jeong, "Machine learning approaches to personalize early prediction of asthma exacerbations", Ann. N. Y. Acad. Sci., 2017, doi: 10.1111/nyas.13218.

[109] R. Dinga et al., "Predicting the naturalistic course of depression from a wide range of clinical, psychological, and biological data: a machine learning approach", Transl. Psychiatry, vol 8, no 1, bl 241, 2018, doi: 10.1038/s41398-018-0289-1.

[110] F. Davis, M. Gostine, B. Roberts, R. Risko, J. Cappelleri, en A. Sadosky, "Characterizing classes of fibromyalgia within the continuum of central sensitization syndrome", J. Pain Res., vol Volume 11, bll 2551–2560, 2018, doi: 10.2147/JPR.S147199.

[111] M. Feres, Y. Louzoun, S. Haber, M. Faveri, L. C. Figueiredo, en L. Levin, "Support vector machine-based differentiation between aggressive and chronic periodontitis using microbial profiles", Int. Dent. J., vol 68, no 1, bll 39–46, 2018, doi: 10.1111/idj.12326.

[112] G. Battineni, G. G. Sagaro, N. Chinatalapudi, en F. Amenta, "Applications of machine learning predictive models in the chronic disease diagnosis", Journal of Personalized Medicine. 2020, doi: 10.3390/jpm10020021.

[113] D. Ichikawa, T. Saito, W. Ujita, en H. Oyama, "How can machine-learning methods assist in virtual screening for hyperuricemia? A healthcare machine-learning approach", J. Biomed. Inform., vol 64, bll 20–24, 2016, doi: 10.1016/j.jbi.2016.09.012.

[114] P. Kaur, M. Sharma, en M. Mittal, "Big Data and Machine Learning Based Secure Healthcare Framework", Procedia Comput. Sci., vol 132, bll 1049–1059, 2018, doi: 10.1016/j.procs.2018.05.020.

[115] K. R. Kruthika, Rajeswari, en H. D. Maheshappa, "Multistage classifier-based approach for Alzheimer's disease prediction and retrieval", Informatics Med. Unlocked, vol 14, no November 2018, bll 34–42, 2019, doi: 10.1016/j.imu.2018.12.003.

[116] X. Liu, D. Tosun, M. W. Weiner, en N. Schuff, "Locally linear embedding (LLE) for MRI based Alzheimer's disease classification", Neuroimage, vol 83, bll 148–157, 2013, doi: 10.1016/j.neuroimage.2013.06.033.

[117] R. XIE, I. Khalil, S. Badsha, en M. Atiquzzaman, "Collaborative extreme learning machine with a confidence interval for P2P learning in healthcare", Comput. Networks, vol 149, bll 127–143, 2019, doi: 10.1016/j.comnet.2018.11.002.

[118] A. Kennedy en D. Inkpen, "Sentiment classification of movie reviews using contextual valence shifters", 2006, doi: 10.1111/j.1467-8640.2006.00277.x.

[119] J. Levman, T. Leung, P. Causer, D. Plewes, en A. L. Martel, "Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines", IEEE Trans. Med. Imaging, 2008, doi: 10.1109/TMI.2008.916959.

[120] A. Danek, B. Landwehrmeyer, A. Ludolph, S. Anderl-Straub, en M. Otto, "Predicting primary progressive aphasias with support vector machine approaches in structural MRI data", NeuroImage Clin., vol 14, bll 334–343, 2017, doi: 10.1016/j.nicl.2017.02.003.

[121] L. Sørensen en M. Nielsen, "Ensemble support vector machine classification of dementia using structural MRI and mini-mental state examination", J. Neurosci. Methods, vol 302, bll 66–74, 2018, doi: 10.1016/j.jneumeth.2018.01.003.

[122] P. W. Wang en C. J. Lin, "Support vector machines", in Data Classification: Algorithms and Applications, 2014.

[123] "Support Vector Machine — Introduction to Machine Learning Algorithms". https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47 (toegang verkry Apr 11, 2019).

[124] A. J. Smola en B. Schölkopf, "A tutorial on support vector regression", Statistics and Computing. 2004, doi: 10.1023/B:STCO.0000035301.49549.88.

[125] C. Chang, C. Lin, en T. Tieleman, "LIBSVM : A Library for Support Vector Machines", ACM Trans. Intell. Syst. Technol., 2008, doi: 10.1145/1961189.1961199.

[126] "Chapter 2 : SVM (Support Vector Machine) — Theory – Machine Learning 101 – Medium". https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72 (toegang verkry Apr 11, 2019).

[127] D. S. Marcus, A. F. Fotenos, J. G. Csernansky, J. C. Morris, en R. L. Buckner, "Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented

and Demented Older Adults", J. Cogn. Neurosci., vol 22, no 12, bll 2677–2684, Des 2010, doi: 10.1162/jocn.2009.21407.

[128] A. F. Fotenos, A. Z. Snyder, L. E. Girton, J. C. Morris, en R. L. Buckner, "Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD", Neurology, 2005, doi: 10.1212/01.WNL.0000154530.72969.11.

[129] B. Ridha en M. Rossor, "The mini mental state examination", Pract. Neurol., 2005, doi: 10.1111/j.1474-7766.2005.00333.x.

[130] J. C. Morris, "The Clinical Dementia Rating (CDR): Current version and scoring rules", Neurology, 2012, doi: 10.1212/wnl.43.11.2412-a.

[131] A. Abramovitch et al., "Body mass index in obsessive-compulsive disorder", J. Affect. Disord., 2019, doi: 10.1016/j.jad.2018.10.116.

[132] F. Er, P. Iscen, S. Sahin, N. Çinar, S. Karsidag, en D. Goularas, "Distinguishing age-related cognitive decline from dementias: A study based on machine learning algorithms", J. Clin. Neurosci., 2017, doi: 10.1016/j.jocn.2017.03.021.

[133] K. Hajian-Tilaki, "Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation", Caspian Journal of Internal Medicine. 2013.

[134] R. Kumar en A. Indrayan, "Receiver operating characteristic (ROC) curve for medical researchers", Indian Pediatrics. 2011, doi: 10.1007/s13312-011-0055-4.

[135] A. I. Schein en L. H. Ungar, "Active learning for logistic regression: An evaluation", Mach. Learn., 2007, doi: 10.1007/s10994-007-5019-5.

[136] J. Davis en M. Goadrich, "The relationship between Precision-Recall and ROC curves", 2006, doi: 10.1145/1143844.1143874.

[137] P. Garrard, V. Rentoumi, B. Gesierich, B. Miller, en M. L. Gorno-Tempini, "Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse", Cortex, 2014, doi: 10.1016/j.cortex.2013.05.008.

[138] E. Pellegrini et al., "Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review", Alzheimer's Dement. Diagnosis, Assess. Dis. Monit., 2018, doi: 10.1016/j.dadm.2018.07.004.

[139] F. Er, P. Iscen, S. Sahin, N. Çinar, S. Karsidag, en D. Goularas, "Distinguishing age-related cognitive decline from dementias: A study based on machine learning algorithms", J. Clin. Neurosci., 2017, doi: 10.1016/j.jocn.2017.03.021.

[140] C. R. Aditya en M. B. S. Pande, "Devising an interpretable calibrated scale to quantitatively assess the dementia stage of subjects with alzheimer's disease: A machine learning approach", Informatics Med. Unlocked, 2017, doi: 10.1016/j.imu.2016.12.004.

[141] G. Battineni, N. Chintalapudi, en F. Amenta, "Comparative machine learning approach in dementia patient classification using principal component analysis",

2020, doi: 10.5220/0009096907800784.
[142] I. Guyon en J. Weston, "An Introduction to Variable and Feature Selection Isabelle", Mach. Learn. Res., 2003, doi: 10.1023/A:1012487302797.
[143] Y. Saeys, I. Inza, en P. Larrañaga, "A review of feature selection techniques in bioinformatics", Bioinformatics. 2007, doi: 10.1093/bioinformatics/btm344.
[144] N. Long, D. Gianola, G. J. M. Rosa, K. A. Weigel, en S. Avendano, "Machine learning classification procedure for selecting SNPs in genomic selection: Application to early mortality in broilers", 2008, doi: 10.1159/000317279.
[145] I. Rish, "An empirical study of the naive Bayes classifier", IJCAI 2001 Work. Empir. methods Artif. Intell., 2001, doi: 10.1039/b104835j.
[146] G. Battineni, G. G. Sagaro, C. Nalini, F. Amenta, en S. K. Tayebati, "Comparative machine-learning approach: A follow-up study on type 2 diabetes predictions by cross-validation methods", Machines, 2019, doi: 10.3390/machines7040074.
[147] L. Bougrain, "Practical introduction to artificial neural networks", IFAC Proc. Vol., vol 37, no 15, bll 347–352, 2004, doi: 10.1016/s1474-6670(17)31048-0.
[148] N. Effendy et al., "Artificial neural network prediction on ultrasonic performance of bismuth-tellurite glass compositions", Integr. Med. Res., 2020, doi: 10.1016/j.jmrt.2020.09.107.
[149] C. Behm, L. Nolting, en A. Praktiknjo, "How to model European electricity load profiles using artificial neural networks", Appl. Energy, vol 277, no May, 2020, doi: 10.1016/j.apenergy.2020.115564.
[150] L. Dormehl, "What is an artificial neural network? Here's everything you need to know", EMERGING TECH, 2019. .
[151] S. Arlot en A. Celisse, "A survey of cross-validation procedures for model selection", Stat. Surv., 2010, doi: 10.1214/09-SS054.
[152] C. R. Rao en Y. Wu, "Linear model selection by cross-validation", J. Stat. Plan. Inference, 2005, doi: 10.1016/j.jspi.2003.10.004.
[153] G. Chandrashekar en F. Sahin, "A survey on feature selection methods", Comput. Electr. Eng., 2014, doi: 10.1016/j.compeleceng.2013.11.024.
[154] D. E. Goldberg en J. H. Holland, "Genetic Algorithms and Machine Learning", Machine Learning. 1988, doi: 10.1023/A:1022602019183.
[155] E. Bauer en R. Kohavi, "Empirical comparison of voting classification algorithms: bagging, boosting, and variants", Mach. Learn., 1999, doi: 10.1023/A:1007515423169.
[156] J. Huang en C. X. Ling, "Using AUC and accuracy in evaluating learning algorithms", IEEE Trans. Knowl. Data Eng., 2005, doi: 10.1109/TKDE.2005.50.
[157] Alzheimer's Disease International (ADI), "World Alzheimer Report 2010: The

Global Economic Impact of Dementia", Alzheimer's Dis. Int. (ADI), 2010, doi: 10.1111/j.0963-7214.2004.00293.x.

[158] I. Jaussent et al., "Excessive sleepiness is predictive of cognitive decline in the elderly", Sleep, 2012, doi: 10.5665/sleep.2070.

[159] P. Johnson et al., "Genetic algorithm with logistic regression for prediction of progression to Alzheimer's disease", BMC Bioinformatics, 2014, doi: 10.1186/1471-2105-15-S16-S11.

[160] E. M. Sweeney et al., "OASIS is Automated Statistical Inference for Segmentation, with applications to multiple sclerosis lesion segmentation in MRI", NeuroImage Clin., 2013, doi: 10.1016/j.nicl.2013.03.002.

[161] L. Palumbo et al., "Evaluation of the intra- and inter-method agreement of brain MRI segmentation software packages: A comparison between SPM12 and FreeSurfer v6.0", Phys. Medica, 2019, doi: 10.1016/j.ejmp.2019.07.016.

[162] A. Krizhevsky, I. Sutskever, en G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Commun. ACM, 2017, doi: 10.1145/3065386.

[163] D. Pan, A. Zeng, L. Jia, Y. Huang, T. Frizzell, en X. Song, "Early Detection of Alzheimer's Disease Using Magnetic Resonance Imaging: A Novel Approach Combining Convolutional Neural Networks and Ensemble Learning", Front. Neurosci., 2020, doi: 10.3389/fnins.2020.00259.

[164] A. Khan en S. Zubair, "An Improved Multi-Modal based Machine Learning Approach for the Prognosis of Alzheimer's disease", J. King Saud Univ. - Comput. Inf. Sci., 2020, doi: 10.1016/j.jksuci.2020.04.004.

[165] Y. Shao, Q. T. Zeng, K. K. Chen, A. Shutes-David, S. M. Thielke, en D. W. Tsuang, "Detection of probable dementia cases in undiagnosed patients using structured and unstructured electronic health records", BMC Med. Inform. Decis. Mak., 2019, doi: 10.1186/s12911-019-0846-4.

[166] M. A. Mayer et al., "Reuse of EHRs to Support Clinical Research in a Hospital of Reference", 2015, doi: 10.3233/978-1-61499-512-8-224.

[167] G. Battineni, N. Chintalapudi, F. Amenta, en E. Traini, "A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects", J. Clin. Med., 2020, doi: 10.3390/jcm9072146.

[168] M. Grassi, G. Perna, D. Caldirola, K. Schruers, R. Duara, en D. A. Loewenstein, "A clinically-translatable machine learning algorithm for the prediction of Alzheimer's disease conversion in individuals with mild and premild cognitive impairment", J. Alzheimer's Dis., 2018, doi: 10.3233/JAD-170547.

[169] H. Choi en K. H. Jin, "Predicting cognitive decline with deep learning of brain

metabolism and amyloid imaging", Behav. Brain Res., 2018, doi: 10.1016/j.bbr.2018.02.017.

[170] S. Esmaeilzadeh, D. I. Belivanis, K. M. Pohl, en E. Adeli, "End-to-end alzheimer's disease diagnosis and biomarker identification", 2018, doi: 10.1007/978-3-030-00919-9_39.

[171] V. Jayaraman en H. P. Sultana, "Artificial gravitational cuckoo search algorithm along with particle bee optimized associative memory neural network for feature selection in heart disease classification", J. Ambient Intell. Humaniz. Comput., 2019, doi: 10.1007/s12652-019-01193-6.

[172] C. B. Gokulnath en S. P. Shantharajah, "An optimized feature selection based on genetic approach and support vector machine for heart disease", Cluster Comput., 2019, doi: 10.1007/s10586-018-2416-4.

Annex A- Abbreviations

- AD - Alzheimer's Disease
- KNN- K-nearest neighbours
- LR- logistic regression
- SVM- Support Vector Machine
- PCA- principal component analysis
- AUC- area under the curve
- GDP- Gross Domestic Product
- APOE- Apolipoprotein E
- EHR- electronic health records
- BMI- body mass index
- ML- machine learning
- AI- Artificial Intelligence
- ANN- artificial neural network
- CV- cross-validation
- DNN- Deep neural networks
- CT- computed tomography
- CNN- convolution neural networks
- IT- Information Technology
- IVD- in-vitro medical diagnostics
- OTSCC- oral tongue squamous cell carcinoma
- NB- Naive Bayes
- BDT- Boosted Decision Tree
- DF- Decision Forest
- CRC- colorectal cancer
- RF- random forecast
- WHO- World Health Organization
- RTE- Real-time tissue elastography
- LFI- liver fibrosis index
- SFFS- sequential forward floating selection
- LFDA- local Fisher discriminant analysis
- FDA-Fisher discriminant analysis
- BPNN- back proportion neural networks
- FAHP- fuzzy analytic hierarchy process
- COPD- Chronic obstructive pulmonary disease

- SBN- smooth Bayesian network
- GFR- glomerular filtration rate
- AKI- acute kidney injury
- FM- Fibromyalgia
- CRBM- Conditional Restricted Boltzmann Machine
- MRI- magnetic resonance imaging
- p-MCI- progressive Mild Cognitive Impairment
- PET- positron emission tomography
- XG-Boost- extreme gradient boosting
- CSF- cerebrospinal fluid
- RQ- research questions
- ROC- receiver operating characteristic
- LLE- locally linear embedding
- RBF- radial basis function
- PPA- Primary progressive aphasia
- MMSE- mini-mental state examination
- ND- non-demented
- D- demented
- C- converted
- ADRC- Alzheimer's disease research Centre
- MP-RAGE- Magnetization Prepared Rapid Acquired Gradient Echo
- n-WBV- Normalized Whole-brain Volume
- CDR- clinical dementia ratio
- OASIS- Open Access Series of Imaging Studies
- MCI- Mild Cognitive Impairment
- SD- standard deviation
- ARCD- age-related cognitive decline

Annex B- Publications

1. Battineni G, Chintalapudi N, Amenta F, Traini E. A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects. J Clin Med. 2020;9(7):2146. doi:10.3390/jcm9072146

2. Battineni G, Chintalapudi N, Amenta F. Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). Informatics Med Unlocked. Published online 2019. doi:10.1016/j.imu.2019.100200

3. Battineni G, Sagaro GG, Chinatalapudi N, Amenta F. Applications of machine learning predictive models in the chronic disease diagnosis. J Pers Med. Published online 2020. doi:10.3390/jpm10020021

4. Gopi B, Nalini C, Francesco A. Late-Life Alzheimer's Disease (AD) Detection Using Pruned Decision Trees. Int J Brain Disord Treat. 2020;6(1). doi:10.23937/2469-5866/1410033

5. Battineni G, Chintalapudi N, Amenta F. Comparative machine learning approach in dementia patient classification using principal component analysis. In: ICAART 2020 - Proceedings of the 12th International Conference on Agents and Artificial Intelligence. ; 2020. doi:10.5220/0009096907800784

6. Chintalapudi N, Battineni G, Amenta F. COVID-19 virus outbreak forecasting of registered and recovered cases after sixty day lockdown in Italy: A data driven model approach. J Microbiol Immunol Infect. 2020;53(3):396-403. doi:10.1016/j.jmii.2020.04.004

7. Battineni G, Chintalapudi N, Amenta F. AI Chatbot Design during an Epidemic Like the Novel Coronavirus. Healthcare. Published online 2020. doi:10.3390/healthcare8020154

8. Chintalapudi N, Battineni G, Sagaro GG, Amenta F. COVID-19 outbreak reproduction number estimations and forecasting in Marche, Italy. Int J Infect Dis. Published online 2020. doi:10.1016/j.ijid.2020.05.029

9. Nittari G, Khuman R, Baldoni S, et al. Telemedicine Practice: Review of the Current Ethical and Legal Challenges. Telemed e-Health. Published online 2020. doi:10.1089/tmj.2019.0158

10. Battineni G, Mittal M, Mohan Goyal L, et al. Cloud-based framework to mitigate

the impact of COVID-19 on seafarers' mental health. Int Marit Health. 2020;71(3):213-214. doi:10.5603/IMH.2020.0038

11. Battineni G, Baldoni S, Chintalapudi N, et al. Factors affecting the quality and reliability of online health information. Digit Heal. 2020;6:1-11. doi:10.1177/2055207620948996

12. Nittari G, Arcese A, Battineni G, et al. Design and evolution of the Seafarer's Health Passport for supporting (tele)-medical assistance to seafarers. Int Marit Health. Published online 2019. doi:10.5603/IMH.2019.0024

13. Battineni G, Chintalapudi N, Amenta F. Model discovery, and replay fitness validation using inductive mining techniques in medical training of CVC surgery. Appl Comput Informatics. Published online January 18, 2020. doi:10.1016/j.aci.2020.01.001

14. Battineni G, Canio MD, Chintalapudi N, Amenta F, Nittari G. Development of physical training smartphone application to maintain fitness levels in seafarers. Int Marit Health. 2019;70(3). doi:10.5603/IMH.2019.0028

15. Battineni G, Chintalapudi N, Amenta F, Tayebati SK. Report on market analysis and preventions need to provide medications for rural patients of Italy using ICT technologies. Int J Innov Technol Explor Eng. Published online 2019. doi:10.35940/ijitee.A4025.119119

16. Mirmoeini SM, Shooshtari SSM, Battineni G, Amenta F, Tayebati SK. Policies and challenges on the distribution of specialists and subspecialists in rural areas of Iran. Med. Published online 2019. doi:10.3390/medicina55120783

17. Battineni G, Sagaro GG, Nalini C, Amenta F, Tayebati SK. Comparative machine-learning approach: A follow-up study on type 2 diabetes predictions by cross-validation methods. Machines. Published online 2019. doi:10.3390/machines7040074

18. Battineni G, Chintalapudi N, Amenta F. SARS-CoV-2 epidemic calculation in Italy by SEIR compartmental models. Appl Comput Informatics. 2020;ahead-of-p(ahead-of-print). doi:10.1108/ACI-09-2020-0060

19. Nittari G, Pallotta G, Battineni G, et al. Comparative analysis of the medicinal compounds of the ship's "medicine chests" in European Union maritime countries. Need for improvement and harmonization. Int Marit Health. Published online 2019. doi:10.5603/IMH.2019.0023

20. Sagaro GG, Battineni G, Chintalapudi N, Di Canio M, Amenta F. Telemedical assistance at sea in the time of COVID-19 pandemic. Int Marit Health. 2020;71(4):229-236. doi:10.5603/IMH.2020.0041

21. Battineni G, Chintalapudi N, Amenta F. Is Really Weather or Tropical Conditions Affects an Outbreak of COVID-19: A Brief Report. Pharm Biomed Res. Published online 2020.

22. Chintalapudi N, Battineni G, Canio M Di, Sagaro GG, Amenta F. Text mining with sentiment analysis on seafarers' medical documents. Int J Inf Manag Data Insights. 2021;1(1):100005. doi:10.1016/j.jjimei.2020.100005

23. Battineni G, Chintalapudi N, Amenta F. Forecasting of COVID-19 epidemic size in four high hitting nations (USA, Brazil, India and Russia) by Fb-Prophet machine learning model. Appl Comput Informatics. Published online 2020. doi:10.1108/aci-09-2020-0059

24. Battineni G, Amenta F. Designing of an Expert system for the management of Seafarer's health. Digit Heal. 2020;1(6):6-8. doi:10.1177/2055207620976244

25. Sagaro GG, Battineni G, Amenta F. Barriers to Sustainable Telemedicine Implementation in Ethiopia: A Systematic Review. Telemed Reports. 2020;1(1):8-15. doi:10.1089/tmr.2020.0002

26. Mittal M, Battineni G, Kumar P, Sharma P, Panwar A. IoT based image defogging system for road accident control during winters. In: Proceedings of the 2020 International Conference on Computing, Communication and Security, ICCCS 2020. Institute of Electrical and Electronics Engineers Inc.; 2020. doi:10.1109/ICCCS49678.2020.9277430

27. Battineni G, Sagaro GG, Chintalapudi N, Amenta F, Schwartz SG. Conceptual Framework and Designing for a Seafarers' Health Observatory (SHO) Based on the Centro Internazionale Radio Medico (C.I.R.M.) Data Repository. Sci World J. Published online 2020. doi:10.1155/2020/8816517

28. Mirmoeini SM, Shooshtari SSM, Battineni G, Amenta F, Tayebati SK. Telepediatric assistance in Iran: Specialist and subspecialty challenges. EAI Endorsed Trans Pervasive Heal Technol. 2020;6(23). doi:10.4108/eai.22-9-2020.166356

29. Battineni G, Chintalapudi N, Karami V, Nittari G, Amenta F, Tayabati SK. Process mining case study approach: Extraction of unconventional event logs to improve performance in Hospital Information Systems (HIS) Corresponding Author. Int J Comput Sci Inf Secur. 2019;17(4):117. https://sites.google.com/site/ijcsis/

30. Chhetri B, Goyal LM, Mittal M, Battineni G. Estimating the prevalence of stress among Indian students during the COVID-19 pandemic: A cross-sectional study from India. J Taibah Univ Med Sci. Published online January 18, 2021. doi:10.1016/j.jtumed.2020.12.012

31. Battineni G, Pallotta G, Nittari G, Amenta F. Telemedicine framework to mitigate the impact of the COVID-19 pandemic. J Taibah Univ Med Sci. Published online January 14, 2021. doi:10.1016/j.jtumed.2020.12.010

32. Battineni G, Sagaro GG, Chintalapudi N, Di Canio M, Amenta F. Assessment of Awareness and Knowledge on Novel Coronavirus (COVID-19) Pandemic among Seafarers. Healthcare. 2021;9(2):120. doi:10.3390/healthcare9020120

33. Amenta F, Battineni G, Traini E, Pallotta G. Choline-containing phospholipids and treatment of adult-onset dementia disorders. In: Diagnosis and Management in Dementia. ; 2020. doi:10.1016/b978-0-12-815854-8.00030-6

34. Battineni G, Chintalapudi N, Amenta F. Epidemic Models in Prediction of COVID-19. In: Predictive and Preventive Measures for Covid-19 Pandemic. Springer, Singapore; 2021:19-34. doi:10.1007/978-981-33-4236-1_2

35. Nittari G, Battineni G, Messinetti M, Campanozzi L, Sirignano A. Critical reflections and solutions for health problems of Italian refugees. Clin Ter. 2021;172(2):158-162. doi:10.7417/CT.2021.2304

36. Battineni G, Nittari G, Sirignano A, Amenta F. Are telemedicine systems effective healthcare solutions during the COVID-19 pandemic? J Taibah Univ Med Sci. 2021;16(3):305-306. doi:10.1016/J.JTUMED.2021.02.009

37. Sagaro GG, Dicanio M, Battineni G, Samad MA, Amenta F. Incidence of occupational injuries and diseases among seafarers: A descriptive epidemiological study based on contacts from onboard ships to the Italian Telemedical Maritime Assistance Service in Rome, Italy. BMJ Open. 2021;11(3). doi:10.1136/BMJOPEN-2020-044633

38. Battineni G, Kumar S, Mittal M, Amenta F. COVID-19 vaccine on board ships: Current and future implications of seafarers. Int Marit Health. 2021;72(1):76-77. doi:10.5603/IMH.2021.0010

39. Chintalapudi N, Battineni G, Canio M Di, Sagaro GG, Amenta F. Text mining with sentiment analysis on seafarers' medical documents. Int J Inf Manag Data Insights. 2021;1(1):100005. doi:10.1016/J.JJIMEI.2020.100005

40. Battineni G, Mittal M, Jain S. Data visualization in the transformation of healthcare industries. Lect Notes Data Eng Commun Technol. 2021;64:1-23. doi:10.1007/978-981-16-0538-3_1

41. Chintalapudi N, Battineni G, Amenta F. Sentimental analysis of COVID-19 tweets using deep learning models. Infect Dis Rep. 2021;13(2). doi:10.3390/IDR13020032

42. Kansal AK, Gautam J, Chintalapudi N, Jain S, Battineni G. Google trend analysis and paradigm shift of online education platforms during the COVID-19 pandemic. Infect Dis Rep. 2021;13(2):418-428. doi:10.3390/IDR13020040

43. Battineni G, Pallotta G, Nittari G, Chintalapudi N, Varlaro V, Amenta F. Development of quality assessment tool for websites of the international aesthetic medicine societies. Informatics Med Unlocked. 2021;23. doi:10.1016/J.IMU.2021.100559

44. Battineni G, Sagaro GG, Chintalapudi N, Amenta F, Tomassoni D, Tayebati SK. Impact of obesity-induced inflammation on cardiovascular diseases (Cvd). Int J Mol Sci. 2021;22(9). doi:10.3390/IJMS22094798

45. Battineni G. Machine Learning and Deep Learning Algorithms in the Diagnosis of Chronic Diseases. Stud Comput Intell. 2021;968:141-164. doi:10.1007/978-981-16-0935-0_7

46. Chintalapudi N, Battineni G, Amenta F. Second wave of COVID-19 in Italy: Preliminary estimation of reproduction number and cumulative case projections. Results Phys. 2021;28. doi:10.1016/J.RINP.2021.104604

47. Mittal M, Battineni G, Singh D, Nagarwal T, Yadav P. Web-based chatbot for Frequently Asked Queries (FAQ) in Hospitals. J Taibah Univ Med Sci. 2021;16(5):740-746. doi:10.1016/J.JTUMED.2021.06.002

48. Sagaro GG, Battineni G, Di Canio M, Amenta F. Self-reported modifiable risk factors of cardiovascular disease among seafarers: A cross-sectional study of prevalence and clustering. J Pers Med. 2021;11(6). doi:10.3390/JPM11060512

49. Chakraborty T, Jamal RF, Battineni G, Teja KV, Marto CM, Spagnuolo G. A review of prolonged post-covid-19 symptoms and their implications on dental management. Int J Environ Res Public Health. 2021;18(10). doi:10.3390/IJERPH18105131

50. Savva D, Battineni G, Amenta F, Nittari G. Hypersensitivity reaction to hyaluronic

acid dermal filler after the Pfizer vaccination against SARS-CoV-2 virus. Int J Infect Dis. Published online September 2021. doi:10.1016/J.IJID.2021.09.066

51. Teja K V., Ramesh S, Vasundhara KA, Janani KC, Jose J, Battineni G. A new innovative automated root canal device for syringe needle irrigation. J Taibah Univ Med Sci. Published online August 2021. doi:10.1016/J.JTUMED.2021.07.011

52. Carotenuto A, Traini E, Fasanaro AM, Battineni G, Amenta F. Tele-neuropsychological assessment of Alzheimer's disease. J Pers Med. 2021;11(8). doi:10.3390/JPM11080688

53. Battineni G, Sagaro GG, Chintalapudi N, Amenta F. The benefits of telemedicine in personalized prevention of cardiovascular diseases (CVD): A systematic review. J Pers Med. 2021;11(7). doi:10.3390/JPM11070658

54. Battineni G, Nittari G, Pallotta G, Sagaro GG, Chintalapudi N, Amenta F. Telehealth and Pharmacological Strategies of COVID-19 Prevention: Current and Future Developments. Published online 2022:897-927. doi:10.1007/978-3-030-72834-2_26

55. Mittal M, Battineni G, Ahmad W, Kumar N, Upreti R. Mathematical Scanner (M-Scan) Mobile Application for Solving Simple Math Equations. Published online 2022:345-353. doi:10.1007/978-981-16-2597-8_29

56. Badr C, Spagnuolo G, Amenta F, et al. A two-year comparative evaluation of the clinical performance of a nanohybrid composite resin to a flowable composite resin. J Funct Biomater. 2021;12(3). doi:10.3390/JFB12030051

ANNEX C

Published works on AD diagnosis through ML techniques

Journal of *Personalized Medicine*

MDPI

Review

Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis

Gopi Battineni [1,*], Getu Gamo Sagaro [1], Nalini Chinatalapudi [1] and Francesco Amenta [1,2]

[1] Center for Telemedicine and Tele pharmacy, School of Medicinal and Health Sciences Products, University of Camerino, Via Madonna Della carceri 9, 62032 Camerino, Italy; getugamo.sagaro@unicam.it (G.G.S.); nalini.chintalapudi@unicam.it (N.C.); francesco.amenta@unicam.it (F.A.)

[2] Research Department, International Medical Radio Center Foundation (C.I.R.M.), 00144 Roma, Italy

* Correspondence: gopi.battineni@unicam.it; Tel.: +39-333-172-8206

Received: 10 February 2020; Accepted: 23 March 2020; Published: 31 March 2020

Abstract: This paper reviews applications of machine learning (ML) predictive models in the diagnosis of chronic diseases. Chronic diseases (CDs) are responsible for a major portion of global health costs. Patients who suffer from these diseases need lifelong treatment. Nowadays, predictive models are frequently applied in the diagnosis and forecasting of these diseases. In this study, we reviewed the state-of-the-art approaches that encompass ML models in the primary diagnosis of CD. This analysis covers 453 papers published between 2015 and 2019, and our document search was conducted from PubMed (Medline), and Cumulative Index to Nursing and Allied Health Literature (CINAHL) libraries. Ultimately, 22 studies were selected to present all modeling methods in a precise way that explains CD diagnosis and usage models of individual pathologies with associated strengths and limitations. Our outcomes suggest that there are no standard methods to determine the best approach in real-time clinical practice since each method has its advantages and disadvantages. Among the methods considered, support vector machines (SVM), logistic regression (LR), clustering were the most commonly used. These models are highly applicable in classification, and diagnosis of CD and are expected to become more important in medical practice in the near future.

Keywords: chronic diseases; prediction models; pathologies; accuracy; disease classification

1. Introduction

Artificial intelligence (AI) is defined as the technology that uses computer knowledge to represent intelligent behavior with nominal human involvement, and machine learning (ML) is considered as a subset of AI techniques. Usually, this kind of intelligence is commonly acknowledged as having begun with the innovation of robotics [1]. With the rapid growth of electronic speeds and programming, computers may display intelligent behavior similar to that of humans in the near future. This is because of the large advancements happening in contemporary ideas in the development of AI [2]. Artificial intelligence can be defined as human intelligence which is performed by machines. In computer science, it is defined as the machine's capacity to emulate intelligent behavior by itself, using nothing but ML [3].

The applications of AI in medicine are developing quickly. In 2016, AI projects coupled with medicine drew in more speculation from the global economy than other projects [4]. In medicine, AI refers to the utilization of automated diagnosis processes and the treatment of patients who require care. Increased AI utilization in prescription will allow a considerable amount of the role to be automated, opening up medicinal experts' time to be used in performing different obligations, ones that cannot be automated. As such, this technology promises progressively significant utilization in the field of human resources (HR).

In general, ML is categorized as supervised (i.e., consists of output variables that are predicted from input variables) [5] or unsupervised (i.e., deals with clustering of different groups for a particular intervention). ML is used to determine complex models, and extract medical knowledge, exposing novel ideas to practitioners, and specialists [2]. In clinical practice, ML predictive models can highlight enhanced rules in the decision-making regarding individual patient care. These are also capable of autonomous diagnosis of different diseases under clinical regulations [4,6–8]. In [9], the incorporation of these models in drug prescription can save doctors and offer new medical opportunities in pathology identification.

With ML models, it can also be possible to improve quality of medical data, reduce fluctuations in patient rates, and save in medical costs. Therefore, these models are frequently used to investigate diagnostic analysis when compared with other conventional methods [10]. To reduce the death rates caused by chronic diseases (CDs), early detection and effective treatments are the only solutions [11]. Therefore, most medical scientists are attracted to the new technologies of predictive models in disease forecasting [12]. These new advancements in medical care have been expanding the accessibility of electronic data and opening new doors for decision support and productivity improvements [13]. ML methods have been effectively utilized in the computerized interpretation of pneumonic capacity tests for the differential analysis of CDs. It is expected that the models with the highest accuracies could gain large importance in medical diagnosis.

Due to the low-progress nature of CDs, it is important to make an early prediction and provide effective medication. Therefore, it is essential to propose a decision model which can help to diagnose chronic diseases and predict future patient outcomes. While there are many ways to approach this in the field of AI, the present study focuses distinctly on ML predictive models used in the diagnosis of CDs, which highlights the importance of this study. In this study, we conducted a systematic literature review of different state-of-art of predictive models, and our significant contribution in this paper is to develop comparative model analysis to propose model optimization. In comparison to the conventional data analysis techniques, this review article will able to find promising results that enhance the quality of patient data and analysis of specific items that are related to ML algorithms in medical care.

2. Methods

2.1. Search Strategy

The systematic literature search was conducted through the libraries of PubMed (Medline) and Cumulative Index to Nursing and Allied Health Literature (CINAHL). Keywords like ‘chronic diseases’, ‘predictive models’, ‘ML in CD diagnosis’, and ‘model classifiers’ were used during the document search. The search was conducted in January 2020 and resulted in 453 documents. The documents were filtered based on its publication dates ranging from 2015 to 2019 to evaluate the latest literature on ML classifiers in CD prediction.

2.2. Selection Criteria

The title and abstract of the individual articles were retrieved based on the mentioned search terms. Finally, a few of the items were found to be eligible to fulfill the research objectives. This research only describes predictive models used to perform CD diagnosis and does not concentrate on overall trends in AI medicine. Further article revision was conducted to filter the duplicates between the two databases. Moreover, the inclusion and exclusion criteria of our review were based on time, methodological quality and language. Reports and other studies published before 2015 were excluded as outside the limitations on the timeframe of this study. The inclusion criteria used in Pub Med and CINAHL are as follows: free full text, English, original papers and research outcomes. We excluded 276 items among the total search documents because of duplication. The remaining 177 were screened to match the methodologies related to the current research topic.

2.3. Data Extraction

Data evaluation was conducted in two phases. In the first phase, depending on the inclusion criteria, 55 documents were identified for extended revisions. In the second phase, two individual researchers (GB and GGS) were equally distributed for quality check. As discussed, the proposal of a precise model in CD diagnosis was considered as the main focus of this paper. Therefore, articles were extracted based on the authors' information, the study design of sampling pattern and method types, and diagnostic criteria. The analysis of each article was individually revised and recorded.

2.4. Quality Evaluation

Quality assessment check was accomplished by the adoption of the Newcastle–Ottawa Scale (NOS), which is a renowned method in the assessment of study relevance and research interest [14]. The quality of each published article was evaluated as weak (0–4), moderate (5–6), or strong (7–9). Each selected study score was recorded in separate excel sheets to compute whether an individual paper was suitable or not for this review. Ultimately, 22 studies were selected, which are in line with the predictive models in the CD diagnosis (Figure 1). Based on their content, the selected papers were tabled into predictive models used in CD identification (Table 1) and pathologies with model usage, along with their strengths and limitations (Table 2).

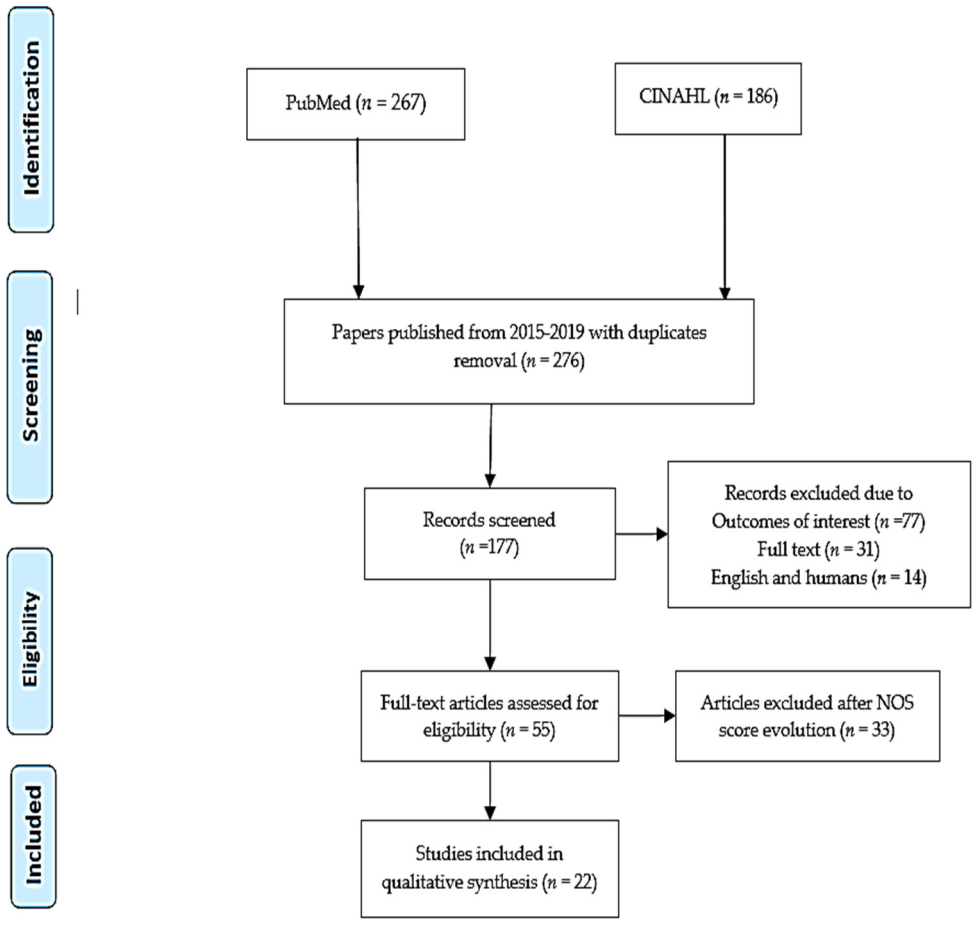

Figure 1. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) diagram [15].

Table 1. Machine learning (ML) algorithms on different pathologies along with input features and outcomes measures.

CD Diagnosis	Study Type	Input Features	Outcomes	Models	Reference
Hepatic fibrosis	Cross-sectional	Age, sex and RTE images	Accuracy, Sensitivity, and Specificity	NB, RF, KNN, SVM, and NN	[16–18]
Chronic hepatitis B stages	Case study	Gene expressions	Precision and AU-ROC	RF, KNN, SVM	[19]
COPD exacerbation events	Retrospective	COPD symptoms	TP, FP, ROC	BN	[20]
Aggravating event identification of COPD	Longitudinal	EDGE digital health system	AU-ROC	LR	[21]
Exacerbations of COPD patients	Case-control	Equi-ripple bandpass (BP)	Sensitivity, specificity, accuracy, PPV, NPV	PCA coupled SVM	[22]
Diabetes classification	Case study	Age and clinical data	Sensitivity, specificity, accuracy, AU-ROC	LR, ANN, NB, KNN, and RF	[23]
Glomerulus filtration rate estimation	Retrospective cohort study (RCT)	Age, sex, and serum creatinine 99mTc-DTPA imaging	Accuracy	ANN, SVM	[24]
Asthma exacerbations events	Case-control	Telemonitoring data	Sensitivity, specificity, accuracy	NB, adaptive Bayesian network, and SVM	[25]
Stage of lung cancer	Prospective cohort study	Cyrano's 320 sensor device, age	Accuracy, sensitivity, and specificity	SVM	[26]
Pulmonary function tests	RCT	Blood analysis, lung images	Accuracy	DT	[27]
Dementia prediction	Case-control	MRI	Accuracy, precision, and specificity	SVM	[28]
Identification of ischemic stroke lesions	Cross-sectional	MRI	Accuracy	NB	[29]
Course of depression	Case study	A shortened version of the IDS (QIDS)	Accuracy	LR	[30]
Late-life dementia assessment	Prospective cohort	MRI/CT, Blood Tests	ROC, AUC and, MCA	SVM	[31]
Degenerative movement disorders	Cross-sectional	Pathological	Not defined	Hierarchical clustering analyses	[32]
Checking CT imaging effectiveness	Case study	CT images, Age, and sex	Accuracy, AU-ROC	NN	[33]
Discriminatory peptide identification of heart failures	Experimental	Age, sex, and renal function	Sensitivity, specificity	SVM	[34]

Table 1. *Cont.*

CD Diagnosis	Study Type	Input Features	Outcomes	Models	Reference
Classification of chronic periodontitis patients	Case-control	Age and PH subjects	Accuracy, Sensitivity, Specificity	SVM	[35]
Classification of fibromyalgia	Case Study	ICD-9 codes	Mean	K-means clustering	[36]
Chronic diseases assessment	Prospective Cohort	Community question answers	Accuracy	NB, SVM, and RNN	[37]

Table 2. Pathology types with used models and their strengths and weakness.

Pathology Type	Name	Models	Accuracy (%)	Strengths	Limitations	Future Developments
Liver	Hepatic fibrosis stage[16], and chronic hepatitis-B [19]	NB, RF, KNN, SVM, and NN	78.1–82.7	Liver related diseases produce large patient information, metabolomics analyses, and EHR. Deep learning algorithms help in the prediction of liver therapeutic discovery.	There is currently no complete AI system that can able to detect a couple of abnormalities overall through the human body [38].	Further studies are needed to develop an advanced deep learning algorithm to remedy greater complicated medical imaging troubles, along with ultrasound or Positron-emission tomography (PET) [18].
Pulmonary	COPD exacerbation, asthma exacerbation[25], lung cancer stages [26]	Bayesian Network, LR, SVM, NB, and PCA	62.3–76.1	Studies proposed a data-driven methodology that can help to produce COPD predictive models and asthma exacerbations. It would be useful to support both patients and physicians [39].	Even it is less cost of devices like spirometers to check lung functionality but it is not likely to replaced by quantified computed tomography.	It is highly recommended in future studies to incorporate ML models in the predictive analysis [40].
Nervous system	Dementia, Ischemic stroke lesions identification [29], late-life dementia [31], degenerative moment disorders [32]	SVM, LR, NB, RF, Hierarchical clustering analyses, and DSI	69–80	ML studies in Nervous systems can help to improve the diagnosis of Nerve system conditions	AI-based behavioral systems are still in early to understand the discrete behavior of patient chronic conditions	Future AI might be able to represent these features into one cognitive reinforcement-mastering model [41].
Diabetes	Type 2 Diabetes Mellitus [23]	LR, ANN, NB, KNN, and RF	73.2–91.6	These techniques in diabetic studies can be helpful in symptoms recognition, and disease forecasting	Technological advancements in AI need to more effective with large data sets in diabetes prediction [42]	ML applications need to produce facts on big data mining of medical data sets [42,43].
Kidney Diseases	Glomerular filtration rate estimation [24]	ANN, SVM, Regression and ensemble learning	73.1–76.0	Risk prediction can highly effective in kidney diseases	The research gap in the artificial kidney implantation needs to be addressed [44].	Many demanding situations need to be a success before it becomes a fact and a part of medical practice in nephrology.

Table 2. *Cont.*

Pathology Type	Name	Models	Accuracy (%)	Strengths	Limitations	Future Developments
Disease-related to muscle pains	Fibromyalgia (FM) [36]	KNN	-	In FM class division, K-means clusters can helpful for categorization of pain, clinical procedure usage, and symptom severity	KNN is a self-learner in trained data classification [45].	Future studies are needed to propose feasible algorithms to forecast FM causes.
Heart diseases	peptides for heart failure [34]	NB, and SVM	84–91	Optimized data-driven ML techniques are helped to predict heart diseases that improve total research and preventive care. Also, it will make sure that many people can happily lead a healthy lifestyle	To predict the risk quality of the heart dataset is needed in clinical practice to support high-quality datasets of heart patients.	Scientists' are needed to propose precise models to predict the risk of heart failures [46]
Infections	Periodontitis [35]	SVM, NN	Not defined	NN and SVM algorithms are useful in the diagnosis and prediction of periodontal diseases	Lack of optimal datasets and model improvements	A computer-aided classification system can be expected to become an efficient and effective procedure for these inflectional diseases [47]

3. Results

3.1. Predictive Models Applied in Diagnosis of CD

From Table 1, it is evident that about 45% of studies used SVM models, 23% of the studies used K-Nearest Neighbor (KNN), and Naïve Bayes (NB) models, 18% of studies applied LR, and 14% of studies applied random forest (RF) models in the CD diagnosis. Regression-based ML models were largely used to predict liver, gas chromatography, and pathological changes. Two studies successfully implemented the random forest (RF) model to do a prediction of the liver fibrosis stages [16,19]. These studies also applied the linear regression (LR) statistical analysis to understand the relationship of image parameter and liver fibrosis stages. The results highlight that RF models are better at identifying the liver fibrosis index (LFI) degree than other statistical approaches [16]. A review of 427 patients on hepatitis-C produced better predictions through the decision trees [17], and multilayer perceptron (MLP) neural networks were best in predicting late-stage liver fibrosis.

In [20], COPD patient's data were analyzed by Bayesian network models. The usage of support vector machines (SVM), LR, Bayesian network, and K-Nearest Neighbor (KNN) models is useful in forecasting the aggravating events of COPD patients [20–22]. Among them, SVM models show better accuracy in predicting exacerbations and COPD detection. In addition, artificial neural networks (ANN) and LR models can effectively be used to understand whether a patient is diabetic or not [23]. In [24], scholars estimated the glomerular filtration rate of the kidney can be done through ensemble models.

3.2. Model Accuracies along with advantages and limitations

Model accuracy is defined as a percentage of true predictions from total predictions. From Table 2, it is evident that diabetic predictions show an accuracy of 73.1–91.6% [23]. Cardiac diseases produce a

prediction accuracy of 84–91% [34], while in the prediction of liver diseases, NB, RF, KNN, SVM, and NN models produce an accuracy in a range of 78.1–82.7%. In particular, RF and NN models are found to have the highest accuracy. In nervous system pathologies, the LR model with feature extraction techniques has identified reasons for depression with accuracy between 72% and 80% [48]. KNN models could be better used to identify disease patterns [49]. In Fibromyalgia, these have the capability to classify pain, clinical usage, and symptom severity [36]. In pulmonary diseases, Bayesian models produced low accuracy range in between 62.3% and 76.1%, because these are not recommended in high dimensional data sets. In contrast, ANN models produced the highest accuracy of 76% in kidney diseases, since they can detect the possible connection between classification variables [50].

Despite the above results, it is still debatable matter on the existence of particular microbial profiles for distinct periodontal conditions. Studies conducted on the ML model development in classifying patient data based on bacterial species [35] have shown that SVM with kernel methods are more helpful. At the same time, dementia is one of the chronic diseases that happen in older people, and, in particular, Alzheimer's is associated with 60–70% of dementia cases. AD prediction through ML models concluded that prediction accuracy depends on the data type and model input [28,31]. These studies with the Disease State Index (DSI) technique produced an accuracy of 79%. All the mentioned studies found that age, cognition, subjective memory complaints, and vascular factors were input features, which can affect the chances of dementia. Therefore, it is understandable that dataset type, input features, and user outcomes can differ by individual study and no model can predict diseases with 100% accuracy.

4. Discussion

The present study analyses the distinct prediction models of machine learning in the diagnosis of chronic diseases. Sometimes, it could be hard to propose the best learning method in disease predictions since it depends on dataset size and user access. Supervised machine learning (SML) approaches are followed in the highest number of studies, with the integration of easy and simple predictive modeling. The implementation of these models in clinical practice certainly can help to provide better health services and enhance specialist decision-making.

It is rudimentary to confirm the different algorithms based on a specific problem, and review studies could help to analyze the performance and determine optimal machine learning models. Before machine learning, recommendations for practice in medicine development depended on individual studies. Therefore, it is affecting the data science because all this medical information is coming from different platforms and people. Due to contemporary trends in computational models, healthcare services are quickly transformed by having the ability to record large amounts of patient data. However, it is highly impossible to analyze huge medical records with human knowledge. On the other hand, with the evolvement of big data in biomedical and medicinal service networks, accurate analysis of medical data becomes possible that could improve patient care [51]; if there is an unavailability of quality medical data, it could result in poor decision-making. Eventually, machine-learning techniques can able to find clear data patterns that can empower health experts in clinical care such as precision medicine. As mentioned in Table 1, different studies have used different predictive models coupled with their results. However, a proper interpretation of medical data will not only help to recommend suitable machine learning models, but also for physicians in the provision of immediate medication.

Aging is one of the significant difficulties facing in the western world, meaning that lowering the weight of continuous infection and enhancing life span is necessary [52,53]. Few researchers have anticipated the safety of each aging compound with a collection of deep neural network classifiers [53]. In addition, predictive models can help in active decision support, in particular if a couple of identical cases are accessible, or where diagnostic symptom knowledge is not precisely available [54]. In this study, we compared and discussed the recent advancements in ML methods used for disease predictions. Studies like pulmonary patient data classification demand the algorithm to anticipate discrete values by distinguishing the patient information either as an individual, or group [27]. The final clinical

diagnosis depended on the integration of a full pulmonary function identification decided by medical expert choice. In the end, clinical diagnosis was classified into ten individual groups, which were validated by experts in the panel. Medical imaging data such as MRI, CT scans, and RTE images follow the classification type SVM models. In contrast, regression problems look at continuous data, and most of the adopted studies followed these model examples of pathologies in [16,30,33,48] gene expression [19], and others.

On the other hand, unsupervised ML deals with a deep learning model containing a medical data set that it can handle without having a clear direction regarding what and how to proceed. The neural network attempts automatic detection of structured data and performs key feature extraction. Depending on the pathology type, it can follow different patterns like clusters that involve a group of particular information [37]. However, some models in machine learning can make immediate decisions on chronic diseases thanks to recent developments in AI. Our findings suggest that stimulating the power of these predictive models in the CDs diagnosis and in structuring medical data will empower medical experts or physicians that will result in a significant tendency decision making at medical centers.

It is also evident that SVM and LR models significantly implemented in the large number of studies to do CD diagnosis. Sixteen studies were adopted these models especially for hepatitis B, COPD, diabetes, and others. An SVM model is popular among others to identify COPD from the beginning, and it could be greatly assisted in the relationship between doctor and patient. Bayesian networks and NB models help to forecast the diagnosis of asthma problems. These models encompass old patient records to look up clinical symptoms and footing on Bayesian networks to present the relationship individual case and diagnose future possible symptoms. The KNN algorithm is associated with five studies for diagnosis, forecasting, and to critically follow the CD's stages with the help of different primary and secondary data.

The main limitation of the present study is most of the adopted literature on disease prediction or classification was adopted with supervised models, and it is important to adopt unsupervised (clustering) and deep learning (neural networks) models as well in future works.

5. Conclusions

The present study evaluated the studies associated with the diagnosis of chronic diseases. However, the implementation of correct methods or selection of the right models is a prerequisite to perform ideal decisions, as modern researchers are claiming that some ML models are compromised by enlarging contained datasets with malicious data that can have severe consequences. On the other hand, diagnosis limitations may lead to life-threatening attacks, and sometimes it might be a driving factor of fatality. In contrast, the wrong diagnosis prompts the skepticism in machine learning use, that can lead policy makers to avoid predictive model usage. Therefore, reviews on predictive models can provide evidence to propose excellent methods for the CDs diagnosis.

In the future, AI techniques like ML, cognitive computing and deep learning may play a critical role in the interpretation of chronic diseases. However, researchers are progressively attracted by predictive model techniques in the advancement of health care. As new advancements in medical care are being established and are expanding the access to electronic data, this opens new doors to decision support and productivity improvements. These models are designed to emphasize the responsibility of patient care quality and cut down medical costs.

Author Contributions: G.B.: Study planning; design and data collection; methods; paper reviews; manuscript preparation, G.G.S. and N.C.: Study selection; research background F.A.: Study approval; final manuscript revision. All authors have read and agreed to the published version of the manuscript.

Funding: Part of this work was supported by a grant of the ITF Trust (No. 1276/2018) for the epidemiological analysis part.

Conflicts of Interest: Authors do not have any potential conflicts of interest.

Acronyms

ML: Machine Learning; AI: Artificial Intelligence; SML: Supervised Machine Learning; DSS: Decision-support systems; CVD: Cardio Vascular Diseases; RTE: Real-Time Electrographs; CHB: Chronic Hepatitis B; NB: Naïve Bayes; RF: Random Forest; ANN: Artificial Neural Networks; RNN: Recurrent neural networks; KNN: K-Nearest Neighborhood; SVM: Support Vector Machine; NN: Neural networks; LR: Logistic Regression; PCA: Principal Component Analysis. DSI: Discrete Smooth Interpolation; FP: False positives; TP: True positives; FN: False negatives; TN: True negatives; AU-ROC: Area under receiver operating characteristics; COPD: chronic obstructive pulmonary disease; PPV: Positive Predictive Value; NPV: Negative Predictive Value. QIDS: Quick Inventory of Depressive Symptomatology; CT: Computerized Tomography; MRI: Magnetic Resonance Imaging; MCA: Multiple Correspondence Analysis; DTPA: Diethylene Thiamine Penta acetate; PET: Positron-emission tomography; EHR: Electronic Health Record.

References

1. Hamet, P.; Tremblay, J. Artificial intelligence in medicine. *Metabolism* **2017**, *69*, S36–S40. [CrossRef]
2. Johnson, K.W.; Soto, J.T.; Glicksberg, B.S.; Shameer, K.; Miotto, R.; Ali, M.; Dudley, J.T. Artificial intelligence in cardiology. *J. Am. Coll. Cardiol.* **2018**, *71*, 2668–2679. [CrossRef] [PubMed]
3. Bini, S. Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? *J. Arthroplast.* **2018**, *33*, 2358–2361. [CrossRef] [PubMed]
4. Buch, V.H.; Ahmed, I.; Maruthappu, M. Artificial intelligence in medicine: Current trends and future possibilities. *Br. J. Gen. Pract.* **2018**, *68*, 143–144. [CrossRef] [PubMed]
5. Kotsiantis, S.B.; Zaharakis, I.; Pintelas, P. Supervised machine learning: A review of classification techniques. *Emerg. Artif. Intell. Appl. Comput. Eng.* **2007**, *160*, 3–24.
6. Deo, R.C. Machine Learning in Medicine. *Circulation* **2015**, *132*, 1920–1930. [CrossRef] [PubMed]
7. Peek, N.; Combi, C.; Marin, R.; Bellazzi, R. Thirty years of artificial intelligence in medicine (AIME) conferences: A review of research themes. *Artif. Intell. Med.* **2015**, *65*, 61–73. [CrossRef]
8. Battineni, G.; Sagaro, G.G.; Nalini, C.; Amenta, F.; Tayebati, S.K. Comparative Machine-Learning Approach: A Follow-Up Study on Type 2 Diabetes Predictions by Cross-Validation Methods. *Machines* **2019**, *7*, 74. [CrossRef]
9. Lo, Y.-C.; Rensi, S.; Torng, W.; Altman, R.B. Machine learning in chemoinformatics and drug discovery. *Drug Discov. Today* **2018**, *23*, 1538–1546. [CrossRef]
10. Napolitano, G.; Marshall, A.; Hamilton, P.; Gavin, A.T. Machine learning classification of surgical pathology reports and chunk recognition for information extraction noise reduction. *Artif. Intell. Med.* **2016**, *70*, 77–83. [CrossRef]
11. Polat, H.; Mehr, H.D.; Cetin, A. Diagnosis of Chronic Kidney Disease Based on Support Vector Machine by Feature Selection Methods. *J. Med. Syst.* **2017**, *41*, 55. [CrossRef] [PubMed]
12. Eslamizadeh, G.; Barati, R. Heart murmur detection based on wavelet transformation and a synergy between artificial neural network and modified neighbor annealing methods. *Artif. Intell. Med.* **2017**, *78*, 23–40. [CrossRef]
13. Martinez, D.; Pitson, G.; Mackinlay, A.; Cavedon, L. Cross-hospital portability of information extraction of cancer staging information. *Artif. Intell. Med.* **2014**, *62*, 11–21. [CrossRef] [PubMed]
14. Wells, G.; Shea, B.; O'Connell, D.; Peterson, J. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses*; Ottawa Hospital Research Institute: Ottawa, ON, Canada, 2000.
15. PRISMA. *PRISMA—Transparent Reporting of Systematic Reviews and Meta-analyses*; Ottawa Hospital Research Institute: Ottawa, ON, Canada, 2015.
16. Chen, Y.; Luo, Y.; Huang, W.; Hu, D.; Zheng, R.-Q.; Cong, S.-Z.; Meng, F.; Yang, H.; Lin, H.; Sun, Y.; et al. Machine-learning-based classification of real-time tissue elastography for hepatic fibrosis in patients with chronic hepatitis B. *Comput. Boil. Med.* **2017**, *89*, 18–23. [CrossRef]

17. Shousha, H.I.; Awad, A.H.; Omran, D.; Elnegouly, M.M.; Mabrouk, M. Data Mining and Machine Learning Algorithms Using IL28B Genotype and Biochemical Markers Best Predicted Advanced Liver Fibrosis in Chronic Hepatitis C. *Jpn. J. Infect. Dis.* **2018**, *71*, 51–57. [CrossRef] [PubMed]
18. Zhou, L.-Q.; Wang, J.-Y.; Yu, S.-Y.; Wu, G.-G.; Wei, Q.; Deng, Y.-B.; Wu, X.-L.; Cui, X.-W.; Dietrich, C.F. Artificial intelligence in medical imaging of the liver. *World J. Gastroenterol.* **2019**, *25*, 672–682. [CrossRef] [PubMed]
19. Zhou, W.; Ma, Y.; Zhang, J.; Hu, J.; Zhang, M.; Wang, Y.; Li, Y.; Wu, L.; Pan, Y.; Zhang, Y.; et al. Predictive model for inflammation grades of chronic hepatitis B: Large-scale analysis of clinical parameters and gene expressions. *Liver Int.* **2017**, *37*, 1632–1641. [CrossRef]
20. Mcheick, H.; Saleh, L.; Ajami, H.; Mili, H. Context Relevant Prediction Model for COPD Domain Using Bayesian Belief Network. *Sensors* **2017**, *17*, 1486. [CrossRef]
21. Shah, S.A.; Velardo, C.; Farmer, A.J.; Tarassenko, L.; Kim, B.; Barber, V. Exacerbations in Chronic Obstructive Pulmonary Disease: Identification and Prediction Using a Digital Health System. *J. Med. Int. Res.* **2017**, *19*, e69. [CrossRef]
22. Granero, M.A.F.; Sanchez-Morillo, D.; Leon-Jimenez, A. Computerised Analysis of Telemonitored Respiratory Sounds for Predicting Acute Exacerbations of COPD. *Sensors* **2015**, *15*, 26978–26996. [CrossRef]
23. Olivera, A.R.; Roesler, V.; Iochpe, C.; Schmidt, M.I.; Vigo, Á.; Barreto, S.M.; Duncan, B.B. Comparação de algoritmos de aprendizagem de máquina para construir um modelo preditivo para detecção de diabetes não diagnosticada—ELSA-Brasil: Estudo de acurácia. *Sao Paulo Med. J.* **2017**, *135*, 234–246. [CrossRef] [PubMed]
24. Liu, X.; Li, N.; Lv, L.; Fu, Y.; Cheng, C.; Wang, C.; Ye, Y.; Li, S.; Lou, T. Improving precision of glomerular filtration rate estimating model by ensemble learning. *J. Transl. Med.* **2017**, *15*, 231. [CrossRef] [PubMed]
25. Finkelstein, J.; Jeong, I.C. Machine learning approaches to personalize early prediction of asthma exacerbations. *Ann. N. Y. Acad. Sci.* **2016**, *1387*, 153–165. [CrossRef] [PubMed]
26. Tirzīte, M.; Bukovskis, M.; Strazda, G.; Jurka, N.; Taivans, I. Detection of lung cancer in exhaled breath with an electronic nose using support vector machine analysis. *J. Breath Res.* **2017**, *11*, 036009. [CrossRef]
27. Topalovic, M.; Laval, S.; Aerts, J.M.; Troosters, T.; Decramer, M.; Janssens, W. Belgian Pulmonary Function Study Investigators. Automated Interpretation of Pulmonary Function Tests in Adults with Respiratory Complaints. *Respiration* **2017**, *93*, 170–178. [CrossRef]
28. Battineni, G.; Chintalapudi, N.; Amenta, F. Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). *Inform. Med. Unlocked* **2019**, *16*, 100200. [CrossRef]
29. Griffis, J.C.; Allendorfer, J.B.; Szaflarski, J.P. Voxel-based Gaussian naïve Bayes classification of ischemic stroke lesions in individual T1-weighted MRI scans. *J. Neurosci. Methods* **2015**, *257*, 97–108. [CrossRef]
30. Dinga, R.; Marquand, A.F.; Veltman, D.J.; Beekman, A.T.F.; Schoevers, R.A.; Van Hemert, A.M.; Penninx, B.W.J.H.; Schmaal, L. Predicting the naturalistic course of depression from a wide range of clinical, psychological, and biological data: A machine learning approach. *Transl. Psychiatry* **2018**, *8*, 241. [CrossRef]
31. Pekkala, T.; Hall, A.; Lötjönen, J.; Mattila, J.; Soininen, H.; Ngandu, T.; Laatikainen, T.; Kivipelto, M.; Solomon, A. Development of a Late-Life Dementia Prediction Index with Supervised Machine Learning in the Population-Based CAIDE Study. *J. Alzheimers Dis.* **2016**, *55*, 1055–1067. [CrossRef]
32. Kuo, S.H.; Lin, C.Y.; Wang, J.; Sims, P.A.; Pan, M.K.; Liou, J.Y.; Faust, P.L. Climbing fiber-Purkinje cell synaptic pathology in tremor and cerebellar degenerative diseases. *Acta Neuropathol.* **2017**, *133*, 121–138. [CrossRef]
33. Oakden-Rayner, L.; Carneiro, G.; Bessen, T.; Nascimento, J.C.; Bradley, A.; Palmer, L.J. Precision Radiology: Predicting longevity using feature engineering and deep learning methods in a radiomics framework. *Sci. Rep.* **2017**, *7*, 1648. [CrossRef] [PubMed]
34. Farmakis, D.; Koeck, T.; Mullen, W.; Parissis, J.; Gogas, B.D.; Nikolaou, M.; Filippatos, G. Urine proteome analysis in heart failure with reduced ejection fraction complicated by chronic kidney disease: Feasibility, and clinical and pathogenetic correlates. *Eur. J. Heart Fail.* **2016**, *18*, 822–829. [CrossRef]
35. Feres, M.; Louzoun, Y.; Haber, S.; Faveri, M.; Figueiredo, L.C.; Levin, L. Support vector machine-based differentiation between aggressive and chronic periodontitis using microbial profiles. *Int. Dent. J.* **2017**, *68*, 39–46. [CrossRef] [PubMed]
36. Davis, F.; Gostine, M.; Roberts, B.; Risko, R.; Cappelleri, J.C.; Sadosky, A. Characterizing classes of fibromyalgia within the continuum of central sensitization syndrome. *J. Pain Res.* **2018**, *11*, 2551–2560. [CrossRef]
37. Singh, S.M.; Hanchate, D.B. Improving disease prediction by machine learning. *Int. J. Res. Eng. Technol.* **2018**, *5*, 1542–1548.

38. Van Der Heijden, M.; Velikova, M.; Lucas, P.J. Learning Bayesian networks for clinical time series analysis. *J. Biomed. Inform.* **2014**, *48*, 94–105. [CrossRef]
39. Swaminathan, S.; Qirko, K.; Smith, T.; Corcoran, E.; Wysham, N.G.; Bazaz, G.; Kappel, G.; Gerber, A.N. A machine learning approach to triaging patients with chronic obstructive pulmonary disease. *PLoS ONE* **2017**, *12*, e0188532. [CrossRef] [PubMed]
40. Gawlitza, J.; Sturm, T.; Spohrer, K.; Henzler, T.; Akin, I.; Schoenberg, S.; Borggrefe, M.; Haubenreisser, H.; Trinkmann, F. Predicting Pulmonary Function Testing from Quantified Computed Tomography Using Machine Learning Algorithms in Patients with COPD. *Diagnostics* **2019**, *9*, 33. [CrossRef]
41. Bzdok, D.; Meyer-Lindenberg, A. Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Boil. Psychiatry Cogn. Neurosci. Neuroimaging* **2018**, *3*, 223–230. [CrossRef]
42. Kavakiotis, I.; Tsave, O.; Salifoglou, A.; Maglaveras, N.; Vlahavas, I.; Chouvarda, I. Machine Learning and Data Mining Methods in Diabetes Research. *Comput. Struct. Biotechnol. J.* **2017**, *15*, 104–116. [CrossRef]
43. Gopi, B.; Getu, G.S.; Chintalapudi, N.; Francesco, A.; Seyed, K.T. Comparative Machine-Learning Approach: A Follow-Up Study on Type 2 Diabetes Predictions by Cross-Validation Methods. *Machines* **2019**, *7*, 74.
44. Hueso, M.; Vellido, A.; Montero, N.; Barbieri, C.; Ramos, R.; Angoso, M.; Cruzado, J.M.; Jönsson, A. Artificial Intelligence for the Artificial Kidney: Pointers to the Future of a Personalized Hemodialysis Therapy. *Kidney Dis.* **2018**, *4*, 1–9. [CrossRef] [PubMed]
45. Zhang, H.; Berg, A.; Maire, M.; Malik, J. SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition—Volume 2 (CVPR 06), New York, NY, USA, 17–22 June 2006; Institute of Electrical and Electronics Engineers (IEEE): Piscataway Township, NJ, USA.
46. Tripoliti, E.E.; Papadopoulos, T.G.; Karanasiou, G.S.; Naka, K.K.; Fotiadis, D.I. Heart Failure: Diagnosis, Severity Estimation and Prediction of Adverse Events Through Machine Learning Techniques. *Comput. Struct. Biotechnol. J.* **2016**, *15*, 26–47. [CrossRef] [PubMed]
47. Lee, J.-H.; Kim, D.-H.; Jeong, S.-N.; Choi, S.-H. Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm. *J. Periodontal Implant. Sci.* **2018**, *48*, 114–123. [CrossRef]
48. Liu, Y.; Yieh, L.; Yang, T.; Drinkenburg, W.H.; Peeters, P.; Steckler, T.; Narayan, V.A.; Wittenberg, G.M.; Ye, J. Metabolomic biosignature differentiates melancholic depressive patients from healthy controls. *BMC Genom.* **2016**, *17*, 669. [CrossRef] [PubMed]
49. Weinberger, K.Q.; Saul, L.K. Distance metric learning for large margin nearest neighbor classification. *J. Mach. Learn. Res.* **2009**, *10*, 207–244.
50. Buscema, P.M.; Massini, G.; Breda, M.; Lodwick, W.A.; Newman, F.; Asadi-Zeydabadi, M. Artificial Neural Networks. In *Studies in Systems, Decision and Control*; Springer: Berlin/Heidelberg, Germany, 2018.
51. Chen, M.; Hao, Y.; Hwang, K.; Wang, L.; Wang, L. Disease Prediction by Machine Learning Over Big Data From Healthcare Communities. *IEEE Access* **2017**, *5*, 8869–8879. [CrossRef]
52. Das, N.; Topalovic, M.; Janssens, W. Artificial intelligence in diagnosis of obstructive lung disease. *Curr. Opin. Pulm. Med.* **2018**, *24*, 117–123. [CrossRef]
53. Aliper, A.M.; Jellen, L.; Cortese, F.; Artemov, A.; Karpinsky-Semper, D.; Moskalev, A.; Swick, A.G.; Zhavoronkov, A. Towards natural mimetics of metformin and rapamycin. *Aging* **2017**, *9*, 2245–2268. [CrossRef]
54. Kuru, K.; Niranjan, M.; Tunca, Y.; Osvank, E.; Azim, T. Biomedical visual data analysis to build an intelligent diagnostic decision support system in medical genetics. *Artif. Intell. Med.* **2014**, *62*, 105–118. [CrossRef]

Informatics in Medicine Unlocked 16 (2019) 100200

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM)

Gopi Battineni[a,*], Nalini Chintalapudi[b], Francesco Amenta[a,c]

[a] Center for Telemedicine and Tele Pharmacy, School of Pharmaceutical Sciences and Health Products, University of Camerino, Via Madonna Delle carceri 9, 62032, Camerino (MC), Italy
[b] Computer Science Department, MRIT, JNT University, India
[c] Studies and Research Department, International Medical Radio Center Foundation (C.I.R.M.), via dell'Architettura 61 Roma (RM), Italy

ARTICLE INFO

Keywords:
Machine learning
OASIS
Support vector machines
Kernel
Gamma
Regularization (C)

ABSTRACT

Machine Learning (ML) is considered as one of the contemporary approaches in predicting, identifying, and making decisions without having human involvement. ML is quickly evolving in the medical industry ranging from diagnosis to visualization of diseases and the study of disease transmission. These algorithms were developed to identify the problems in medical image processing. Numerous studies previously attempted to apply these algorithms on MRI (Magnetic Resonance Image) data to predict AD (Alzheimer's disease) in advance. The present study aims to explore the usage of support vector machine (SVM) in the prediction of dementia and validate its performance through statistical analysis. Data is obtained from the Open Access Series of Imaging Studies (OASIS-2) longitudinal collection of 150 subjects of 373 MRI data. Results provide evidence that better performance values for dementia prediction are achieved by low gamma (1.0E-4) and high regularized (C = 100) values. The proposed approach is shown to achieve accuracy and precision of 68.75% and 64.18%.

1. Introduction

Machine learning (ML) was considered as an integral part of Artificial Intelligence (AI), also a data analysis technique that computerizes the explanatory model structure. In most scenarios, based on the learning method, two types of ML algorithms (supervised & unsupervised) were used [1]. At present, these algorithms are engaging in all the major industries like healthcare, banking, transport, social media, etc. [1,2]. Above all, the medical industry is advancing quickly with high volumes of information and increasing difficulties in inventory and patient outcomes. Economically developed nations such as USA, Japan, European countries are even facing the problems with the enormous collection of medical data [3]. However, by using conventional techniques, it is not possible to analyze this significant volume of information because of time consumption and efforts. Therefore, ML techniques are coming up with various algorithms and programs to avoid these issues. Besides that, the selection of proper algorithm is not an easy task since it depends on multiple factors such as data volume, information type, and outcomes related to industry requirements [1].

Nowadays, ML algorithms are progressively utilized in neuroimaging studies like a prediction of Alzheimer's disease (AD) from auxiliary MRI. Also, many studies attempted different ML strategies in predicting AD and their causes [3,4]. In the study of AD prediction and retrieval, a multistage classifier utilizing ML, including Naive Bayes classifier, support vector machine (SVM), and K-nearest neighbor (KNN) was used to group Alzheimer's illness in the more acceptable and effective way [5]. Similarly, a study from Ref. [6], concluding that the utilization of locally linear embedding (LLE) kind of unsupervised learning was utilized to categorize AD based on fundamental MRI data. Besides, some preliminary studies with ML techniques concluded that these methods are valid and accomplish with high precision (up to 98%) in diagnosing clinical events with analysis of patient medical records [7].

Despite of it, AD is one of common type in dementia and associated mostly with older people [8]. In this paper, we explain how to predict dementia and calculate performance by using support vectors. Typically, SVM's are considered as supervised machine learning, which solves the data issues related to classification and regression analysis [9]. An SVMs give a compelling and adaptable structure for MRI, and that the proposed classifier perception technique has potential as a system for the assessment of characterization solutions [9,10]. This is also used to categorize dementia subjects and is similar to the research that use a uniform algorithm to differentiate 3 Primary progressive aphasia (PPA) subtypes in predicting PPA [11]. Distinguishing early morphological changes in the mind and making initial finding is

* Corresponding author. Ph.D student in E-health and Telemedicine, Italy. Tel.: +39-3331728206.
E-mail address: gopi.battineni@unicam.it (G. Battineni).

https://doi.org/10.1016/j.imu.2019.100200
Received 2 May 2019; Received in revised form 25 June 2019; Accepted 25 June 2019
Available online 27 June 2019

significant for dementia. High-resolution MRI data can be utilized to support finding and forecast of the disease [11]. To do this, we propose to find an optimal solution by experimenting with radial basis function (RBF) kernel in the SVM. The proposed method of calculation is inspired by a new approach of using an ensemble SVM for dementia classification [12], using MRI data and mini-mental state examination parameters (MMSE).

In contrast, we consider the attributes like MR delay; CDR, ASF, AGE, and GENDER included with MMSE that corresponds to subject ID. We strongly believe that it is the novel way of examining the importance of each parameter while forecasting dementia in older patients. Despite it, our work aims to predict dementia in elder individuals by SVM algorithms to accomplish promising outcomes. This paper is organized as follows; section 2 describes the SVM background with its key parameters; Section 3 will explain the data collection and methodology; section 4 will provide experimental results; Section 5 on proposed discussion and little conclusion in section 6.

2. Support vector machines (SVM)

2.1. Background

Support Vector Machines are a well-known ML technique for classification and other learning activities. SVM is a discriminative classifier and formally characterized by an optimal hyperplane. It produces an outcome of the optimal hyperplane, which classifies new examples and datasets that support hyperplane are called support vectors [13]. In two-dimensional (2D) region, this hyperplane is a line isolating into two segments wherein each segment lay in either side. For instance, multiple line data classification had done with two distinct datasets (i.e., squares and dots) and ready to propose an affirmative interpretation (Fig. 1). However, the selection of optimal hyperplane is not an easy job as it should not be noise sensitive, and generalization of data sets should be accurate [14]. Pertinently, SVM trying to find out optimized hyperplane that provides considerable minimum distance to the trained data set [13,14].

In mathematical notation, for 2D space, a line can distinguish the linearly separable data. The equation of the line is $y = ax + b$. By rename x with x_1 and y with x_2, the equation will change to $ax_1 - x_2 + b = 0$. If we specify $X = (x_1, x_2)$ and $w = (a, -1)$, we get $w \cdot x + b = 0$, which is called the equation of the hyperplane.

2.1.1. Derivation of SVM optimization problem

To estimate w & b of the optimal hyperplane, it is mandatory to address a performance issue with the need of the geometric edge for every pattern must be more prominent to M [16].

$$\text{Max}_{\ w,\ b}\ M;\ \text{Subject to}\ \gamma_i \geq M,\ i = 1 \ldots m \tag{1}$$

If $M = \frac{F}{\|w\|}$ the above equation can be rewritten as:

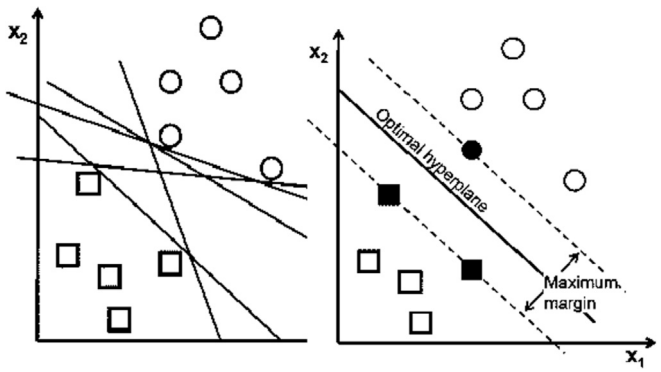

Fig. 1. Data classification using multiple lines [On left] and data classification: optimal hyperplane [On Right] [15].

$$\text{Max}_{\ w,\ b}\ M;\ \text{Subject to fi} \geq F,\ i = 1 \ldots m \tag{2}$$

The case that rescales w and b are yet boosting M, and the enhancement result will not change. Let us rescale w & b and make F = 1; the above equation shift to

$$Max_{w,b}\frac{1}{\|w\|};\ subject\ \text{to fi} \geq 1\ ,\ \text{i= 1...m} \tag{3}$$

This maximization issue is proportionate to the accompanying minimization issue written as

$$\text{Min}_{\ w,\ b}||w||;\ \text{subject to fi} \geq 1,\ i = 1 \ldots m \tag{4}$$

This minimization issue is proportionate to the accompanying minimization issue written as

$$\min_{w,b}\frac{1}{2}\|w\|^2;\ \text{subject to}\ y_1(wx + b) - 1 \geq 0,\ \text{i= 1...m} \tag{5}$$

The above statement refers to the SVM optimization problem.

2.1.2. SVM classifier

When we have the hyperplane, eventually we would be able to utilize the hyperplane to make predictions. The hypothesis function of H is

$$H(x_i) = \begin{cases} +1\ if\ w.x \geq 0 \\ -1\ if\ w.x < 0 \end{cases}$$

2.2. Tuning parameters

To comprehend the SVM working, it is critical to understand about some prerequisites like kernel, regularization, and gamma.

2.2.1. Kernel

In machine learning, the kernel is a technique that is used to solve the non-linear problem with the use of linear classifier and involved in exchanging linearly non-separable data into linearly separable data [17]. The idea behind this concept is linearly non-separated data in N-dimensional space might be linearly separate in high M-dimensional space. Mathematically, kernel indicated as K (a, b) = < F (a), F (b) > , Where K: kernel function and a, b are n-dimensional inputs. 'F' is mapping from N-dimensional to M-dimensional space (i.e., M > N). The mapping in the kernel is defined as K (a, b) = Ø (a). Ø (b).

Kernel Functions: There are several kernels functions some of them listed below here [18].

- Polynomial Type: is well known for nonlinear modeling and is represented as

$$K\ (a,\ b)\ =\ (a,\ b)^{\ d} \tag{6}$$

- Gaussian Radial Basis Type: Radial basis functions mostly with Gaussian form and represented by

$$k(a,b)\ =\ \exp(-\frac{\|a-\ b\|^2}{2\sigma^2}) \tag{7}$$

- Exponential Radial basis: function produces a bitwise linear solution that will be useful when discontinuities are satisfactory

$$k(a,\ b)\ =\ \exp(-\frac{\|a-\ b\|}{2\sigma^2}) \tag{8}$$

In addition to them, there are many more functions such as multilayer perceptron, Fourier, additive, and tensor products type [18].

2.2.2. Regularization

The regularization parameter (C) explains the SVM optimization

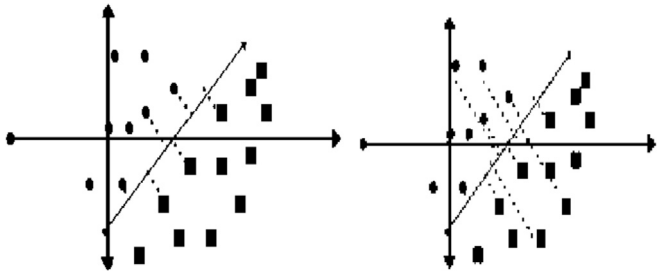

Fig. 2. High Gamma Close points (left) and Low Gamma Far away points found (Right) [20].

and percentage of escaping the misclassified trained data [19]. For high C values, training data will categorize accurately by hyperplane; similarly, for low C, optimizer looks for higher margin separating hyperplane while it will misclassify the more data points.

2.2.3. Gamma

It describes the impact of specific training data [13,17,19]. The high gamma values (Fig. 2 left) results in consideration of datasets that are near to separation line. Similarly, for low gamma values (Fig. 2.Right) datasets that are away from the separation line, will also be taken into consideration while in the calculation of separation line (Chapter 2: SVM (Support Vector Machine) — Theory – Machine Learning 101 – Medium).

3. Data collection and methodology

3.1. Dataset

We consider a longitudinal collection of OASIS - MRI data set [21], comprising of demented and non-demented subjects with right-hand (R) type aging from 60 to 96. A sample size of 150 subjects, including men and woman, have attended scanning sessions more than two visits; sessions were separated by at least one year with 373 MR Sessions. The sample training data (Table 1) included with demographic values of Subject ID, MRI ID, Group, Visit, MR delay, Sex, Age, Social Economic Status (SES), Education level (EDUC), MMSE [22], Clinical Dementia Ratio (CDR) [23], estimated Total Intracranial Volume (e-TIV), normalized Whole Brain Volume (n-WBV) and Atlas Scaling Factor (ASF). Also, Fig. 3. Explaining the present MRI sessions categorization based on the current CDR (0–2) score and total sessions of non-demented (190), demented (146) and converted (37) were evaluated. In particular, some subjects treated as demented at initial visit later transformed into the non-demented managed by converted type. If CDR value is equal to zero, the subjects were considered as mostly non-demented, simultaneously if CDR ≥ 1 the subjects will face the tendency to have dementia.

3.2. Methodology

The methodology layout that used and analyzed in the current study in explaining in Fig. 4.

❖ Data collection

The trained data set was collected from the Open Access Series of Imaging Studies (OASIS) included with longitudinal MRI data of 150 subjects.

❖ Data Preprocessing

Real world data is available more likely incomplete with missing entries. Therefore, data preprocessing is one of the data mining techniques to address this issue. Missing entries were filled-up by averaging of particular attribute values.

❖ Attribute Selection

Select a specific characteristics to predict the outcome to do mapping with input correspondence values. We choose the group column as output variable that corresponds to the dementia status based on other input variables.

❖ Input variable matching

Performance of any ML model largely depends on the number of input attributes taken into consideration. To maintain better performance, selection of the corresponding attributes, instead of selecting multiple ones is very important. Attributes like Subject ID, CDR, MMSE, Age, MR Delay, and n WBV chosen as input to SVM that were directly targeted to the dementia group attribute.

❖ Classifier

We consider three groups of dementia as demented, non-demented, and converted.

Table 1
Example of actual portion dataset of Longitudinal OASIS-2 MRI data.

SUBJECT ID	MRI ID	GROUP	VISIT	MR Delay	M/F	Hand	Age	EDUC	SES	MMSE	CDR	E TIV	n-WBV	ASF
OAS2_0100	OAS2_0100_MR1	Non Demented	1	0	F	R	77	11	4	29	0	1583	0.777	1.108
OAS2_0100	OAS2_0100_MR2	Non Demented	2	1218	F	R	80	11	4	30	0	1586	0.757	1.107
OAS2_0100	OAS2_0100_MR3	Non Demented	3	1752	F	R	82	11	4	30	0	1590	0.760	1.104
OAS2_0101	OAS2_0101_MR1	Non Demented	1	0	F	R	71	18	2	30	0	1371	0.769	1.280
OAS2_0101	OAS2_0101_MR2	Non Demented	2	952	F	R	74	18	2	30	0	1400	0.752	1.254
OAS2_0101	OAS2_0101_MR3	Non Demented	3	1631	F	R	76	18	2	30	0	1379	0.757	1.273
OAS2_0102	OAS2_0102_MR1	Demented	1	0	M	R	82	15	3	29	0.5	1499	0.689	1.171
OAS2_0102	OAS2_0102_MR2	Demented	2	610	M	R	84	15	3	29	0.5	1497	0.686	1.172
OAS2_0102	OAS2_0102_MR3	Demented	3	1387	M	R	86	15	3	30	0.5	1498	0.681	1.171
OAS2_0103	OAS2_0103_MR1	Converted	1	0	F	R	69	16	1	30	0	1404	0.750	1.250
OAS2_0103	OAS2_0103_MR2	Converted	2	1554	F	R	74	16	1	30	0.5	1423	0.722	1.233
OAS2_0103	OAS2_0103_MR3	Converted	3	2002	F	R	75	16	1	30	0.5	1419	0.731	1.236
OAS2_0104	OAS2_0104_MR1	Demented	1	0	M	R	70	16	1	25	0.5	1568	0.696	1.119

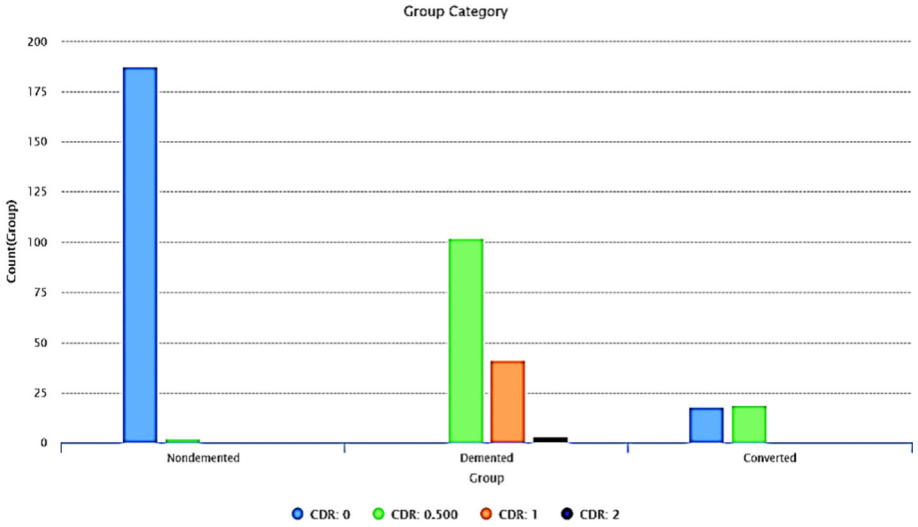

Fig. 3. Categorization of dementia sessions based clinical dementia ratio (CDR).

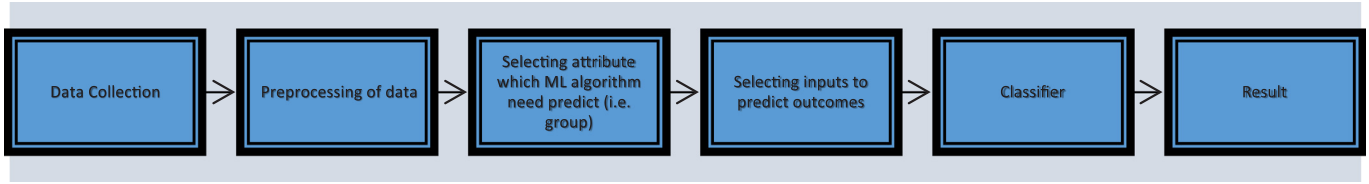

Fig. 4. Methodology layout.

❖ Results

Finally, the classification performance has achieved and analyzed. Performance value calculated as the percentage of correctly predicted outcomes divided by the total number of samples

$$i.\ e,\ performance = \frac{\text{True predicted an outcomes}}{\text{Total number of samples}} \times 100$$

4. Results

Once the mapping has done by input attributes with targeted output group column, the machine will run the SVM algorithm automatically.

4.1. Kernel

The Kernel outcome model with 150 support vectors (Table 2) has generated, and three different categories of training data set are observed. As mentioned, kernel mapping with three input values formulated, as K (ND, D, C) = ø(79).ø(50).ø(21), where K is kernel function with three input class vectors such as non-demented (ND), Demented (D) and converted (C) and corresponding mapping values of 79, 50 and 21. Besides, bias value is equal to −0.3 (offset defines compensate the feature vectors that are not centered around the zero).

Table 2
Kernel outcome statistic values.

Total number of Support Vectors: 150
Bias (offset): 0.3 and Number of classes: three Number of support vectors for class Non-demented - 79 Number of support vectors for class Demented- 50 Number of support vectors for class Converted - 21

4.2. Gamma VS. C

As discussed, Gamma and C values are obligatory to confirm optimal hyperplane. In further, Radial Basis Function (RBF) kernel is one of the novel kernel approaches that related to gamma. Hence, SVM anticipated with following performance conditions

$$if \begin{cases} RBF = 1.0E-4\ ;\ C = 100 & p = 69.2\%; \\ RBF = 1.0E-3\ ;\ C = 100 & p = 69.2\% \\ RBF = 1.0E-1\ ;\ C = 10 & p = 57.1\% \end{cases}$$

Here, two conditions were supporting identical performance gain. However, as per condition of SVM it prefers to choose optimal hyperplane region with low RBF (1.0E-4), and High C (100) which represented by Yellow colored circle explained in Fig. 5.

4.3. Performance, precision, and recall

- ❖ Assessment of performance done by the percentage of true predicted subjects from the total subjects. From Table 3, sum of true predicted subjects were 105, therefore performance was calculated as 70% ($\frac{105}{150}$*100). This value is matching with the optimal system performance 69.2% by utilization of RBF and C values that proves the SVM hypothesis.
- ❖ Precision is define as percentage of positive predictive values for each subject category. For demented subjects precision validates as of 64.18% ($\frac{43}{43+14+10}$*100) and for demented 75%. On the counter note, no valid predicted values for converted category subjects. SVM algorithm predicted two subjects as a converted category, but in a real scenario, it belongs to non-demented ones.
- ❖ In the context of ML, recall is referred as sensitivity or true positive rate. Thus, recall for non-demented subjects validated with 81.13% ($\frac{43}{43+8+2}$*100), and demented 65.85%.

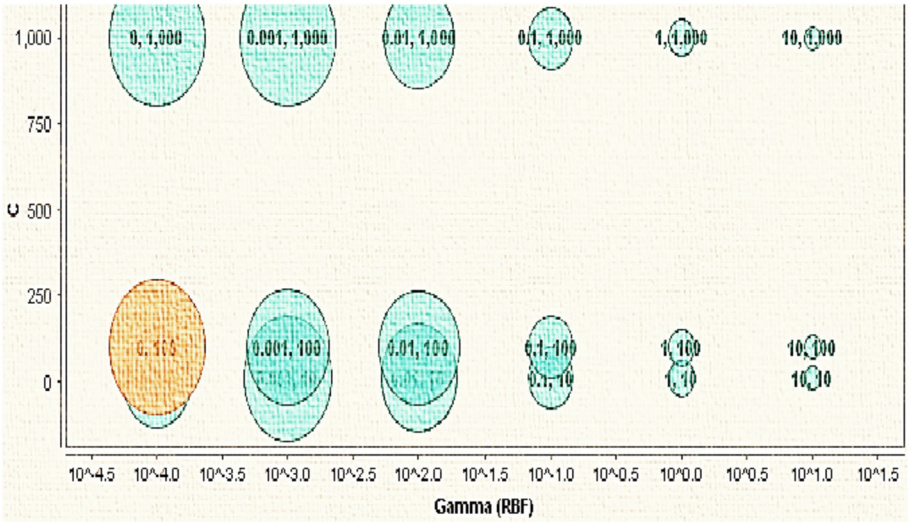

Fig. 5. Spatial distribution of Gamma (RBF) Vs. C values.

Table 3
Confusion Matrix of given subjects TND*: True Non-Demented; TD*: True Demented; TC*: True Converted; PND*: Predict Non-Demented; PD*: Predict Demented, and PC*: Predict Converted.

	TND	TD	TC	precision
PND	43	14	10	64.18%
PD	8	27	1	75.00%
PC	2	0	0	0.00%
Recall	81.13%	65.85%	0.00%	0.00%

5. Discussion

In present study, we considered longitudinal MRI subjects from OASIS datasets, and input information to machine chose as key attributes like MMSE, CDR, MR delay and n WBV. The forecasting of dementia depends on the scores of mentioned attributes. As best of our insight, this is the essential investigation for foreseeing dementia dependent on these scores by utilizing SVM calculations. Additionally, we locate an ideal hyperplane by using RBF and C esteems that is also used in the study of weather forecast datasets [24]. It helped us to make a correlation between hyperplane parameters to investigate better support vectors. We classified MRI sessions into three groups based on the CDR scale (0–2). Additionally, we conduct statistical analysis by bar charts to differentiate subject category. In next sections, we are going to introduce the outcomes of these group-level comparisons, after that, we discuss in more detail about how SVM produces optimized performance values to forecast dementia using kernel functions and study limitations when compared to other methods.

5.1. Dementia prediction by a selection of key attributes

As discussed, current MRI sessions division was done based on the current CDR value. Beyond that, our subject group classifications are in line with the study designed for investigating diagnostic agreements [25]. However, it is not feasible to predict dementia disease with single attribute or parameter. Thus, we examine with other key parameters such as MMSE, AGE, n WBV and MR delay that matched with targeted group column. At the same time we tried to exclude other demographic values like Gender, SES, EDU, and ASF since these parameters not good enough in dementia prediction, also by considering many attributes performance may get low [26]. In addition, outcomes mentioned that 100 subjects (Fig. 6) are predicted non-demented (actually these distributed as 63ND, 24D and 13C types), and 47 subjects predicted as

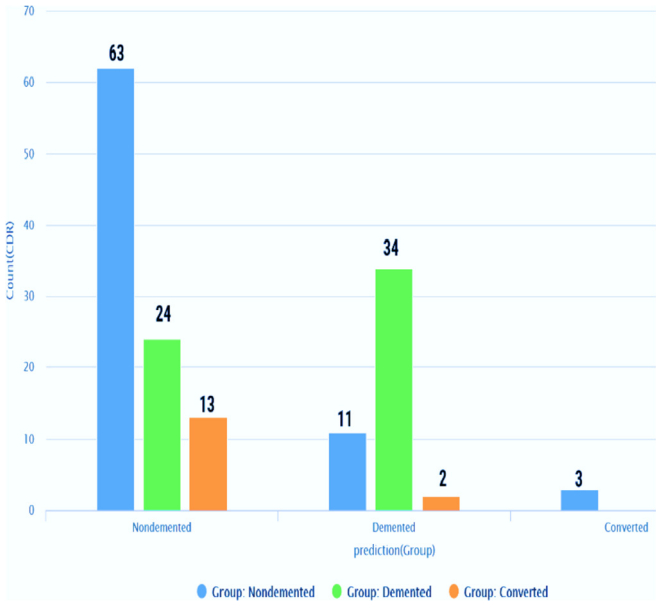

Fig. 6. Subject Classification between Predictions subject groups Vs Actual Subject Groups.

demented (but these distributed as 11ND, 34D and 2C types). Finally, 3 non-demented subjects forecasted as converted type.

The prediction was validated and done on the confidence values of the actual category of each subject (refer appendix). For example, in real time scenario Subject, ID-04 was non-demented based current CDR score (=0) but predicted as demented. This might be caused by the high age [27] or more significant delay in the MR value [28]. Therefore, this will change from subject to subject depending on the present reports.

5.2. Selection of optimal hyperplane

The performance for the given dataset by SVM algorithm producing about 70% and recall or sensitivity providing in the range of 65–82% that depended on the subject category. Until now, only single research tried to develop a new method for an ensemble of SVM for classification of dementia using systematic MRI and MMSE values [12]. The researchers performed ensemble SVM using RBF kernel or linear to achieve distinct class accuracies. In their results, accuracy was

increased from 55% to 59.1%. Our SVM approach by considering total brain value with MMSE and CDR, producing the accuracy nearly 70%. Additionally, we compared the statistical calculation of performance outcomes with optimal hyperplane coordinates to verify whether machine-generated results were performing similar SVM optimal performance (Fig. 5). In the end, outcomes generated by the ML system and Hyperplane are matched to prove the theory of support vector algorithms.

5.3. Limitations

The relatively lowest number of subjects may hamper the speculation of outcomes to the overall population of dementia patients. Despite that, our study closely related to Ref. [12], but we achieve better performance values by introducing an optimal hyperplane study. Classification and normalization of subject groups are not accurate in most cases, and it might tend to underestimation of dementia in older patients that result in getting low accuracies through SVM categorization. Nevertheless, approaching optimal hyperplanes, we tried to increase the performance by a selection of low RBF and high C values. Eventually, the order between sets of different subjects was an optimal hyperplane, which does not reflect the issue regarding accurate differential determination between a few neurological diseases. This issue should be addressed in future researches validating the use of SVM approaches consistently in real life.

6. Conclusion

Dementia is one of the significant health issues that has challenged health experts worldwide. In addition, it mostly happened in older people (age > 60). Unfortunately, there are no proper medicines for completely cure this disease, and sometimes it will directly affect person memory skills and reduce the human ability to perform daily activities. Many healthcare professionals and computer scientists were performing research activities on this problem from last two decades. Still, there is an extreme need for identification of relevant characteristics that can forecast the detection of dementia. We approached support vectors for classification and prediction purposes of dementia and achieved optimized results with efficient performance values.

Conflicts of interest

Authors do not have any conflict of interest.

Informed consent

Trained data gathered from OASIS longitudinal studies and access available from concerned authorities in participating in the mentioned study. Authors do not have permission for direct communication with a human participant in the study.

Author note and contributions

We are certifying that the manuscript is orginal and not published with any journal. All authors are strictly read and validated the final copy. GB: Made primary contributions of data collection,methodology developement,and conduct empirical analysis NC: Done contributions with literature review and discussion sections FA: Final check, approval and manuscript validation.

Acknowledgments

We are thankful to the Principal Investigators: D. Marcus, R, Buckner, J. Csernansky, and J. Morris to provide access on OASIS longitudinal studies. The Italian Ministry of Economic Development (MISE) no. F/080034/01/X35 funded this study. Gopi Battineni Ph.D. bursaries were supported in part by the University of Camerino.

Appendix. Dementia predicted outcome dataset after SVM implementation

N	Age	CDR	MMSE	MR Delay	N WBV	Group	Conf (ND)	Conf (D)	Conf (Con)	Prediction
1	87	0.0	27	0	0.7	Nondemented	1	0	0	Nondemented
2	80	0.5	22	1895	0.7	Demented	1	0	0	Nondemented
3	88	0.0	28	0	0.7	Nondemented	1	0	0	Nondemented
4	90	0.0	27	538	0.7	Nondemented	0	1	0	Demented
5	85	0.0	30	1603	0.7	Nondemented	0	0	1	Converted
6	71	0.5	28	0	0.7	Demented	1	0	0	Nondemented
7	75	1.0	27	1281	0.7	Demented	1	0	0	Nondemented
8	68	0.5	27	0	0.8	Demented	0	1	0	Demented
9	66	0.5	30	0	0.8	Demented	1	0	0	Nondemented
10	68	0.5	29	854	0.8	Demented	1	0	0	Nondemented
11	78	0.0	29	0	0.7	Nondemented	1	0	0	Nondemented
12	80	0.0	29	730	0.7	Nondemented	1	0	0	Nondemented
13	85	0.0	29	1456	0.7	Nondemented	1	0	0	Nondemented
14	81	0.5	27	617	0.8	Nondemented	0	1	0	Demented
15	86	0.0	27	2400	0.8	Nondemented	1	0	0	Nondemented
16	87	0.0	30	0	0.7	Converted	1	0	0	Nondemented
17	88	0.0	29	489	0.7	Converted	1	0	0	Nondemented
18	92	0.5	27	1933	0.7	Converted	1	0	0	Nondemented
19	64	0.0	29	828	0.8	Nondemented	1	0	0	Nondemented
20	82	0.5	27	0	0.7	Demented	0	1	0	Demented
21	71	0.0	30	609	0.8	Nondemented	0	1	0	Demented
22	73	0.0	30	1234	0.8	Nondemented	1	0	0	Nondemented
23	77	0.0	29	0	0.7	Nondemented	1	0	0	Nondemented
24	60	0.0	30	0	0.8	Nondemented	1	0	0	Nondemented
25	86	0.0	30	0	0.7	Converted	1	0	0	Nondemented
26	90	0.5	21	0	0.7	Demented	0	1	0	Demented
27	88	0.0	30	0	0.7	Nondemented	1	0	0	Nondemented
28	89	0.0	27	405	0.7	Nondemented	1	0	0	Nondemented
29	75	0.0	29	2369	0.8	Nondemented	1	0	0	Nondemented
30	85	0.5	29	1123	0.7	Demented	1	0	0	Nondemented
31	89	0.5	26	2508	0.7	Demented	1	0	0	Nondemented
32	83	0.5	25	486	0.7	Demented	1	0	0	Nondemented

33	86	0.5	27	567	0.7	Demented	0	1	0	Demented
34	73	0.0	28	756	0.8	Converted	1	0	0	Nondemented
35	75	0.0	30	0	0.7	Nondemented	1	0	0	Nondemented
36	66	1.0	21	248	0.7	Demented	0	1	0	Demented
37	68	1.0	19	647	0.7	Demented	0	1	0	Demented
38	69	1.0	4	1233	0.7	Demented	0	1	0	Demented
39	78	0.0	30	1510	0.7	Nondemented	1	0	0	Nondemented
40	84	0.0	28	842	0.7	Nondemented	0	1	0	Demented
41	85	0.0	29	0	0.7	Converted	1	0	0	Nondemented
42	87	0.5	24	846	0.7	Converted	0	1	0	Demented
43	67	0.0	27	726	0.8	Nondemented	1	0	0	Nondemented
44	71	0.0	28	0	0.8	Nondemented	1	0	0	Nondemented
45	85	0.0	30	1340	0.7	Nondemented	1	0	0	Nondemented
46	79	0.5	26	212	0.7	Demented	0	1	0	Demented
47	70	0.0	30	873	0.7	Nondemented	1	0	0	Nondemented
48	72	0.0	30	1651	0.7	Nondemented	1	0	0	Nondemented
49	79	0.0	29	0	0.7	Nondemented	1	0	0	Nondemented
50	83	0.0	29	1351	0.7	Nondemented	1	0	0	Nondemented
51	81	0.5	27	490	0.7	Demented	1	0	0	Nondemented
52	81	0.5	26	830	0.7	Demented	0	1	0	Demented
53	82	0.5	18	1282	0.7	Demented	0	1	0	Demented
54	62	0.5	30	497	0.7	Demented	0	1	0	Demented
55	68	0.0	29	451	0.7	Nondemented	1	0	0	Nondemented
56	71	0.0	29	1438	0.7	Nondemented	0	1	0	Demented
57	73	0.0	28	2163	0.7	Nondemented	1	0	0	Nondemented
58	90	0.0	29	743	0.7	Nondemented	1	0	0	Nondemented
59	82	0.0	30	432	0.7	Nondemented	1	0	0	Nondemented
60	82	0.0	29	672	0.7	Nondemented	1	0	0	Nondemented
61	84	0.0	29	1415	0.7	Nondemented	1	0	0	Nondemented
62	86	0.0	30	2386	0.7	Nondemented	1	0	0	Nondemented
63	84	1.0	28	365	0.7	Demented	1	0	0	Nondemented
64	70	0.0	29	0	0.8	Nondemented	1	0	0	Nondemented
65	72	0.0	28	580	0.8	Nondemented	0	1	0	Demented
66	75	0.5	22	567	0.7	Demented	0	1	0	Demented
67	66	0.0	30	0	0.7	Nondemented	1	0	0	Nondemented
68	73	0.0	29	1393	0.7	Nondemented	1	0	0	Nondemented
69	89	0.0	28	0	0.7	Nondemented	1	0	0	Nondemented
70	71	1.0	16	584	0.7	Demented	0	1	0	Demented
71	66	0.5	25	0	0.7	Demented	0	1	0	Demented
72	68	0.5	30	580	0.7	Demented	0	1	0	Demented
73	69	0.5	28	1209	0.7	Demented	1	0	0	Nondemented
74	82	0.5	26	0	0.7	Demented	0	1	0	Demented
75	78	1.0	21	0	0.7	Demented	0	1	0	Demented
76	72	1.0	27	563	0.7	Demented	0	1	0	Demented
77	75	0.0	29	680	0.8	Nondemented	1	0	0	Nondemented
78	76	0.0	30	1345	0.8	Nondemented	1	0	0	Nondemented
79	61	0.0	30	0	0.8	Nondemented	1	0	0	Nondemented
80	67	0.5	28	661	0.8	Demented	1	0	0	Nondemented
81	80	0.5	27	0	0.8	Demented	0	1	0	Demented
82	77	0.0	29	0	0.8	Nondemented	1	0	0	Nondemented
83	76	0.0	30	1631	0.8	Nondemented	1	0	0	Nondemented
84	82	0.5	29	0	0.7	Demented	1	0	0	Nondemented
85	86	0.5	30	1387	0.7	Demented	1	0	0	Nondemented
86	75	0.5	30	2002	0.7	Converted	0	1	0	Demented
87	87	0.0	30	675	0.7	Nondemented	1	0	0	Nondemented
88	70	1.0	22	0	0.7	Demented	0	1	0	Demented
89	65	0.5	17	881	0.7	Demented	0	1	0	Demented
90	78	0.5	20	558	0.7	Demented	0	1	0	Demented
91	75	0.5	28	504	0.7	Demented	0	1	0	Demented
92	76	0.5	27	0	0.7	Demented	0	1	0	Demented
93	74	0.0	30	576	0.8	Nondemented	0	1	0	Demented
94	78	0.0	29	1927	0.7	Nondemented	1	0	0	Nondemented
95	81	0.0	28	0	0.8	Nondemented	1	0	0	Nondemented
96	74	0.0	30	647	0.7	Nondemented	1	0	0	Nondemented
97	86	0.0	30	0	0.7	Nondemented	1	0	0	Nondemented
98	88	0.0	30	597	0.7	Nondemented	0	1	0	Demented
99	71	0.5	27	472	0.7	Demented	1	0	0	Nondemented
100	79	0.0	29	0	0.7	Converted	1	0	0	Nondemented
101	81	0.5	29	1042	0.7	Converted	1	0	0	Nondemented
102	84	0.5	29	2153	0.7	Converted	1	0	0	Nondemented
103	86	0.5	30	2639	0.7	Converted	1	0	0	Nondemented
104	76	0.0	28	0	0.8	Nondemented	1	0	0	Nondemented
105	78	0.0	30	0	0.7	Nondemented	1	0	0	Nondemented
106	82	0.0	29	1591	0.6	Nondemented	0	0	1	Converted
107	65	0.5	30	0	0.8	Converted	1	0	0	Nondemented
108	74	0.0	30	0	0.7	Nondemented	1	0	0	Nondemented
109	78	0.0	27	1146	0.7	Nondemented	1	0	0	Nondemented
110	74	0.5	28	0	0.7	Demented	1	0	0	Nondemented
111	75	0.5	30	636	0.7	Demented	1	0	0	Nondemented
112	73	0.0	29	0	0.8	Nondemented	1	0	0	Nondemented

113	67	0.5	29	0	0.8	Demented	1	0	0	Nondemented
114	76	0.5	26	0	0.7	Demented	0	1	0	Demented
115	65	0.0	30	0	0.8	Nondemented	1	0	0	Nondemented
116	91	0.0	30	561	0.7	Nondemented	0	1	0	Demented
117	93	0.0	29	1553	0.7	Nondemented	0	0	1	Converted
118	68	0.0	30	0	0.8	Converted	1	0	0	Nondemented
119	82	0.0	30	1806	0.7	Nondemented	1	0	0	Nondemented
120	81	0.0	29	0	0.7	Nondemented	1	0	0	Nondemented
121	73	0.5	30	0	0.7	Demented	1	0	0	Nondemented
122	66	0.0	29	0	0.8	Nondemented	1	0	0	Nondemented
123	68	0.0	29	790	0.8	Nondemented	1	0	0	Nondemented
124	77	0.0	28	791	0.7	Nondemented	1	0	0	Nondemented
125	75	1.0	18	764	0.7	Demented	0	1	0	Demented
126	73	0.5	29	0	0.8	Demented	1	0	0	Nondemented
127	76	0.5	28	759	0.8	Demented	1	0	0	Nondemented
128	77	0.0	29	0	0.7	Nondemented	1	0	0	Nondemented
129	82	0.5	23	0	0.7	Demented	0	1	0	Demented
130	84	0.5	22	621	0.7	Demented	0	1	0	Demented
131	77	1.0	23	0	0.8	Demented	0	1	0	Demented
132	79	2.0	25	580	0.8	Demented	0	1	0	Demented
133	73	0.0	30	691	0.7	Nondemented	1	0	0	Nondemented
134	77	0.0	30	493	0.8	Nondemented	1	0	0	Nondemented
135	75	0.5	30	0	0.7	Demented	1	0	0	Nondemented
136	70	0.5	26	0	0.7	Demented	0	1	0	Demented
137	73	0.5	28	1343	0.7	Demented	1	0	0	Nondemented
138	87	0.0	30	774	0.7	Converted	1	0	0	Nondemented
139	68	0.0	26	0	0.8	Nondemented	0	1	0	Demented
140	70	0.0	28	665	0.8	Nondemented	1	0	0	Nondemented
141	89	0.0	29	0	0.8	Nondemented	1	0	0	Nondemented
142	90	0.0	28	600	0.7	Nondemented	0	1	0	Demented
143	79	0.5	26	0	0.7	Demented	0	1	0	Demented
144	74	0.5	26	0	0.7	Demented	0	1	0	Demented
145	73	0.5	23	0	0.7	Demented	0	1	0	Demented
146	66	0.0	30	182	0.7	Nondemented	1	0	0	Nondemented
147	86	0.5	26	2297	0.7	Demented	1	0	0	Nondemented
148	61	0.0	30	0	0.8	Nondemented	1	0	0	Nondemented
149	63	0.0	30	763	0.8	Nondemented	1	0	0	Nondemented
150	62	0.0	26	1180	0.7	Nondemented	1	0	0	Nondemented

References

[1] Huddleston SH, Brown GG. Machine learning. Informs analytics body of knowledge 2018. https://doi.org/10.1002/9781119505914.ch7.

[2] Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nat Rev Genet 2015. https://doi.org/10.1038/nrg3920.

[3] Ichikawa D, Saito T, Ujita W, Oyama H. How can machine-learning methods assist in virtual screening for hyperuricemia? A healthcare machine-learning approach. J Biomed Inform 2016;64:20–4.

[4] Kaur P, Sharma M, Mittal M. Big data and machine learning based secure healthcare framework. Procedia Comput Sci 2018;132:1049–59.

[5] Kruthika KR, Rajeswari, Maheshappa HD. Multistage classifier-based approach for Alzheimer's disease prediction and retrieval. Inf Med Unlocked 2019;14:34–42. November 2018.

[6] Liu X, Tosun D, Weiner MW, Schuff N. Locally linear embedding (LLE) for MRI based Alzheimer's disease classification. Neuroimage 2013;83:148–57.

[7] XIE R, Khalil I, Badsha S, Atiquzzaman M. Collaborative extreme learning machine with a confidence interval for P2P learning in healthcare. Comput Network 2019;149:127–43.

[8] McKhann GM, et al. The diagnosis of dementia due to Alzheimer's disease: re-commendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's Dement; 2011. https://doi.org/10.1016/j.jalz.2011.03.005.

[9] Campbell C, Ying Y. Learning with support vector machines. Synth Lect Artif Intell Mach Learn 2011. https://doi.org/10.2200/s00324ed1v01y201102aim010.

[10] Levman J, Leung T, Causer P, Plewes D, Martel AL. Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines. IEEE Trans Med Imaging 2008. https://doi.org/10.1109/TMI.2008.916959.

[11] Danek A, Landwehrmeyer B, Ludolph A, Anderl-Straub S, Otto M. Predicting primary progressive aphasias with support vector machine approaches in structural MRI data. NeuroImage Clin 2017;14:334–43.

[12] Sørensen L, Nielsen M. Ensemble support vector machine classification of dementia using structural MRI and mini-mental state examination. J Neurosci Methods 2018;302:66–74.

[13] Wang PW, Lin CJ. Support vector machines. Data classification: algorithms and applications 2014. https://doi.org/10.1201/b17320.

[14] Gholami R, Fakhari N. Learn more about support vector machine support vector Machine : principles , Pa- rameters , and applications quantitative structure-activity rela- tionship (QSAR): modeling approaches to biological applications technical aspects of brain rhythms and sp. 2017.

[15] Support Vector Machine. Introduction to machine learning algorithms [Online]. Available:. https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47 [Accessed: 11-Apr-2019].

[16] Understanding the mathematics behind support vector machines [Online]. Available:. https://shuzhanfan.github.io/2018/05/understanding-mathematics-behind-support-vector-machines/ [Accessed: 29-May-2019].

[17] Smola AJ, Schölkopf B. A tutorial on support vector regression. Stat Comput 2004. https://doi.org/10.1023/B:STCO.0000035301.49549.88.

[18] Andrew AM. An Introduction to Support Vector Machines and other Kernel-Based Learning Methods by Nello Christianini and John Shawe-Taylor. Cambridge: Cambridge University Press; 2000. https://doi.org/10.1017/s0263574700232827. 2000, xiii + 189 pp., ISBN 0-521-78019-5 (Hbk, £27.50). Robotica.

[19] Chang C, Lin C, Tieleman T. LIBSVM : a library for support vector machines'. ACM Trans Intell Syst Technol 2008. https://doi.org/10.1145/1961189.1961199.

[20] Chapter 2 : SVM (Support Vector Machine). Theory – machine learning 101 – Medium [Online]. Available:. https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72 [Accessed: 11-Apr-2019].

[21] Marcus DS, Fotenos AF, Csernansky JG, Morris JC, Buckner RL. Open access Series of imaging studies: longitudinal MRI data in nondemented and demented older adults. J Cogn Neurosci Dec. 2010;22(12):2677–84.

[22] Ridha B, Rossor M. The mini mental state examination. Pract. Neurol.; 2005. https://doi.org/10.1111/j.1474-7766.2005.00333.x.

[23] Morris JC. The clinical dementia rating (CDR): current version and scoring rules. Neurology 2012. https://doi.org/10.1212/wnl.43.11.2412-a.

[24] Smolik M, Skala V, Majdisova Z. Vector field radial basis function approximation. Adv Eng Software 2018. https://doi.org/10.1016/j.advengsoft.2018.06.013.

[25] Abramovitch A, Anholt GE, Cooperman A, van Balkom AJLM, Giltay EJ, Penninx BW, van Oppen P. Body mass index in obsessive-compulsive disorder. J. Affect. Disord. 2019;245:145–51.https://doi.org/10.1016/j.jad.2018.10.116. et al. Prevalence of psychiatric disorders in patients with mechanicalvalve prostheses with and without rheumatic fever. J. Bras. Psiquiatr. (2011).

[26] Er F, Iscen P, Sahin S, Çinar N, Karsidag S, Goularas D. Distinguishing age-related cognitive decline from dementias: a study based on machine learning algorithms. J Clin Neurosci 2017. https://doi.org/10.1016/j.jocn.2017.03.021.

[27] Li J, Ogrodnik M, Devine S, Auerbach S, Wolf PA, Au R. Practical risk score for 5-, 10-, and 20-year prediction of dementia in elderly persons: framingham Heart Study. Alzheimer's Dement; 2018. https://doi.org/10.1016/j.jalz.2017.04.013.

[28] Mok VCT, et al. Delayed-onset dementia after stroke or transient ischemic attack. Alzheimer's Dement; 2016. https://doi.org/10.1016/j.jalz.2016.05.007.

ISSN: 2469-5866

Battineni et al. Int J Brain Disord Treat 2020, 6:033
DOI: 10.23937/2469-5866/1410033
Volume 6 | Issue 1
Open Access

International Journal of
Brain Disorders and Treatment

ORIGINAL RESEARCH

Late-Life Alzheimer's Disease (AD) Detection Using Pruned Decision Trees

Gopi Battineni*, Nalini Chintalapudi and Francesco Amenta

E-Health and Telemedicine Center, University of Camerino, Italy

***Corresponding author:** *Gopi Battineni, E-health and Telemedicine Center, University of Camerino, Camerino, 62032, Italy, Tel: +39-3331728206*

Abstract

Machine Learning (ML) is a contemporary technique of artificial intelligence. These methods are exponentially rising in the medical field, especially in diagnosis and disease predictions. The present study was aimed to develop a decision tree model to predict late-life Alzheimer's disease (AD). A dataset of 150 subjects along with 373 MRI sessions demographic values were considered in this paper. Pruned decision trees (J48) were employed to do predictive analysis on AD subjects. Model validation was conducted with cross fold (k = 10) methods. Performance measures were evaluated by accuracy, precision, and receiver operating characteristic (ROC) curve. Results were provided an accuracy of 88.7%, precision of 86.7%, and ROC of 91.8% was recorded.

Keywords

Machine learning, AD, Cross-validation, Decision tree, ROC

Introduction

There are some established plans and proposals for a medical practice on some external examinations and hard-coded into their software. However, these programs are restrained the data precision because they are generated from different people and conditions. Dementia is one of the global medical issues that was high in demand. Most of the studies are related to dementia causes explaining the risk reduction, early medication, and immediate disease finding in older adults. Therefore, it is mandatory to conduct some advanced studies dealing with these diseases.

In general, subjects with Mild Cognitive Impairment (MCI) are relevant groups for the cure as they are at the prodromal stages and a higher risk of Alzheimer's disease (AD). AD and different kinds of dementia were becoming a global challenge and tending to the death in one of three elder peoples in the USA. While the reasons for these diseases have not yet completely understood, they can effectively affect discourse, memory, and other essential psychological abilities.

Machine learning (ML) is a category of an algorithm that allows software applications to become more accurate in outcome prediction without being explicitly programmed [1]. The basic premise of these methods is to build algorithms that can receive input data and use statistical analysis to predict an output. Nowadays, it is hard to exclude these techniques because most of them used in real-time purposes, and many researchers are thinking that it is an ideal approach to gain grounds toward human-level artificial intelligence (AI) [2,3]. Furthermore, ML methods are similar to data mining, and prediction algorithms as of both require data exploration to search for examples and to change program activities in the same manner [2,4]. Recently, these techniques are gradually increasing in the medical filed for prediction or visualization of patient data [5], development of medical diagnostics case studies [6,7]. Present concentrate on late-life AD detection with the help of MRI demographic data and AD prediction were evaluated with feature characteristics. Pruned decision trees (J48) model was employed to conduct this analysis, and model performance was assessed by accuracy, precision, and receiver operating characteristic curve (ROC).

Methods

Patients

Citation: Battineni G, Chintalapudi N, Amenta F (2020) Late-Life Alzheimer's Disease (AD) Detection Using Pruned Decision Trees. Int J Brain Disord Treat 6:033. doi.org/10.23937/2469-5866/1410033

Accepted: February 06, 2020: **Published:** February 08, 2020

A dataset of consists of 150 patients (i.e., subjects) of demographic MRI data with age ranging from 60 to 96 were considered. All Subjects are giving in informed consent and extracted from the open Access Series of Imaging Studies (OASIS) [8]. Each subject was exposed at least two scanning sessions and a total of 373 MR session information was available. All subjects associated with right hand irrespective gender. Present AD status was decided by the clinical dementia ratio (CDR) and each session was categorized into 3 groups of 146 AD (demented), 190 AD_{non} (Non-demented), and 37 AD_{con} (converted).

Decision trees

Decision trees are the conventional model of machine learning techniques and produce results with higher accuracy when compared to others. An algorithmic methodology developed these that data splitting was done by distinct conditions [9]. Many studies were considered decision trees as a great approach to conducting a predictive analysis. In AD prediction, we begin from the tree root feature and compare this feature with other tree node features. Based on the correlation, we pursue the branch relating to that value and jump to the next node [10]. It is important to keep different AD groups and other tree internal nodes until we achieve a leaf node with a predicated class.

Results

Based on the AD group, all the features were exposed to the J48 decision trees model. Cross (k-fold) validation (CV) techniques were employed to validate the model. The CV is a resampling technique with a unique parameter 'k,' which used in model evaluation on a limited data sample. Based 'k' value data can split into test and train groups. The cross-validation was conducted with k = 5 to avoid fitting issues, which means of five data folds (or subgroups) for testing and k-5 folds for training purposes had used. For generating pruning decision tree, we considered limited features of CDR, MMSE, n-WBV, gender, and MR delay since these are highly correlated with the group category.

Model performance was evaluated by accuracy, precision, and area under receiver operating characteristic curve (AU-ROC). Data preprocessing was conducted by selection of highly correlated features coupled with AD group. Model training was held between the target AD group and rest of the features, model drive the operation of dementia forecasting along performance measures and confusion matrix (Figure 1).

From Figure 1, it is evident 331 were correctly classified among 373 MRI sessions with an accuracy of 88.7%. Weighted average of true positive prediction (i.e., precision) of 86.7% was recorded. Precision (or sensitivity) was calculated by the ratio of true positives and a total number of positive predictions. For example, the precision of true AD subjects is evaluated as 91.3% (Equation 1).

True AD predictive AD

$$\frac{TrueADsubjects}{True(AD+ADnon+ADcon)subjects}=\frac{188}{188+18}*100=91.3\% \quad (1)$$

Correctly Classified Instances	331	88.7399%
Incorrectly Classified Instances	42	11.2601%
Kappa statistic	0.7992	
Mean absolute error	0.1085	
Root mean squared error	0.2609	
Relative absolute error	28.1314%	
Root relative squared error	59.4574%	
Total Number of Instances	373	

--------- Detailed Accuracy By Class ----------

	Precision	ROC Area	Class
	0.913	0.937	Nondemented
	0.924	0.962	Demented
	0.409	0.650	Concerted
Weighted Avg.	0.867	0.918	

---------- Confusion Matrix ----------

a	b	c	< -- classified as
188	1	1|	a = Nondemented
0	134	12|	b = Demented
18	10	9|	c = Converted

Figure 1: Model outcome.

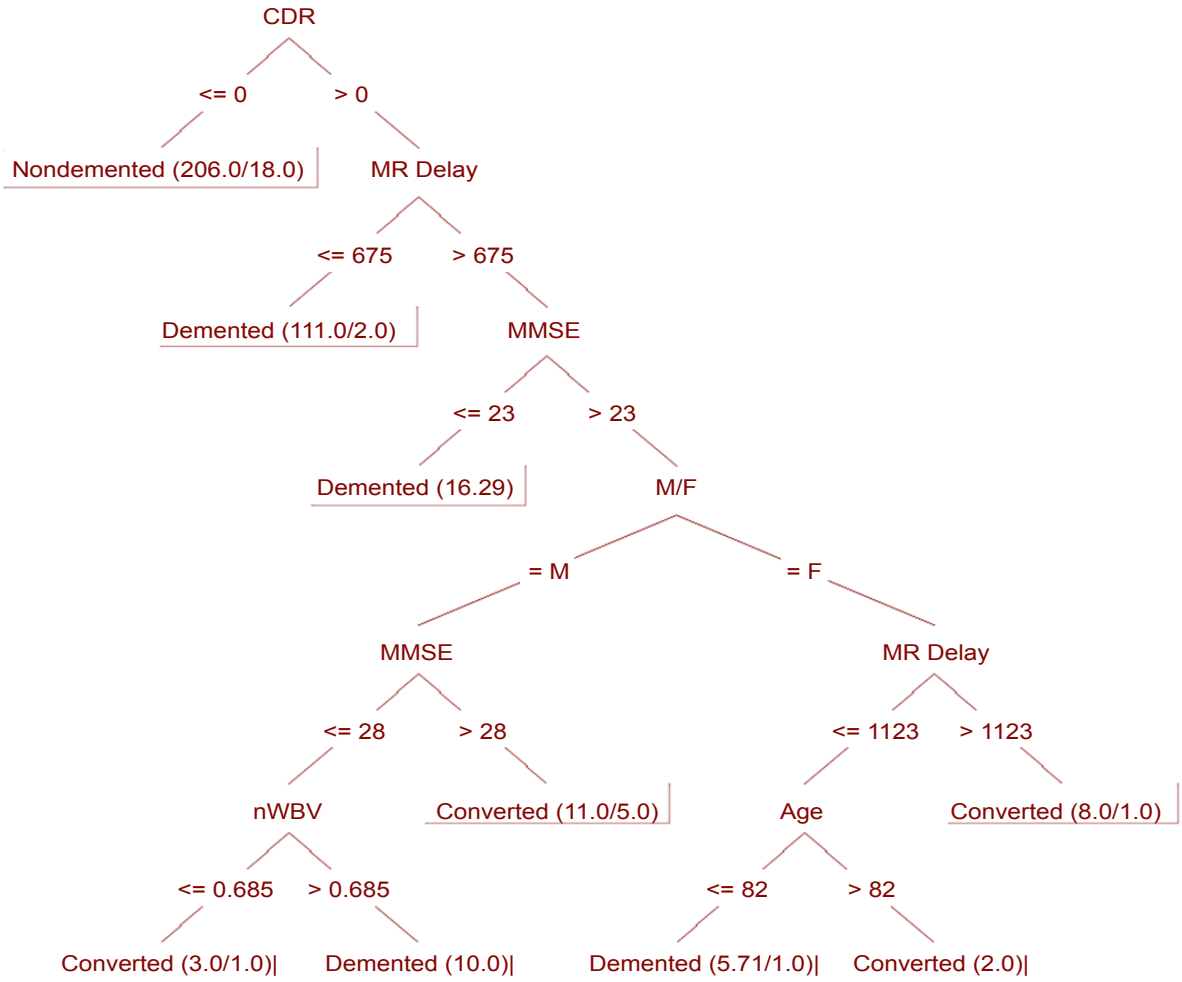

Figure 2: Pruned decision tree outcome (CDR: Clinical Dementia Ratio; MMSE: Mini Mental State Examination; n-WBV: Total Brain Volume; MRI: Magnetic Resonance Imaging).

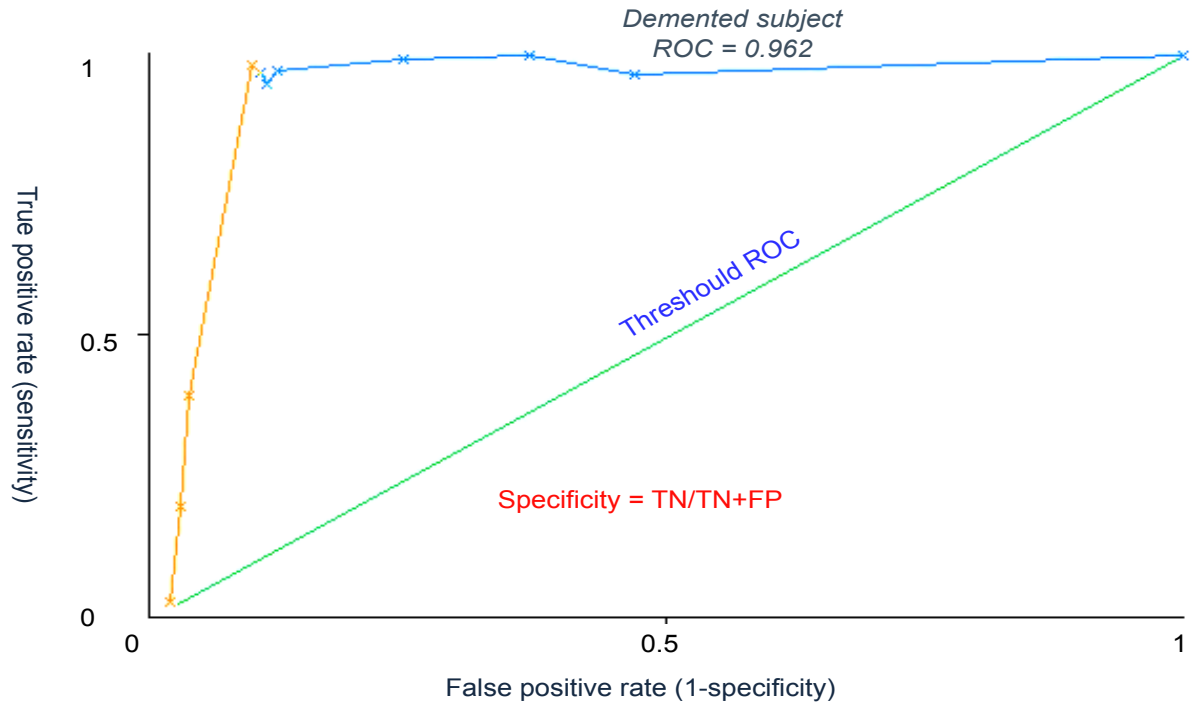

Figure 3: ROC curve of AD subjects.

The J48 pruned decision tree with a central node of CDR can be observed (Figure 2). If the branch CDR ≤ 0, MRI session classification was done as AD_{non} with an accuracy of 92%. The second branch CDR > 0, splitting

into two branches of MR delay as the central node. It generated AD subject accuracy of 98.2%, along with another branch with an MMSE central node. This tree follows the bottom node with a group category. As mentioned, there are some specific cases AD_{con} (i.e., characterized as non-demented at first visit and subsequently described into demented at a later visit and vice versa), which are having a more significant effect on other dementia factors. Generated decision tree predictions have correctly mapped and analyzed with confidence values of dementia status. Eventually, the highest confidence value of dementia will predict the future dementia status of the particular patient, and the mentioned model explains and predict the patient's condition by utilizing specific benefits to help patients by assisting them in advance.

Discussion

In AD diagnosis of most MCI studies, MRI demographic information along with other features highly important in AD forecasting [11]. In this study, we have developed a machine-learning model with a feature reduction (pruning) technique to enhance classification accuracy. Distinct medical diagnostics have developed with the connection of ML implementation. But, few studies were only associated with AD classification. AD is one of the complex data analysis because it requires test information, physical test, cognitive testing, research facility studies, and MR images [11-13]. As of this, we consider specific features such as CDR, MR delay, MMSE, and n-WBV.

At first, the AD group was mapped with the rest of the features, which were highest correlated with present AD status. The CDR value evaluated late-life AD prediction. Despite age, if CDR ≤ 0, then subjects were classified as $AD_{non,}$ and CDR > 0 highest percent of subjects were classified as AD, and rest were as AD_{con}. The outcome tree was generated with different sub-branches and left a decision at the end, considered as a leaf of the corresponding branch. In the end, outcomes suggesting that pruned decision tree models are one of the best approaches with an accuracy of 88.7%.

ROC curve value was evaluated as fundamental analysis in medical diagnosis [14], and it's a plot of true positive rate on y-axis and false positive rate on x-axis (Figure 3). According to [15], in diagnosis classification an excellent model possess ROC near to one that means it has effective measure of separability. If it near to zero said to have worst measure of separability. In this experiment, we got ROC of AD classification is 0.962, which means that comprehensive classification of AD patients was done.

Conclusions

Highest percent of mortality rates were happened due to the lack of early disease diagnosis, and AD is one among them. Especially, old patients were facing the dementia problem; such patients can overcome this issue by some extent through early doctor approach. At the same time, reduction of MR delay could also a comprehensive precaution to overcome probability of AD happening. Therefore, there is more chance to save AD patients in future before they turn into helpless situations.

Conflicts of Interest

No author possesses any conflict of information.

References

1. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI (2015) Machine learning applications in cancer prognosis and prediction. Computational and Structural Biotechnology Journal 13: 8-17.
2. Domingos P (2012) A few useful things to know about machine learning. Communications of the ACM 55: 10.
3. Learning M, Zheng A (2015) Evaluating Machine Learning Models: A Beginner's Guide to Key Concepts and Pitfalls. O'Reilly.
4. Bhatia P (2019) Introduction to Data Mining. Data Mining and Data Warehousing.
5. Darcy AM, Louie AK, Roberts LW (2016) Machine learning and the profession of medicine. JAMA 315: 551-552.
6. Battineni G, Chintalapudi N, Amenta F (2019) Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). Informatics in Medicine Unlocked 16: 100200.
7. Giger ML (2018) Machine Learning in Medical Imaging. Journal of the American College of Radiology 15: 512-520.
8. Smith SS (2009) Predicting Alzheimer's dementia mortality using Medicare Outcome Assessment and Information Set (OASIS).
9. Podgorelec V, Kokol P, Stiglic B, Rozman I (2002) Decision trees: An overview and their use in medicine. Journal of Medical Systems 26: 445-463.
10. Ritchie LJ, Tuokko H (2011) Clinical decision trees for predicting conversion from cognitive impairment no dementia (CIND) to dementia in a longitudinal population-based study. Archives of Clinical Neuropsychology 26: 16-25.
11. Eckerström C, Olsson E, Borga M, Ekholm S, Ribbelin S, et al. (2008) Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: The Göteborg MCI study. J Neurol Sci 272: 48-59.
12. Facal D, Valladares-Rodriguez S, Lojo-Seoane C, Pereiro AX, Anido-Rifon L, et al. (2019) Machine learning approaches to studying the role of cognitive reserve in conversion from mild cognitive impairment to dementia. International Journal of Geriatric Psychiatry 34: 941-949.
13. Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I, et al. (2011) Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. BMC Research Notes 4: 299.
14. Hajian-Tilaki K (2013) Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med 4: 627-635.
15. Kumar R, Indrayan A (2011) Receiver operating characteristic (ROC) curve for medical researchers. Indian Pediatrics 48: 277-287.

Comparative Machine Learning Approach in Dementia Patient Classification using Principal Component Analysis

Gopi Battineni[a], Nalini Chintalapudi and Francesco Amenta
e-Health and Telemedicine Centre, School of Pharmaceutical Sciences and Health Products, University of Camerino, Camerino, 62032, Italy
{gopi.battineni, nalini.chintalapudi, francesco.amenta}@unicam.it

Keywords: Dementia, Machine Learning, PCA, Model Prediction, Classifiers, AUC.

Abstract: Dementia is one of the brain diseases that were significantly affecting the global population. Mainly it is exposed to older people with an association of memory loss and thinking ability. Unfortunately, there are no proper medications for dementia prevention. Doctors are suggesting that early prediction of this disease can somehow help the patient by slowdown the dementia progress. Nowadays, many computer scientists were using machine learning (ML) algorithms and data-mining operations in the healthcare environment for predicting and diagnosing diseases. The current study designed to develop an ML model for better classification of patients associated with dementia. For that, we developed a feature extraction method with the involvement of three supervised ML techniques such as support vector machines (SVM), K-nearest neighbor (KNN), and logistic regression (LR). Principal component analysis (PCA) was selected to extract relevant features related to the targeted outcome. Performance measures were assessed with accuracy, precision, recall, and AUC values. The accuracy of SVM, LR, and KNN was found as 0.967, 0.983, and 0.976, respectively. The AUC of LR (0.997) and KNN (0.966) were recorded the highest values. With the highest AUC values, KNN and LR were considered optimal classifiers in dementia prediction.

1 INTRODUCTION

Dementia is a broad category of brain diseases, and this can be happening very often in older adults. Neurodegenerative disorders are one of the leading causes of the development of this disease (Barragán Martínez et al. 2019). There are different types of dementia, like Alzheimer's disease (AD), Lewy body dementia, and front temporal disorders. More than 50-60% of dementia was associated with AD type (McKhann et al. 2011). Sometimes AD can generate the loss of mental ability, individual thinking, memory loss, and visual perception (Barragán Martínez et al. 2019; Mahalingam and Chen 2019).

At present, there are is no proper prevention methods for dementia. Early prediction of dementia could enhance patient life expectancy and slow down the progress of this disease. Despite, machine learning (ML) is emerged as a branch of artificial intelligence (AI) and associated with techniques that allow computers to autonomous learning with nominal human involvement (Başţanlar and Özuysal 2014). Machine self-learning means that machines can be able to understand and identify input data. Ultimately, it can develop relations and predictions based on data feeding (Domingos 2012). Nowadays, these techniques are globally evolving health care from diagnosis to drug discovery.

Many studies were associated with the integration of ML approaches in automatic analysis of biomedical data. Glomerular diseases (Liu et al. 2017), detection of liver pathologies (Li, Jia, and Hu 2015), cancer predictions (Guyon et al. 2002; Kourou et al. 2015), Type 2 diabetes classifications (Luo 2016), dementia prediction (Battineni, Chintalapudi, and Amenta 2019), and cardiovascular disease (CVD) risk assessments (Kakadiaris et al. 2018) were the some of the applications in machine learning. Despite that, many researchers were attempted to find out the best ML algorithm in dementia predictions. For example, a study on the identification of developing dementia patients through ML obtained 84% accuracy (Mathotaarachchi et al. 2017). The risk factors associated with dementia were well-validated

[a] https://orcid.org/0000-0003-0603-2356

in (Aditya and Pande 2017; Pekkala et al. 2017), with the usage of supervised machine learning approaches. However, there has been little discussion on the involvement of feature extraction methods in dementia forecasting. As of this, the present study aimed to propose supervised machine learning algorithms for AD patients to understand the patterns associated with knowledge discovery in AD. We adopt longitudinal MRI data in demented and non-demented patients whose ages from 60 to 98. In this, we have studied the performance of three different models: SVM, Linear regression (LR), and K-nearest neighbor (KNN) algorithms to forecast dementia in older adults.

Table 1: Statistical report of OASIS longitudinal studies (where EDUC: education; SES: social-economic status; MMSE: mini-mental state examination; CDR: clinical dementia rating; e-TIV: estimated total intracranial volume; n-WBV: normalized whole brain volume; ASF: atlas scaling factor; D: demented; ND: Non-demented; Con: Converted.

N	Variable	Min-Max	Range (N)	Percentage
1	Subject ID	-	150	100
2	MRI ID	-	373	100
3	Group	-	D (146) ND (190) Con (37)	39.14 50.93 9.91
4	Visit	1-5	1-1.4 (150) 1.8-2.2(144) 3.0-3.4 (58) 3.8-5.0 (21)	40.21 38.60 15.54 5.62
5	MR delay	0-2639	0-880 (280) 881-1759 (71) 1760-2639 (22)	75.06 19.03 5.89
6	Sex	-	Male (160) Female (213)	42.89 57.10
7	Hand (R)	-	373	100
8	Age	60-98	60-73 (106) 74-85 (213) 86-98 (54)	28.41 57.10 14.47
9	EDUC	6-23	6-11 (23) 12-17 (270) 18-23 (80)	6.16 72.38 21.44
10	SES	1-5	1-3 (191) 4-5 (163)	51.20 43.69
11	MMSE	4-30	4-12.5 (2) 12.6-21.3 (33) 21.4-30 (336)	0.05 8.84 90.08
12	CDR	0-2	0-1(329) 1-2 (44)	88.19 11.81
13	e-TIV	1106-2004	1106-1555(263) 1556-2004(110)	70.51 29.49
14	n-WBV	0.644-0.837	373	100
15	ASF	0.876-1.587	0.87-1.23 (229) 1.23-1.58 (144)	61.39 38.61

2 MATERIALS AND METHODS

2.1 Data Selection

An open-access series of imaging studies (OASIS) dataset with 150 patients with at least 60years of age was considered (Smith 2009). Each patient exposed to at least two MRI sessions, and a total of 373 MRI sessions were analyzed. Current AD status (i.e., along with 15 independent variables) classified into three groups: Demented, Non-demented, and Converted, had mentioned in Table 1.

2.2 Feature Extraction

Feature extraction is a method that can be used to remove irrelevant (redundant) features from the actual dataset (Guyon and Elisseeff 2006). In model design, feature extraction is an essential step because the reduction of irrelevant or partially relevant features can tend to have a high-performance model. In this study, the selection of high correlated attributes was measured to conduct the feature extraction technique. The principal component analysis (PCA) method was adopted to reduce the actual dataset features (Ruby-Figueroa 2015).

We considered OASIS longitudinal dataset to find a combination of input attribute that matches actual data distribution. Feature extraction experiment was performed with the help of auto package PCA (auto.pca) in the 'R' platform (https://cran.r-project.org/web/packages/auto.pca/index.html).

2.3 Classifiers

2.3.1 Support Vector Machines (SVM)

SVM is a supervised machine learning (SML) approach; it is one of the highly used classification algorithms in machine learning (Wang and Lin 2014). In SVM, each data segment was represented as a single point in N-dimensional (where N is the total number of features in the actual dataset) space, with the forecasting of each element is being the estimation of specific coordinates. At that point, we perform classification action by finding the hyperplane (i.e., decision boundaries to classify data points) that correctly separates the output classes. The best hyper-plane can be chosen among the number of hyper-planes on the premise of the separation between the two categories that isolates. The plane, which has the highest margin between the two classes, is called the high margin hyper-plane.

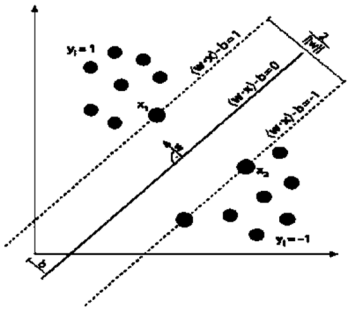

Figure 1: SVM representation example.

The hyperplane can be described by w. x + b = 0, where w is a normal vector and $\frac{b}{\|w\|}$ is the hyperplane offset along w vector.

For n data points, SVM defined as(x1, y1)... (xn, yn), and optimization can be written as

If $yn(w.xn+b)-1=0$ then (xn,yn) are support vectors and save parameters w, b
else if $yn(w.xn+b)-1>0$ then save parameters w, b
else if $yn(w.xn+b)-1<0$ then update parameters w, b

In the example (Figure 1), two hyperplanes are passing through support vectors (y=±1): (w. x) − b = −1 and (w. x) − b = 1. The distance between the two hyperplanes and origin is

$$\frac{1-b}{\|w\|}-\frac{-1-b}{\|w\|} \text{ and margin } M=\frac{2}{\|w\|}$$

2.3.2 Linear Regression (LR)

LR is utilized to finding the linear relation between the target variable and the predictor variable. It explores the relationship between two variables by the linear equation to the test data. One variable is viewed as a logical type, and the other variable is considered to be a dependent type (Kumar 2006).

In the present study, a dataset of 150 patients' information (trained data) about the relationship between "14 different features" and "group attribute." We aimed to design a model that can predict a patient group based on other features. A regression line was obtained (with minimum error) by using trained data. Thus, if trained data exposed to the feature extraction technique, the model should predict the patient group with less or no error.

$$y(pred)=b_0+b_1 \times x \text{ here } b_0, \text{ and } b_1 \text{ need to select for error minimization}$$

$$Error=\sum_{n=1}^{k}(actutal\ input-actutal\ output)**2, \text{ and}$$

$$coeffeciant\ b_1=\frac{\sum_{n=1}^{k}(x_i-x)(y_i-y)}{\sum_{n=1}^{k}(x_i-x)^2}$$

2.3.3 K-nearest Neighbor (KNN)

KNN is easy to understand and address the issues of classification and regression. It uses similar features to predict the estimations of new data points. Therefore, the new data point will be allotted a value based on how closely it coordinates the points in the trained dataset (Chen, Li, and Tang 2013).

3 RESULTS AND DISCUSSION

3.1 Model Outcome

A comparison of the three machine-learning classifiers' performance was done. Initially, OASIS longitudinal dataset exposed to the R platform (Figure 2) and model testing conducted with two datasets: an actual data set and dataset after PCA. Preprocessing involved with the prediction of missing values by the imputation of K-NN. Feature extraction was performed with the help of the PCA technique. Highly correlated features were selected for better outcomes. Each ML classifier was evaluated independently by cross-validation techniques (with k=10).

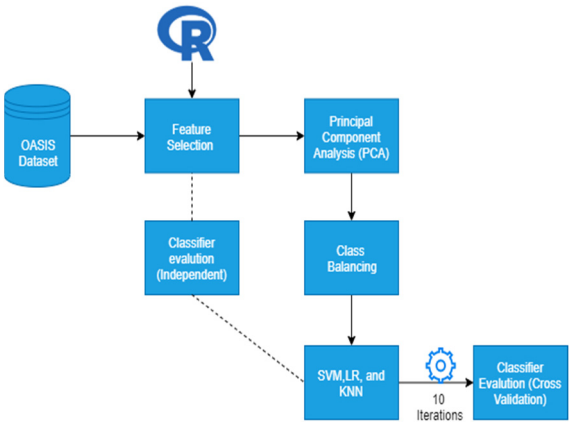

Figure 2: Experimental workflow and design.

3.2 Performance Parameters

To predict specific patient associated with AD or not, a predictive model should be correctly classified the instances. Accuracy (A) is a ratio of correctly predicted outcomes to a total number of input samples (Powers 2011). Three supervised ML techniques (SVM, LR, and KNN) were used to develop predictive models (Table 2). The performance of three predictive models was analyzed using parameters such as precision (Davis and Goadrich

2006), recall, and area under the curve (AUC) (Davis and Goadrich 2006; Powers 2011). LR produced the highest accuracy of about 98.3%. Followed to LR, KNN and SVM produced accuracy about 97.6%, and 96.7%, respectively. Three models were generating similar accuracy rates. Sometimes, accuracy is not only enough to judge the model performance. Therefore, analysis of other parameters such as precision, recall, and AUC is mandatory to define model validation.

Precision can define positive outcomes from total predicted positive instances. In this study, we found similar accuracy for two models (LR and KNN) about 98± 0.04%. When compared with the other two models, SVM was producing a low positive prediction rate of 97.1%. On the other hand, recall (sensitivity) can define true positives from total actual positives. Both precision and recall are based on the understanding of the relevance of positive outcomes. From Table2, the sensitivity for LR predictive model found at about 97.4%. Alternatively, KNN was with the highest sensitivity rate of 98.3%, and SVM with the lowest sensitivity rate of 96.6% can found. Despite this, in machine learning, AUC can help to overcome classification problems. It is one of the key performance tools for model performance checks. Generally, the AUC was ranging in between [0, 1]. By definition, if AUC ≈ 1, then the model was correctly distinguishing the target class. The AUC values of LR, KNN, and SVM were 99.7%, 99.6%, and 98.3%, respectively.

Table 2: Performance metrics of different predictive models.

Model	Accuracy	Precision	Recall	AUC
SVM	0.967	0.971	0.966	0.983
LR	0.983	0.986	0.974	0.997
KNN	0.976	0.982	0.983	0.996

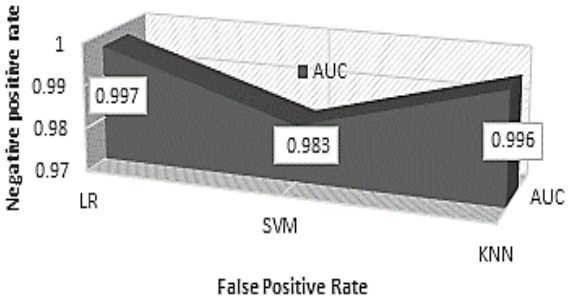

Figure 3: Graphical representation of AUC values.

4 CONCLUSIONS

In this study, three supervised ML algorithms (SVM, LR, and KNN) were defined to classify dementia patients. Feature extraction performed using the principal component analysis method using the R platform. Different performance parameters set was defined the model validation. Results validated that the three models are accurately classifying dementia patients with better rates from 96.7-98.3%. In unbalanced datasets, accuracy is not only the parameter to validate the model. Therefore, other metrics, such as precision, recall, and AUC, were also considered. The AUC of LR and KNN reached the highest value of one, such that these two predictive models were well classified the dementia patients. This work is concluding that employment PCA techniques were much better than the manual selection of attributes with minimum medical knowledge. Therefore, with limited features and integration of the PCA method, we were achieved better accuracy rates when compared with previous studies in dementia classifications.

CONFLICTS OF INTEREST

The authors do not possess any conflicts during the publication.

ACKNOWLEDGMENTS

We are thankful to the Principal Investigators: D. Marcus, R, Buckner, J. Csernansky, and J. Morris, to provide access to OASIS longitudinal studies.

REFERENCES

Aditya, C. R., and M. B.Sanjay Pande. 2017. “Devising an Interpretable Calibrated Scale to Quantitatively Assess the Dementia Stage of Subjects with Alzheimer’s Disease: A Machine Learning Approach.” Informatics in Medicine Unlocked.

Barragán Martínez, D., M. A. García Soldevilla, A. Parra Santiago, and J. Tejeiro Martínez. 2019. “Alzheimer’s Disease.” Medicine (Spain).

Baştanlar, Yalin, and Mustafa Özuysal. 2014. “Introduction to Machine Learning.” Methods in Molecular Biology.

Battineni, Gopi, Nalini Chintalapudi, and Francesco Amenta. 2019. “Machine Learning in Medicine: Performance Calculation of Dementia Prediction by

Support Vector Machines (SVM).” Informatics in Medicine Unlocked.

Chen, Qifeng, Dingzeyu Li, and Chi Keung Tang. 2013. “KNN Matting.” IEEE Transactions on Pattern Analysis and Machine Intelligence.

Davis, Jesse, and Mark Goadrich. 2006. “The Relationship between Precision-Recall and ROC Curves.” In Proceedings of the 23rd International Conference on Machine Learning - ICML ’06,.

Domingos, Pedro. 2012. “A Few Useful Things to Know about Machine Learning.” Communications of the ACM.

Guyon, Isabelle, and Andre Elisseeff. 2006. “Feature Extraction, Foundations and Applications: An Introduction to Feature Extraction.” Studies in Fuzziness and Soft Computing.

Guyon, Isabelle, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. 2002. “Gene Selection for Cancer Classification Using Support Vector Machines.” Machine Learning.

Kakadiaris, Ioannis A. et al. 2018. “Machine Learning Outperforms ACC/AHA CVD Risk Calculator in MESA.” Journal of the American Heart Association.

Kourou, Konstantina et al. 2015. “Machine Learning Applications in Cancer Prognosis and Prediction.” Computational and Structural Biotechnology Journal.

Kumar, K. Vasanth. 2006. “Linear and Non-Linear Regression Analysis for the Sorption Kinetics of Methylene Blue onto Activated Carbon.” Journal of Hazardous Materials.

Li, Wen, Fucang Jia, and Qingmao Hu. 2015. “Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks.” Journal of Computer and Communications.

Liu, Xun et al. 2017. “Improving Precision of Glomerular Filtration Rate Estimating Model by Ensemble Learning.” Journal of Translational Medicine 15(1): 1–5.

Luo, Gang. 2016. “Automatically Explaining Machine Learning Prediction Results: A Demonstration on Type 2 Diabetes Risk Prediction.” Health Information Science and Systems.

Mahalingam, Sowmya, and Ming Kai Chen. 2019. “Neuroimaging in Dementias.” Seminars in Neurology.

Mathotaarachchi, Sulantha et al. 2017. “Identifying Incipient Dementia Individuals Using Machine Learning and Amyloid Imaging.” Neurobiology of Aging.

McKhann, Guy M. et al. 2011. “The Diagnosis of Dementia Due to Alzheimer’s Disease: Recommendations from the National Institute on Aging-Alzheimer’s Association Workgroups on Diagnostic Guidelines for Alzheimer’s Disease.” Alzheimer’s and Dementia.

Pekkala, Timo et al. 2017. “Development of a Late-Life Dementia Prediction Index with Supervised Machine Learning in the Population-Based CAIDE Study.” Journal of Alzheimer’s Disease 55(3): 1055–67.

Powers, David M. W. 2011. “Evaluation: From Precision, Recall And F-Measure To Roc, Informedness, Markedness & Correlation.” Journal of Machine Learning Technology.

Ruby-Figueroa, René. 2015. “Principal Component Analysis (PCA).” In Encyclopedia of Membranes,.

Smith, Susan Spivock. 2009. Predicting Alzheimer’s Dementia Mortality Using Medicare Outcome Assessment & Information Set (oasis) “Predicting Alzheimer’s Dementia Mortality Using Medicare Outcome Assessment and Information Set (OASIS).”

Wang, Po Wei, and Chih Jen Lin. 2014. “Support Vector Machines.” In Data Classification: Algorithms and Applications,.

Journal of
Clinical Medicine

MDPI

Article

A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects

Gopi Battineni *, Nalini Chintalapudi, Francesco Amenta and Enea Traini

Telemedicine and Telepharmacy Center, School of Medicinal and Health Products Sciences, University of Camerino, 62032 Camerino, Italy; nalini.chintalapudi@unicam.it (N.C.); francesco.amenta@unicam.it (F.A.); enea.traini@unicam.it (E.T.)
* Correspondence: gopi.battineni@unicam.it; Tel.: +39-33-3172-8206

Received: 27 April 2020; Accepted: 29 May 2020; Published: 8 July 2020

Abstract: Increasing evidence suggests the utility of magnetic resonance imaging (MRI) as an important technique for the diagnosis of Alzheimer's disease (AD) and for predicting the onset of this neurodegenerative disorder. In this study, we present a sophisticated machine learning (ML) model of great accuracy to diagnose the early stages of AD. A total of 373 MRI tests belonging to 150 subjects (age ≥ 60) were examined and analyzed in parallel with fourteen distinct features related to standard AD diagnosis. Four ML models, such as naive Bayes (NB), artificial neural networks (ANN), K-nearest neighbor (KNN), and support-vector machines (SVM), and the receiver operating characteristic (ROC) curve metric were used to validate the model performance. Each model evaluation was done in three independent experiments. In the first experiment, a manual feature selection was used for model training, and ANN generated the highest accuracy in terms of ROC (0.812). In the second experiment, automatic feature selection was conducted by wrapping methods, and the NB achieved the highest ROC of 0.942. The last experiment consisted of an ensemble or hybrid modeling developed to combine the four models. This approach resulted in an improved accuracy ROC of 0.991. We conclude that the involvement of ensemble modeling, coupled with selective features, can predict with better accuracy the development of AD at an early stage.

Keywords: MRI; machine learning; feature selection; ensemble methods; ROC

1. Introduction

Adult-onset dementia disorders are among the prevalent global medical issues in industrialized countries that have a high impact on individuals' lifestyles. These disorders represent a great challenge for the community over their advancement from early diagnosis to end of life [1]. Statistical studies have estimated that every three seconds, a new dementia case is developing in the world. This means that approximately 50 million patients are suffering from this disease worldwide [1,2]. These numbers might double every twenty years and possibly reach 100 million patients by 2040.

Dementia is a syndrome that develops largely in older adults. It affects brain functionality, daily activities, and communication efficiency [1,3]. Alzheimer's disease (AD) represents the prevalent form of adult-onset dementias. Some studies have highlighted that the early diagnosis of dementia is useful for starting treatments and for predicting outcomes of the disease but did not offer reliable methods for the early diagnosis of AD [4–6]. At the same time, some forms of mild cognitive impairment (MCI) do not evolve into overt dementia, whereas other forms of MCI represent a very mild form of AD [7]. In view of this, advanced computer techniques may represent a tool for the early diagnosis of AD and for predicting the evolution of prodromal forms of the disease or MCI into dementia.

Magnetic resonance imaging (MRI) techniques are becoming a relevant tool for prodromal AD and MCI evaluation [8]. A few studies based on the comparative analysis of cognitive testing and neuroimaging have hypothesized that AD neuroimaging may be enough to predict disease [8–10]. On the other hand, dementia forecasting with machine learning (ML) is becoming a more diffused approach in clinical practice [11]. In spite of the practical interest to quantify AD evolutions based on MRI data, only a few studies have calculated AD incidence rates based on MRI.

Neuroimaging and primarily MRI provide essential information for AD dementia classification and prediction [12–14]. ML models, coupled with MRI information, can provide high diagnostic accuracy of age-related cognitive decline (ARCD) in dementia subjects [15]. It has been hypothesized that ML-supervised methods generate the knowledge of features necessary to correlate AD sample data [16]. It is also reported that logistic regression, coupled with cross-validation, can enhance the accuracy of AD prediction by speech amalgamation [17]. On the other hand, support vectors, along with feature reduction techniques, were able to classify dementia subjects with 70% accuracy [4].

The present study was designed to detect AD based on MRI findings along with the use of four ML models, such as naive Bayes, neural networks, k-nearest neighbor, and support vectors. Each model was validated separately by tenfold cross validation (CV). The receiver operating characteristic (ROC) curve value was used to evaluate the model accuracy. Three individual experiments were designed to test the model, and model performance was separately evaluated with given MRI characteristic information. The experiments that were done included

1. Models with manual selection of MRI features,
2. Models with automatic feature selection, and
3. A single model with ensemble learning or hybrid modeling.

The subsequent part of this paper is organized as follows. In Section 2, subject information of MRI features, feature selection techniques, and adopted models of AD prediction are analyzed. In Section 3, the experimental results of the four models are presented. In Section 4, each model is discussed and compared by accuracy and ROC parameters. Finally, Section 5 summarizes the main results of the present work.

2. Materials and Methods

2.1. Subjects

A longitudinal collection of 150 subjects and 373 MRI sessions was considered for this study. Each subject had undergone full screening of complete clinical assessment conducted at the Alzheimer's Disease Research Center (ADRC) of Washington University. All subjects included, both men and women, were right-handed with a minimum age of 60 years and a maximum age of 96 years [18]. The subjects included 72 nondemented (ND) individuals and 64 demented (D) individuals (including 51 with mild to moderate AD). The remaining 14 subjects were identified as nondemented at the initial visit but resulted as demented when examined in subsequent visits. These subjects were defined as belonging to the converted (C) type.

Subjects undergoing age-related normal brain changes, such as leukoaraiosis, mild atrophy, and regular dementia cases of AD, were included in this study. All MRI sessions were done in one year. These sessions were followed by clinical tests made on 0–352 days (mean—111 days) after MRI. Twelve confirmed demented subjects were scanned with a delay ranging from 374 to 924 days (mean—653 days) and were included in this study as they had a clinical dementia rating (CDR) higher than zero in previous clinical assessments. Two nondemented subjects, with a scan delay range of 392 to 431 days, were also included because they did not display dementia symptoms in successive clinical evaluations. With this approach, each subject had at least two individual scan sessions with a mean delay of 719 days (range: 183–1707 days) between each visit. The demographic characteristics of the subjects are presented in Table 1. Diagnostic characteristics of subjects of different age groups on the initial clinical visit are detailed in Table 2.

Table 1. Demographic characteristics of the subjects investigated.

Subjects	78 D	72 ND
Male	40 D	22 ND
Female	38 D	50 ND
Age range (years)	60–96	
Median	77.0	
Mean ± SD	77.01 ± 7.3	

D: demented; ND: nondemented; SD: standard deviation.

Table 2. Age and characteristics of the individuals investigated on the first clinical visit [18].

	Non-Demented						Demented				
Age Group	**N**	**n**	**Mean**	**Male**	**Female**	**Convert**	**n**	**Mean**	**Male**	**Female**	**CDR 0.5/1**
60s	34	23	65.71	6	17	3	11	65.67	8	3	8/3
70s	71	35	74.91	11	24	4	36	73.97	20	16	29/7
80s	41	26	84.30	9	17	7	15	82.33	7	8	13/2
90s	4	2	92.50	0	2	0	2	93.00	1	1	1/1
Total	150	86	75.82	26	59	14	64	74.95	36	29	52/13

CDR: clinical dementia rating.

2.2. MRI Acquisition Methods

Three or four separate T1-weighted MRI scans were acquired with a 1.5T Siemens Vision MRI scanner for each single subject. A high-resolution Magnetization Prepared Rapid Acquired Gradient Echo (MP-RAGE) was used to handle the classification of subject scans. For each subject, separate scan files were generated using Siemens proprietary IMA to 16-bit NiFTI1 format by employing the traditional conversion program. The MR images were corrected for interscan head rotation and wrapped spatially into atlas space. The transformation outcome placed the brains in a correlated coordinate system, with the bounding box as the actual atlas. With this procedure, every image was turned out as a unique, high contrast, averaged MP-RAGE image in an atlas-space. The insight explanation on image acquisition and postprocessing steps are detailed in [18].

The estimated total intracranial volume (eTIV) was defined manually across intracranial volume on an atlas. Normalized whole-brain volume (nWBV) was computed with the FAST program of the FSL software suite. Image segmentation was done to classify brain tissue as spinal fluid or white or gray matter. This segmentation process was iteratively assigned as voxels to tissue classes based on high probability estimates of hidden Markov random field models. In the end, nWBV was calculated as the proportion of accumulated voxels across the brain mask, and the normalized volume was expressed in a percentage of total gray and white matter voxels of eTIV [18]. The atrophy rates were estimated as the slope of the line that connects to nWBV. Details of the MRI acquisition characteristics are summarized in Table 3.

Table 3. Magnetic resonance imaging (MRI) acquisition details [17].

MR Characteristics	Values
Sequence	MP-Rage
TR (repetition time)	9.7 msec
TE (echo time)	4.0 msec
Flip angle	10°
TI	20 msec
TD	200 msec
Orientation	Sagittal
Thickness	1.25 mm
Gap	0 mm
Slice number	128
Resolution	256 × 256 (1 × 1 mm)

MP-RAGE: Magnetization Prepared Rapid Acquired Gradient Echo; TI: Inversion time; TD: Dead time.

2.3. Feature Description

The dataset included 373 pieces of MRI information with 15 independent characteristics (attributes). The description of each feature is detailed in Table 4. The subject attribute "Group" specifies the dementia status (Demented/Nondemented) and is considered as an outcome of a binary classifier. In this study, scoring rules of Clinical Dementia Rating (CDR), Mini-Mental State Evaluation (MMSE), and Visit were used to determine the dementia status (Table 5). All subjects underwent similar procedures and received the same tests, including MMSE.

Table 4. Dataset feature description.

Features	Description
Subject ID	Subject identification number
MRI ID	Image identification number of an individual subject
Visit	Number of subject visits
Gender	Male/Female
Hand	Right/Left handed
EDUC	Subject education level (in years)
SES	Socioeconomic status
MMSE	Mini-mental state examination score
CDR	Clinical dementia rating score
eTIV	Estimated total intracranial volume result
nWBV	Normalized whole brain volume result
ASF	Atlas scaling factor
Age	Subject age while scanning
Group	Demented/Nondemented/Converted
MR delay	Magnetic resonance (MR) delay is the delay time that is prior to the image procurement

Table 5. Scoring rules.

Features	Range	Condition
CDR	0–3	None—0, Very mild—0.5, Mild—1, Moderate—2, Extreme—3
MMSE	1–30	Extreme impairment (<10) Moderate dementia (10–19) Early-stage Alzheimer's aliment (19–24) Normal (>25)
Visit	0 or 1	Low status—0High status—1

2.4. Feature Selection

In this step, the machine performed an autonomous selection of input features that correlates to the subject group [19]. Selection techniques are largely used and standardized to reduce unnecessary features and to enhance model accuracy [20]. Moreover, this approach measures the relationship between independent variables and the target outcome. Feature selection can be conducted by three approaches, namely, filtering, regularization, and wrapping [20,21]. In this study, the wrapping technique was used because it amplifies model performance with limited features.

2.5. Feature Importance

This method results in a "feature score" assigned to independent characteristics and a defined score to each characteristic that is highly correlated with the subject "group". The correlation between each characteristic-associated group variable is shown in Figure 1. The CDR rating was excluded during model development because it did not have the highest relevance, but it helps in subject groupings.

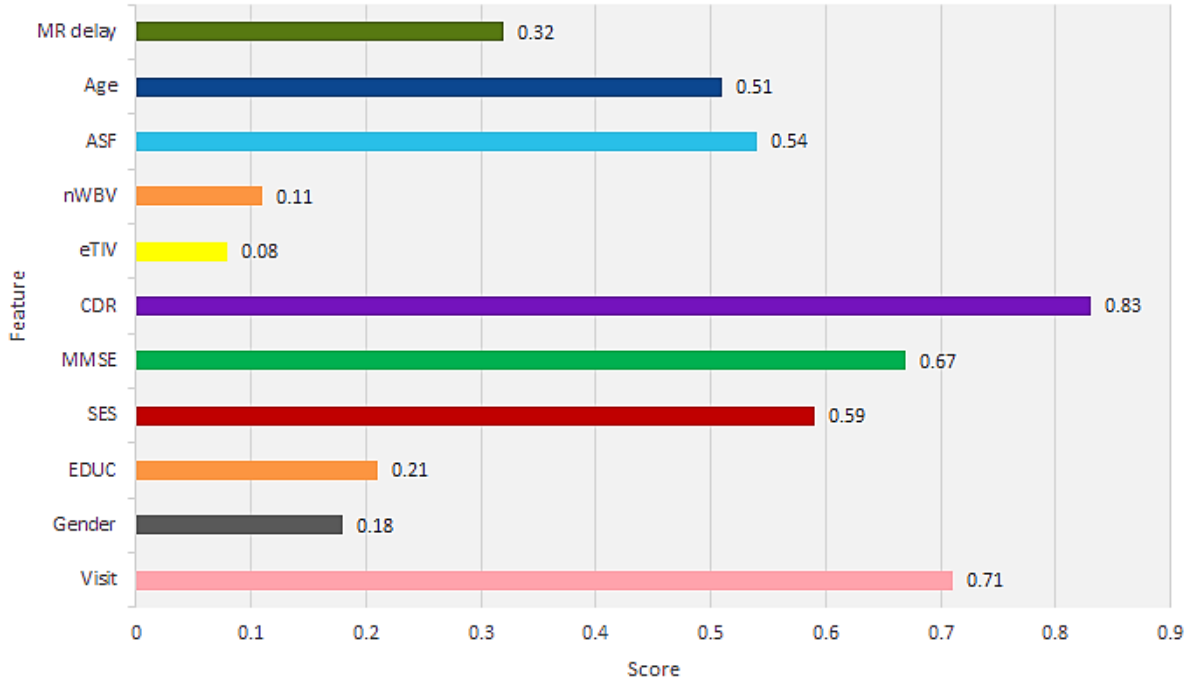

Figure 1. Individual feature scores.

2.6. Feature Selection with Wrapping

In the wrapping method, feature search represents a big challenge in calculating model accuracy [22]. Feature selection can be made as either step backward or forward, and exhaustive. Feature search helps the identification of primary features in the enhancement of model performance. The MRI characteristics with a correlation of at least 0.5 can automatically help to develop a model. Figure 2 shows the scatter plot of feature results following the wrapping method.

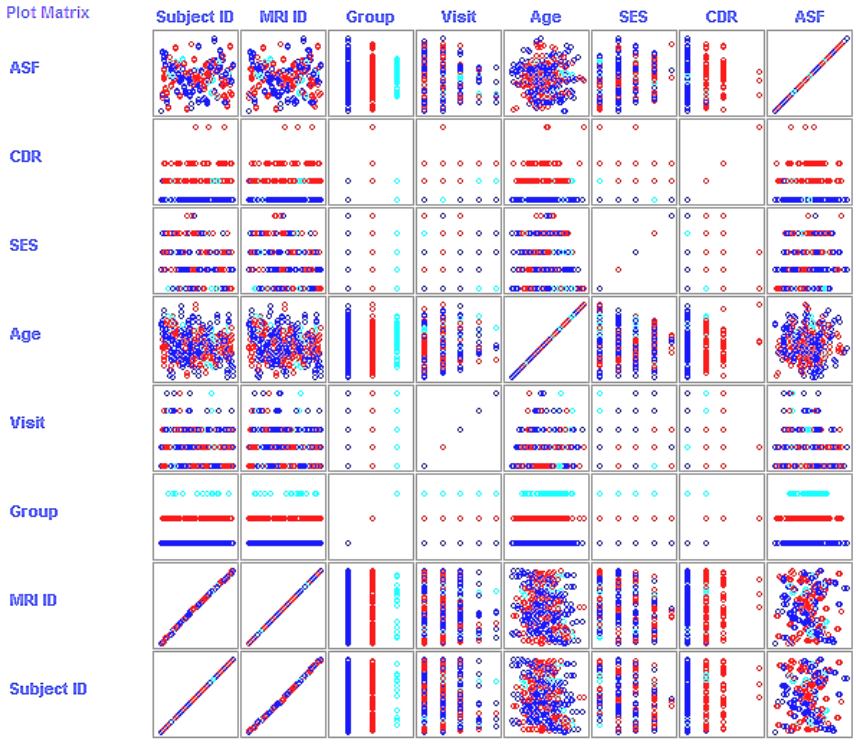

Figure 2. Scatter plot of selective features. Blue dots (ND), red dots (D), light blue dots (C).

2.7. Model Classifiers

The purpose of the present study is to develop a sophisticated ML model of dementia detection in aged subjects based on MRI findings. It is unanimously recognized that advanced age is the greatest risk factor for AD [23]. In this work, four popular ML models such as neural networks (NN) [24], k-nearest neighbor (KNN) [25], naive Bayes (NB) [26], and support vector machines (SVM) [27] were used. These models were selected because of the easy implementation and production of high accuracy during model development. A short description of each model is provided below.

Neural networks are able to learn from independent features to predict target outcomes. They allow the design of an artificial neural network (ANN) to admit machines with the integration of new data [28]. ANN is largely associated with clustering (combining the unlabeled data of similar features) and classification (trained data grouping) procedures. One of the conventional and popular neural networks is the multilayer perception (MLP) type, which includes one or more neuron layers [29]. These neuron layers largely intervene to develop predictive models for forecasting clinical diagnoses [30].

KNN is a comprehensive model used to perform both regression and classification problems [25]. It is also called a "lazy" learner because instead of the model development approach, it calculates the nearest neighbors during prediction. When KNN initiates predictive analysis, it searches for nearest neighbors (i.e., K) in the trained dataset. The neighboring distance is then calculated with the Euclidean function, which defines the similarity between two points [31].

NB is a probabilistic model that predicts output based on Bayes' principle. It calculates the outcome value of individual groups, which is not associated with other variables [26]. Due to its simplicity during target prediction, it has become popular in classification and multiclass predictions [32].

SVM is another algorithm developed for subject classification. In SVM plotting, dataset features are described in n-dimensional space (here, "n" is feature count), and classification is done to decide the optimal hyperplane [27]. In more detail, SVM produces an optimal hyperplane with the trained label data that classifies new feature examples. This hyperplane is a line of binary classification and tuning parameters, such as "kernel", "gamma", and "C", that can help to improve SVM model performance [33,34].

2.8. Performance Measures

After model development, it is important to evaluate individual model performance. This is calculated through the prediction of the trained model of a test dataset. Different parameters like accuracy (A_{cc}), sensitivity (S_e), specificity (S_p), and receiver operating characteristic (ROC) curve define model performance. To calculate each parameter, the confusion matrix (CM) was used to identify misclassifications in tabular form (Table 6). A subject is true-positive when it is diagnosed as demented (X = D), and a subject is true-negative when it diagnosed as "nondemented" (Y = ND).

Table 6. Simple confusion matrix (CM).

Prediction	X	Y
X = D	TP	FN
Y = ND	FP	TN

D: demented; ND: nondemented; TP: true-positive; TN: true-negative; FP: false-positive; FN: false-negative.

The performance measures evaluated by CM are given below:

- Accuracy: Percentage of total true predicted outcomes from total outcomes, i.e., Accuracy (%) = $\left(\frac{TP+TN}{TP+TN+FP+FN} * 100\right)$.
- Sensitivity: It measures the proportion of true-positives, i.e., Sensitivity (%) = $\left(\frac{TP}{TP+FN} * 100\right)$.
- Specificity: It measures the proportion of true-negatives, i.e., Specificity (%) = $\left(\frac{TN}{TP+FN} * 100\right)$.

- ROC: ROC is a performance visualization tool of binary classifiers with the false-positive rate (FPR) on the *X*-axis and the true-positive rate (TPR) on the *Y*-axis. In this study, we mainly highlight the ROC value to determine model performance because it is frequently used in medical diagnosis.

2.9. Model Validation and Framework

Model validation can be done by either holdout (spilt) or cross-validation (CV) techniques. During his study, we adopted the CV technique because of its popularity in target prediction, with low bias. Simultaneously, it also applies a resampling method with limited features during model validation [35]. In CV, the dataset is distributed into N-folds of equal size. The first fold is used for validation, and the remaining k-1 folds are kept for training. The model framework used during simulation is represented in Figure 3.

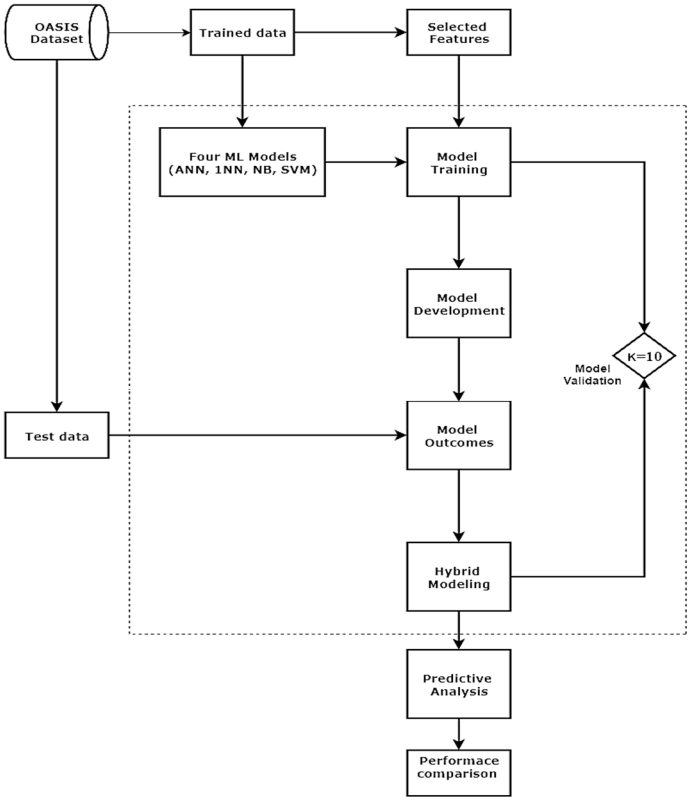

Figure 3. Model framework for evaluating predictive classifications.

2.10. Experiments Design

A large number of MRIs for a low number of subjects could generate bias in dementia detection. Therefore, we considered final MRI scans that define the status of each subject. Three experiments were conducted, including manual and automatic feature selection techniques.

In the first experiment, model training was done using the original dataset with manual feature selection. In ANN, the number of layers (N) is used as a search parameter during model evaluation. In KNN, k is tuned to one (i.e., 1NN). In SVM, the linear kernel coupled to regularization parameter "C" and a standard deviation of radial basis function "r" are implemented in model tuning. Finally, model validation was done with a 10-fold CV to avoid data fitting issues [36]. The model performance was, therefore, assessed by the above parameters.

In the second experiment, limited features that occurred as the result of wrapping were considered for conducting model training. For NB and KNN, an exhaustive search was used to calculate model accuracy with potential feature alliance in order to select the best of them [37]. In SVM, genetic algorithms (GAs) were used for the feature search. GAs are frequently applied in bioinformatics to

generated models with high accuracy [38]. For ANN, the feature search was excluded, and the search consisted of the identification of the hidden neuron layers. Model tuning was adjusted by maintaining batch size as 100 in NB, (C, gamma) as (1.0, 1.0×10^{-12}) in SVM, and k = 1 in KNN. MRI characteristics that were highly correlated (≥0.5) with subject groups were selected (see Figure 2).

In the third experiment, the four models were combined to develop an ensemble or hybrid model. By doing this, there is the advantage of getting a high prediction accuracy of the adopted dataset. Moreover, combining several models can enable noise reduction (bagging), low bias (boosting), and better predictions (voting). We used a voting technique in this experiment because of the capability to create standalone models from trained data [39].

3. Results

3.1. Experiment 1: Handling of the Feature Set Prior to Autonomous Feature Selection

Table 7 summarizes the performance outcomes of the four models in manual feature selection. The CDR rating was excluded as it represents a dementia factor that can affect model accuracy. From the performance comparison matrix, it can be seen that the 1NN model offers better performance compared to the other tested models in terms of accuracy, sensitivity, and specificity. As already mentioned, the ROC curve plays a relevant role in diagnostic assessments to differentiate the true state subjects and to find optimal cutoff values. Moreover, a higher ROC offers better dementia prediction in given subjects [40]. In view of this, the ANN model correctly discriminates against the true demented subjects, with a ROC of 0.812. The ROC of NB, 1NN, and SVM models produced ROCs of 0.753, 0.787, and 0.796, respectively.

Table 7. Performance comparison matrix (4 × 4) of four classifiers.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	ROC
NB	88.76	82.43	85.72	0.753
ANN	83.56	89.92	88.84	0.812
1NN	91.32	89.92	89.56	0.787
SVM	89.67	89.24	89.45	0.796

NB: naive Bayes; ANN: artificial neural networks; 1NN: 1-nearest neighbor; SVM: support vector machines; ROC: Receiver operating charactersitcs.

3.2. Experiment 2: Automatic Feature Selection with Wrapping

Table 8 shows the model performance outcomes obtained with automatic feature selection. With this approach, progress in terms of accuracy and ROC compared to manual feature selection was noticeable. SVM resulted in high accuracy (96.12%), and 1NN, NB, and ANN produced an accuracy of 95.92%, 93.44%, and 83.56%, respectively. With regard to ROC, NB was a better diagnosis predictor, with 0.942, followed by 1NN, SVM, and ANN, with 0.916, 0.834, and 0.817, respectively.

The results of the present experiment, in which performance results were better than those obtained in the previous one, stimulated the identification of other approaches for maximizing prediction accuracy. We, therefore, extended our work to explore the outcomes of joint modeling with limited features.

Table 8. Model performance evaluation after feature selection (with selective features).

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	ROC
NB	93.44	98.21	97.32	0.942
ANN	83.56	89.92	88.84	0.817
1NN	95.92	94.92	97.36	0.916
SVM	96.12	94.94	98.23	0.834

3.3. Experiment 3: AD Predictions with Hybrid Modeling

To check if a model correctly predicted the target variable (occurrence of dementia), a confusion matrix was used. In this analysis, vertical labeling presents actual subjects, and horizontal labeling presents predicted subjects. As shown in Figure 4, 76 subjects were correctly predicted as AD among 78 subjects, and 71 subjects were correctly predicted as non-AD among 72. Collectively, 147 subjects were properly predicted out of 150 subjects. This results in 98% accuracy. For reaching these conclusions, a hybrid-modeling technique, combining the four adopted models, was introduced.

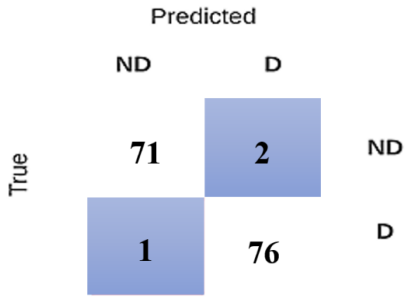

Figure 4. Confusion matrix outcome of the hybrid model (D: Demented; ND: Non demented).

The performance of the individual subject group is presented in Table 9. Nondemented and demented subjects were correctly diagnosed with 98.6% and 97.4% accuracy, respectively. The weighted average ROC curve of both subjects nearly touches one. Hence, maximum AD subject predictions have been made without bias because of hybrid modeling. The sensitivity and specificity rates produced were 98.05% and 98%, respectively. The ROC curve of the hybrid model is shown in Figure 5. Based on the evaluation of performance differences in the above three experiments, the intervention of hybrid modeling with limited features resulted in being good practice in AD-related studies.

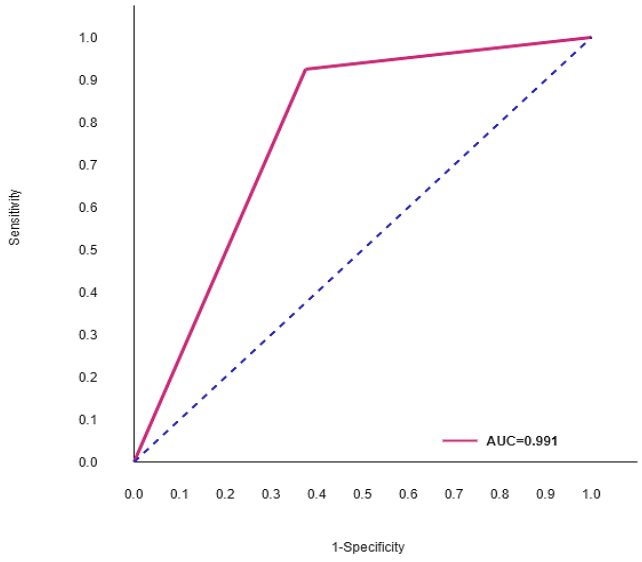

Figure 5. Receiver operating characteristic (ROC) curve of the hybrid model.

Table 9. Performance statistics of hybrid modeling.

Accuracy (%)	Sensitivity (%)	Specificity (%)	ROC	Class
98.6	98.7	98.6	0.992	ND
97.4	97.4	97.4	0.989	D
98.0	98.05	98.0	0.991	Weighted average

4. Discussion

ML models are highly acknowledged in real-time clinical practice and also in diagnosis and AD treatment selection [41]. Several MRI works have been integrated into ML models to make AD predictions [12,17,42], but there has been no comprehensive model to amplify model accuracy. In view of this, we introduced a hybrid model to enhance the precise detection of AD based on the analysis of MRIs.

In this paper, the significance of joint ML modeling for AD-onset prediction in elderly people has been demonstrated. Three different experiments were conducted, including manual and automatic feature selection techniques. Fourteen independent MRI features were used to identify the AD group using standard diagnostic approaches. Four supervised predictive models (NB, ANN, KNN, and SVM) were used, and the obtained results indicate the prediction accuracy of each model, constantly increasing between experiments. Figure 6 compares the prediction accuracy of the three experiments. 1NN generated 91.32% accuracy by manual feature selection; SVM had a high 96.12% accuracy by automatic feature selection, whereas joint or hybrid modeling enabled 98% accuracy in predicting AD in older adults. The outcomes suggest that joint modeling, with limited features, is a best practice to assess AD-onset by subject prediction.

In the first experiment, all the designed classifiers revealed enough performance values in terms of true-positive rates (sensitivity). ANN and 1NN produced the highest sensitivity (89.92%), followed by SVM (89.24%) and NB (82.43%). As mentioned, ROC curve values between 0.5 and 0.7 indicate low prediction accuracy, between 0.7 and 0.9 indicate moderate prediction accuracy, and between 0.9 and 1 indicate high prediction accuracy [43]. From Table 7, it is obvious that the four adopted models produce moderate prediction accuracy when checking with manual feature selection.

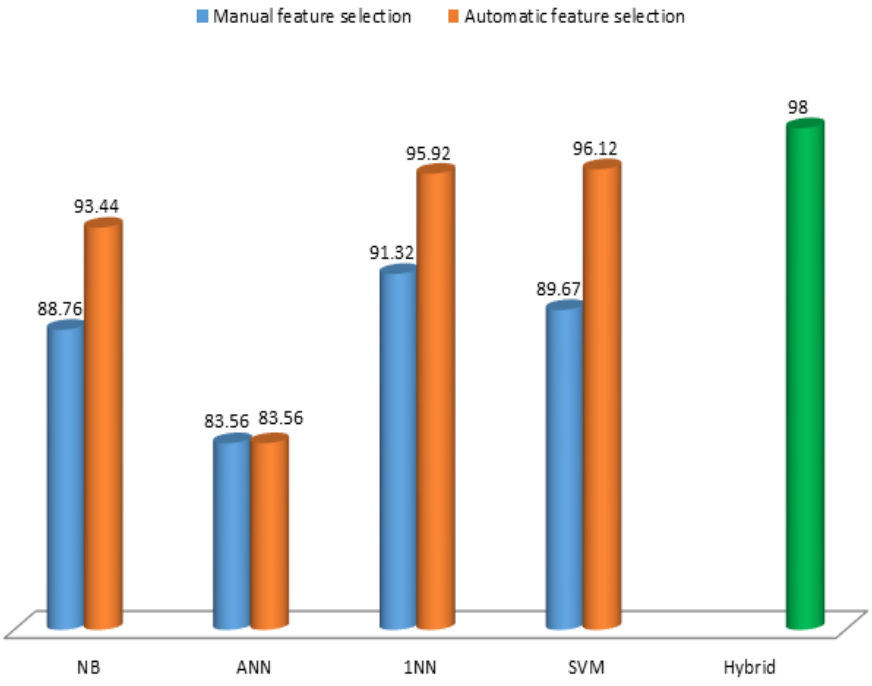

Figure 6. Prediction accuracy (in %) comparisons of three experiments.

To amplify model performances, the second experiment was conducted with selective features after wrapping. This resulted in NB of 98.21% sensitivity, followed in descending order by SVM (94.94%), ANN (94.92%), and 1NN (89.92%). Both NB and 1NN predict subject class in a comparatively better manner, with ROC of 0.942 and 0.916, respectively. However, we argued that there could be other possibilities for enhancing prediction accuracy to values higher than those identified in the above two experiments. To support this claim, a hybrid model was developed by combining the four

investigated models. A simulation of four recruited models was then performed, and thanks to this approach, the sensitivity of the model attained the highest predicted value of 97.4%, and its ROC was nearly equal to one (Figure 7).

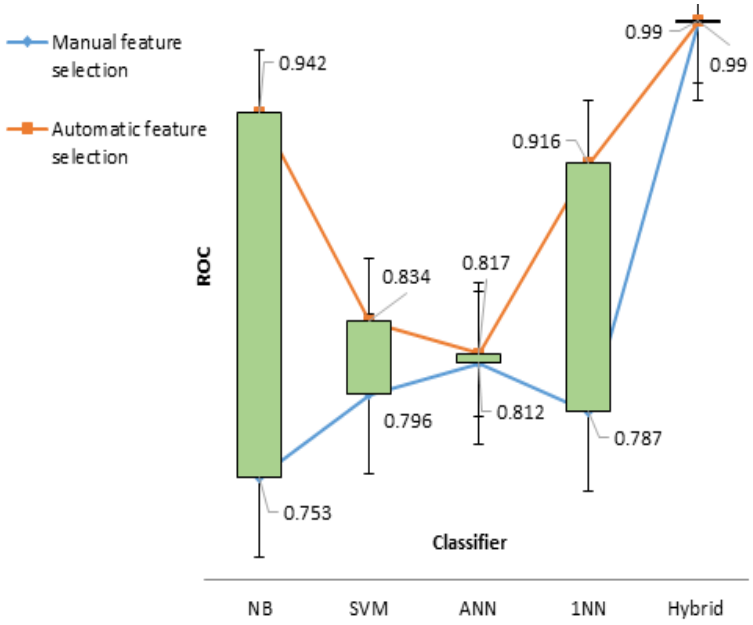

Figure 7. ROC comparison of hybrid modeling with other experiments.

The developed model produced better accuracy than other conventional models, but the present study has some limitations. First, the limited number of subjects investigated could hamper the final dementia subject prediction to the overall AD subjects; second, the outcome of the integration of three experiments may have influenced the results. The use of external MRI information does not guarantee data quality and can affect the significance of the study as a whole.

Brain studies corroborated with artificial intelligence analysis may offer relatively faster investigation methods to modern neurological research. However, it would be preferable to avoid data limitations and, therefore, to enlarge as much as possible the size of the sample investigated in future studies. At the same time, it is also recommended to apply hybrid modeling to younger subjects or subjects with mild AD and to anticipate prediction accuracy with other biological tests like cerebrospinal fluid (CSF) or blood markers.

5. Conclusions

Adult-onset dementia disorders are serious brain pathologies caused by the loss of neuron functions and to progressive atrophy. AD is the most common of these pathologies. It affects primarily elderly people and has a tremendous impact on the lives of people suffering from it. In view of the long time passing between brain lesions bringing about dementia and the onset of clinical symptomatology, early identification of the preclinical and prodromal forms of the disease represents a challenge for medicine. This will reduce medical costs and could contribute to undertaking therapeutic approaches for delaying the conversion of the disease into overt dementia.

Unfortunately, the identification of AD at very early stages is extremely difficult, and there are no tools for its simple detection. We have developed different ML models to predict dementia in the elderly based on MRI findings. The hybrid model with selective features was found to enhance the accuracy of dementia prediction. Experiments with manual feature selection prior to automatic feature selection with 1NN produced 91.32% of accuracy, and the experiment of automatic feature selection generated 96.12% accuracy by SVM. This value significantly increased using multi modeling and produced 98% accuracy. The predictive models developed in this study forecast early AD diagnosis and

the associated risk of developing dementia. Although it is difficult to develop longitudinal projection models in older adults as compared to the younger population, future research in the field should consider addressing both genetic and nongenetic features of multifactorial hazards.

Author Contributions: G.B. and E.T.: study design, manuscript preparation, experiments, and statistical analysis; N.C.: data analysis, methods, and results; F.A.: final revision and study approval. All authors have read and agreed to the published version of the manuscript.

Funding: This study received no external funding.

Acknowledgments: This study was supported by an institutional grant of Camerino University. Gopi B., Nalini C. bursaries were supported by the University of Camerino.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Harvey, R.J.; Skelton-Robinson, M.; Rossor, M.N. The prevalence and causes of dementia in people under the age of 65 years. *J. Neurol. Neurosurg. Psychiatry* **2003**, *74*, 1206–1209. [CrossRef] [PubMed]
2. Prince, M.; Comas-Herrera, A.; Knapp, M.; Guerchet, M.; Karagiannidou, M. *World Alzheimer Report 2016: Improving Healthcare for People living with Dementia. Coverage, Quality and Costs Now and in the Future*; Alzheimer's Disease Int.: London, UK, 2016; pp. 1–140. Available online: https://www.alz.co.uk/research/world-report-2016 (accessed on 6 July 2020).
3. McMurtray, A.; Clark, D.G.; Christine, D.; Mendez, M.F. Early-onset dementia: Frequency and causes compared to late-onset dementia. *Dement. Geriatr. Cogn. Disord.* **2006**, *21*, 59–64. [CrossRef] [PubMed]
4. Battineni, G.; Chintalapudi, N.; Amenta, F. Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM). *Inform. Med. Unlocked* **2019**, *16*, 100200. [CrossRef]
5. Chen, R.; Herskovits, E.H. Machine-learning techniques for building a diagnostic model for very mild dementia. *Neuroimage* **2010**, *52*, 234–244. [CrossRef]
6. Alam, M.A.U.; Roy, N.; Holmes, S.; Gangopadhyay, A.; Galik, E. Automated Functional and Behavioral Health Assessment of Older Adults with Dementia. In Proceedings of the 2016 IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE 2016), Washington, DC, USA, 27–29 June 2016; pp. 140–149. [CrossRef]
7. Angelucci, F. Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) Patients are Characterized by Increased BDNF Serum Levels. *Curr. Alzheimer Res.* **2009**, *5*, 272–273. [CrossRef]
8. Eckerström, C.; Olsson, E.; Borga, M.; Ekholm, S.; Ribbelin, S.; Rolstad, S.; Starck, G.; Edman, Å.; Wallin, A.; Malmgren, H. Small baseline volume of left hippocampus is associated with subsequent conversion of MCI into dementia: The Göteborg MCI study. *J. Neurol. Sci.* **2008**, *272*, 48–59. [CrossRef]
9. Maroco, J.; Silva, D.; Rodrigues, A.; Guerreiro, M.; Santana, I.; de Mendonça, A. Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Res. Notes* **2011**, *4*, 299. [CrossRef]
10. Facal, D.; Valladares-Rodriguez, S.; Lojo-Seoane, C.; Pereiro, A.X.; Anido-Rifon, L.; Juncos-Rabadán, O. Machine learning approaches to studying the role of cognitive reserve in conversion from mild cognitive impairment to dementia. *Int. J. Geriatr. Psychiatry* **2019**, *34*, 941–949. [CrossRef]
11. Darcy, A.M.; Louie, A.K.; Roberts, L.W. Machine learning and the profession of medicine. *JAMA—J. Am. Med. Assoc.* **2016**, *315*, 551–552. [CrossRef]
12. Garrard, P.; Rentoumi, V.; Gesierich, B.; Miller, B.; Gorno-Tempini, M.L. Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex* **2014**, *55*, 122–129. [CrossRef]
13. Karami, V.; Francesco, A.; Giuseppe, N.; Claudio, D.P.; Roberta, L.; Maria, T.P.; Marina, B.; Claudio, B. P68-F Abnormalities of cortical neural synchronization mechanisms in patients with Alzheimer's diseases dementia: An EEG study. *Clin. Neurophysiol.* **2019**, *130*, e86–e87. [CrossRef]
14. Pellegrini, E.; Lucia, B.; Maria, D.C.V.H.; Francesca, M.C.; Victor, G.-C.; Devasuda, A.; Samuel, D.; Susana, M.-M.; Dominic, J.; Cyril, P.; et al. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: A systematic review. *Alzheimer's Dement. Diagn. Assess. Dis. Monit.* **2018**, *10*, 519–535. [CrossRef] [PubMed]

15. Er, F.; Iscen, P.; Sahin, S.; Çinar, N.; Karsidag, S.; Goularas, D. Distinguishing age-related cognitive decline from dementias: A study based on machine learning algorithms. *J. Clin. Neurosci.* **2017**, *42*, 186–192. [CrossRef] [PubMed]
16. Aditya, C.R.; Pande, M.B.S. Devising an interpretable calibrated scale to quantitatively assess the dementia stage of subjects with alzheimer's disease: A machine learning approach. *Inform. Med. Unlocked* **2017**, *6*, 28–35. [CrossRef]
17. Liu, L.; Zhao, S.; Chen, H.; Wang, A. A New Machine Learning Method for Identifying Alzheimer's Disease. *Simul. Model. Pract. Theory* **2020**, *99*, 102023. [CrossRef]
18. Marcus, D.S.; Fotenos, A.F.; Csernansky, J.G.; Morris, J.C.; Buckner, R.L. Open access series of imaging studies: Longitudinal MRI data in nondemented and demented older adults. *J. Cogn. Neurosci.* **2010**, *22*, 2677–2684. [CrossRef]
19. Battineni, G.; Chintalapudi, N.; Amenta, F. Comparative machine learning approach in dementia patient classification using principal component analysis. In Proceedings of the ICAART 2020—12th International Conference on Agents and Artificial Intelligence, Valletta, Malta, 22–24 February 2020; pp. 780–784. [CrossRef]
20. Iguyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182. [CrossRef]
21. Saeys, Y.; Inza, I.; Larrañaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **2007**, *23*, 2507–2517. [CrossRef]
22. Long, N.; Gianola, D.; Rosa, G.J.M.; Weigel, K.A.; Avendano, S. Machine learning classification procedure for selecting SNPs in genomic selection: Application to early mortality in broilers. *Dev. Biol.* **2008**, *132*, 373–376. [CrossRef]
23. Guerreiro, R.; Bras, J. The age factor in Alzheimer's disease. *Genome Med.* **2015**, *7*, 106. [CrossRef]
24. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
25. Peterson, L.E. K-nearest neighbor. *Scholarpedia* **2009**, *4*, 1883. [CrossRef]
26. Rish, I. An empirical study of the naive Bayes classifier. In Proceedings of the IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, Seattle, WA, USA, 4 August 2001; Volume 3, pp. 41–46. [CrossRef]
27. Campbell, C.; Ying, Y. Learning with Support Vector Machines. *Synth. Lect. Artif. Intell. Mach. Learn.* **2011**, *5*. [CrossRef]
28. Schmidhuber, J. Deep Learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
29. Brookes, M.J.; Prejaas, K.T.; Benjamin, A.E.; Hunt, S.E.R.; Lauren, E.G.; Elizabeth, B.L.; Peter, F.L.; Peter, G.M. A multi-layer network approach to MEG connectivity analysis. *Neuroimage* **2016**, *132*, 425–438. [CrossRef]
30. Gaonkar, B.; Hovda, D.; Martin, N.; Macyszyn, L. Deep learning in the small sample size setting: Cascaded feed forward neural networks for medical image segmentation. In *Medical Imaging 2016: Computer-Aided Diagnosis*; International Society for Optics and Photonics: Bellingham, WA, USA, 2016. [CrossRef]
31. Weinberger, K.Q.; Saul, L.K. Distance metric learning for large margin nearest neighbor classification. *J. Mach. Learn. Res.* **2009**, *10*, 207–244. [CrossRef]
32. Battineni, G.; Sagaro, G.G.; Nalini, C.; Amenta, F.; Tayebati, S.K. Comparative machine-learning approach: A follow-up study on type 2 diabetes predictions by cross-validation methods. *Machines* **2019**, *7*, 74. [CrossRef]
33. Lee, L.H.; Wan, C.H.; Rajkumar, R.; Isa, D. An enhanced Support Vector Machine classification framework by using Euclidean distance function for text document categorization. *Appl. Intell.* **2012**, *37*, 80–99. [CrossRef]
34. Zhou, J.; Shi, J.; Li, G. Fine tuning support vector machines for short-term wind speed forecasting. *Energy Convers. Manag.* **2011**, *54*, 1990–1998. [CrossRef]
35. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. *Stat. Surv.* **2010**, *4*, 40–79. [CrossRef]
36. Rao, C.R.; Wu, Y. Linear model selection by cross-validation. *J. Stat. Plan. Inference* **2005**, *128*, 231–240. [CrossRef]
37. Chandrashekar, G.; Sahin, F. A survey on feature selection methods. *Comput. Electr. Eng.* **2014**, *40*, 16–28. [CrossRef]
38. Goldberg, D.E.; Holland, J.H. Genetic Algorithms and Machine Learning. *Mach. Learn.* **1988**, *3*, 95–99. [CrossRef]
39. Bauer, E.; Kohavi, R. An Empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Mach. Learn.* **1999**, *36*, 105–139. [CrossRef]

40. Pencina, M.J.; D'Agostino, R.B.; D'Agostino, R.B.; Vasan, R.S. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat. Med.* **2008**, *27*, 157–172. [CrossRef]
41. Battineni, G.; Sagaro, G.G.; Chinatalapudi, N.; Amenta, F. Applications of machine learning predictive models in the chronic disease diagnosis. *J. Pers. Med.* **2020**, *10*, 21. [CrossRef] [PubMed]
42. Gopi, B.; Nalini, C.; Francesco, A. Late-Life Alzheimer's Disease (AD) Detection Using Pruned Decision Trees. *Int. J. Brain Disord. Treat.* **2020**, *6*, 033. [CrossRef]
43. Huang, J.; Ling, C.X. Using AUC and accuracy in evaluating learning algorithms. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 299–310. [CrossRef]

Deep Learning Type Convolution Neural Network Architecture for Multiclass Classification of Alzheimer's Disease

Gopi Battineni[a], Nalini Chintalapudi[b], Francesco Amenta[c] and Enea Traini
Telemedicine and Tele Pharmacy Centre, School of Medicinal and Health Products Sciences, University of Camerino, Camerino, 62032, Italy

Keywords: Alzheimer's Disease (AD), OASIS-3, MRI Images, Deep Learning, CNN.

Abstract: Alzheimer's disease (AD) is one of the common medical issues that the world is facing today. This disease has a high prevalence of memory loss and cognitive decline primarily in the elderly. At present, there is no specific treatment for this disease, but it is thought that identification of it at an early stage can help to manage it in a better way. Several studies used machine learning (ML) approaches for AD diagnosis and classification. In this study, we considered the Open Access Series of Imaging Studies-3 (OASIS-3) dataset with 2,168 Magnetic Resonance Imaging (MRI) images of patients with very mild to different stages of cognitive decline. We applied deep learning-based convolution neural networks (CNN) which are well-known approaches for diagnosis-based studies. The model training was done by 70% of images and applied 10-fold cross-validation to validate the model. The developed architecture model has successfully classified the different stages of dementia images and achieved 83.3% accuracy which is higher than other traditional classification techniques like support vectors and logistic regression.

1 INTRODUCTION

Alzheimer's Disease (AD) is the most well-known and largely diffused neurodegenerative disorder occurring in the elderly. AD negatively affects patients' everyday lives, causing an advanced decline of cognitive capabilities such as memory, language, behaviour, and critical thinking (Alzheimer's Disease International (ADI) 2010). Changes in cognitive impairment of AD patients start slowly and evolve rapidly over the long run.

Similar to other body parts, brain can change as people get older. Some people lost thinking and incidental issues with recollecting certain things. Excessive cognitive decline, and other significant changes in the manner in which brain function is impaired (Jaussent et al. 2012). The first symptoms of AD are trouble recalling recently learned data because Alzheimer's progressions regularly start in the brain areas involved in learning and memory. As Alzheimer's progresses progressively severe symptoms like confusion, mood changes, disorientation, unwarranted doubts about family and companions, and trouble talking appear. Individuals with cognitive decline or other potential indications [1]of AD may think that it's difficult to remember they have an issue.

AD is a type of dementia with several implications on the cognitive domain, affecting primarily thinking and memory. Specialists and different parental figures screen the movement of AD in patients by assessing the level of decrease in the patients' psychological capacities that are often classified into three stages: very mild (normal cognitive), mild cognitive impairment (MCI), and demented (Gaugler et al. 2016). Figure 1 presents the magnetic resonance image (MRI) images of different AD conditions. Although the MCI and dementia patients both are experiencing a reduction of cognitive abilities, dementia patients would suffer from more pronounced difficulties with thinking or hampered judgment.

[a] https://orcid.org/0000-0003-0603-2356
[b] https://orcid.org/0000-0003-0818-306X
[c] https://orcid.org/0000-0002-0555-1034

Battineni, G., Chintalapudi, N., Amenta, F. and Traini, E.
Deep Learning Type Convolution Neural Network Architecture for Multiclass Classification of Alzheimer's Disease.
DOI: 10.5220/0010378602090215
In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2021) - Volume 2: BIOIMAGING*, pages 209-215
ISBN: 978-989-758-490-9

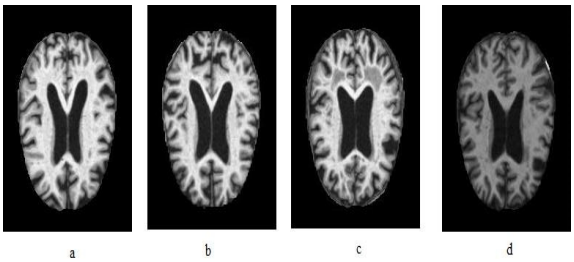

Figure 1: AD presented by MRI images (a) mild dementia; (b) moderate demented; (c) nondemented; and (d) very mild demented.

In clinical practice, the capacity to accurately forecast the patient diagnosis can help by adding appropriate medical decisions on treatment approaches. Recently, machine learning (ML) algorithms are largely applying to forecast and predict diseases and helping in quick decision making (Battineni, Sagaro, et al., 2020). Pattern-related approaches like logistic regression (Johnson et al., 2014), support vector machines (Battineni, Chintalapudi, en Amenta 2019), and linear discriminant analysis (Rathore et al. 2017) are giving promising results in the prediction of AD development and early AD detection.

Deep learning models were used unlabeled data during preprocessing. These are well suited for imbalanced datasets and achieve a knowledge base (Mittal et al. 2019). At present these are largely involved in all other problems that are not able to be addressed by traditional artificial intelligence (AI) techniques. Neural networks are the latest deep learning algorithms that have discovered the functionality of different situations. Convolutional neural networks (CNN) are characterized contributions to profits through a complex composition of layers that presents building blocks including nonlinear functions and transformations.

Medical experts feel that deep learning could be a promising solution in AD identification and stage detection (Khan et al., 2020). For instance, (Basheera en Sai Ram, 2019) applied CNN modeling for AD diagnosis based on T2 weighted magnetic resonance imaging (MRI) and achieved 90.47% accuracy. A Siamese CNN can also help to categorize the AD and studies reported 99.05% of accuracy (Mehmood et al. 2020). It is also reported that AD prediction from MCI using the CNN model reported 79.9% of accuracy(Lin et al., 2018). Therefore, it is assumed that an effective and comprehensive deep learning model can help to identify early AD prediction and ultimately provide timely treatment to the suffered patients. In this work, we proposed convolutional neural networks (CNN) model of deep learning type for detection of early-stage AD and successfully classify the MRI images on four different dementia stages presented in Figure 2.

Experiments were conducted on longitudinal neuroimages of the OASIS-3 database that include MR scans of T1-weighted, T2 weighted, ASL, SWI, DTI sequences, FLAIR, time of flight, and resting-state BOLD. The rest of the paper is structured according to the following outline: Section 2 presents the dataset and proposed model architecture; section 3 presents the experimental results, and section 4 makes a discussion which is followed by the conclusion in section 5.

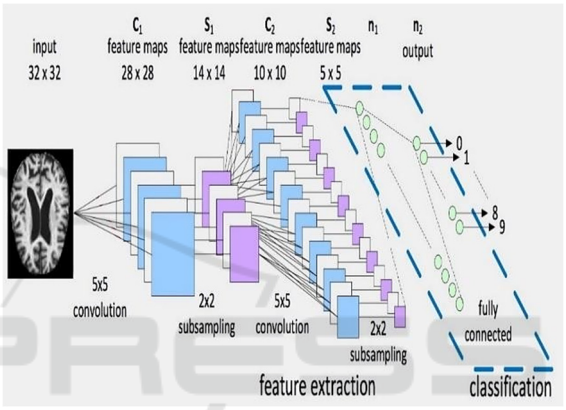

Figure 2: Brain image classification with the CNN model framework.

2 METHODS

2.1 Dataset

The Open Access Series of Imaging Studies (OASIS) contains MR scanning information that is openly accessible to scientific communities. They released OASIS-1 (cross-sectional) and OASIS-2 (longitudinal) MRI datasets among different subjects and these datasets are widely used in many studies (Sweeney et al. 2013; Palumbo et al. 2019). OASIS-3 is the extension of previous datasets. It includes 1,098 patients aging from 42 to 95 years. Among participants, 609 are associated with normal cognitive decline (very mild), and 489 were associated with different cognitive decline stages. OASIS-3 dataset incorporated both functional and structural features of more than 2,000 MRI images. The dataset outcome of four categories of MR images has presented in Figure 3.

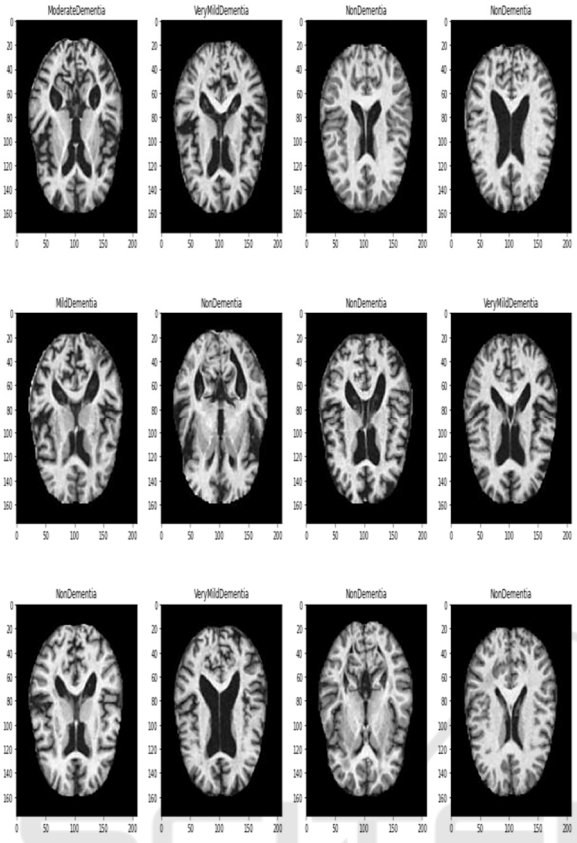

Figure 3: Dataset outcome of different dementia stages (3*4 image matrix).

2.2 CNN Model Architecture

A convolutional neural network (ConvNet) is deep learning type algorithms that take images as input, assign features based on their importance (biases and learnable weights) to different image objects, and also be able to separate one from the other (Krizhevsky, Sutskever, en Hinton 2017). When compared with other classification models, ConvNet possesses low complex pre-processing steps. In CNN, each input image is gone through sequence convolution layers namely pooling layers, filtering layers (kernels), and fully connected layers (FCs).

To make the proposed model easier for understanding, we created a dense layer block and convolution block. The architecture of the CNN model is inspired by the article (Pan et al. 2020). We built the CNN model by using five convolutional slabs covered with convolution layers, feature engineering, max pooling, and classification. We have used cross-entropy as a loss function and Adam as an optimizer. SoftMax has been used to classify the multiclass AD stages since it is associated with a mutually exclusive relationship. The feature representation (f_k) works as an input to the SoftMax layer and interprets output brain stages. Aprobability score P (k) for each class as defined as

$P_k=\frac{\exp(fk)}{\sum_{k=1}^{k} exp\ (fk)}$;where fi feature representation, and Cross entropy loss function as

(L)= $\sum_{k=1}^{k} tk.\log(pk)$; where t_k ground truth of MRimage then $\frac{\partial L}{\partial fk}=P_k-t_k$.

2.3 Experimental Setup

Figure 4 presents the most relevant procedures followed to construct the feature data of brain images and extraction of AD images developed in thispaper. After pre-processing steps, the given image dataset has been divided into training and validation files with standard (80:20) division.

The procedures indicated red line are MR images that fed to the CNN model for training purposes. The model extracts the input image features of trained images under present parameters and supplies them to the SoftMax classifier for testing. The SoftMax function calculates the loss and model accuracy. For avoiding high loss, network parameters are adjusted by the back-propagation algorithm. After applying several iterations (epochs) the better-trained parameters have been achieved. The model visualization metrics like loss and receiver operating characteristic area under the curve (ROC AUC) have been taken as the performance parameter for AD classification since it has been considered one of the key metrics in multi-image classification techniques. The experimental setup and AD detection and classification have been done through TensorFlow and python language.

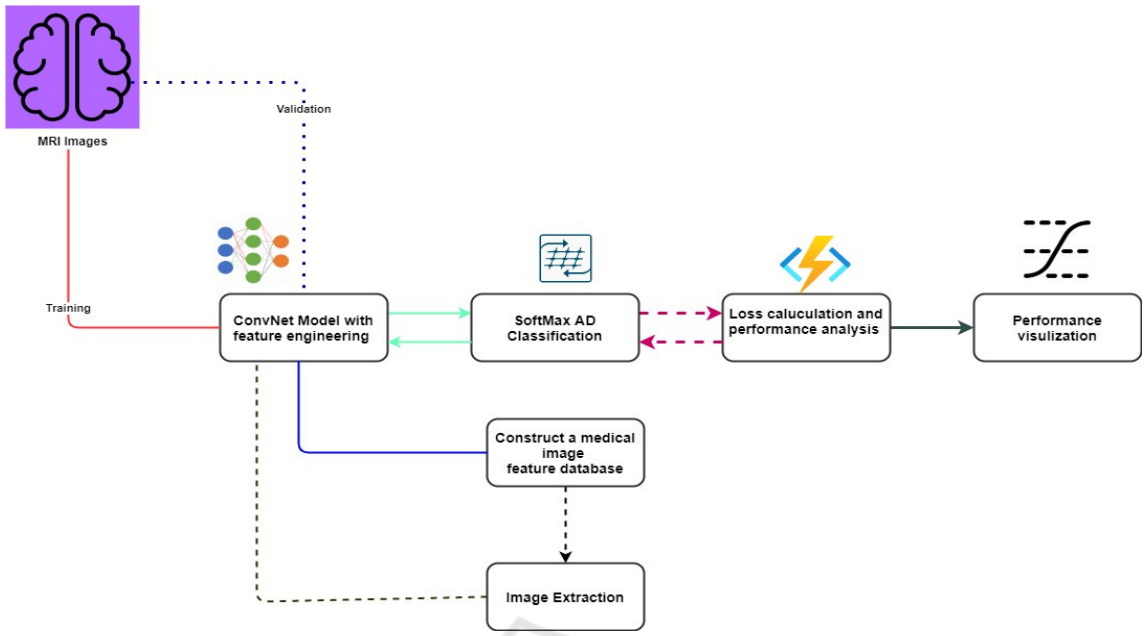

Figure 4: Experimental setup of the work.

3 RESULTS

To do efficient training on our CNN model, a back-propagation algorithm is set to adjust the rate of learning and stop the model automatically once it reaches maximum accuracy. Since the learning rate is one of the hyperparameters that decides model accuracy and time to process the model. OASIS-3 dataset consisted of 2168 independent MRI scanners. Among the given images, 1,734 are used for training and 434 were used for validation purposes. Because of the large image dataset, 10-fold cross-validation has been used and we have used each fold 70% as training, 10% as validation, and 20% images are used testing. The distribution of the dataset is presented in Table 1.

Table 1: Total image distribution.

Total Images: 2168	
Type	***Percentage***
Trained images	1517 (70%)
Testing images	434 (20%)
Validation images	217 (10%)

The model-fitting has to be done on a sample of 100 epochs and to prevent model overfitting we stop the model early at the 80th iteration. The model took a run time of 138 min to process the trained images. Figure 5 presents a graphical representation of ROC. AUC and loss metrics after each iteration on both training and validation image data.

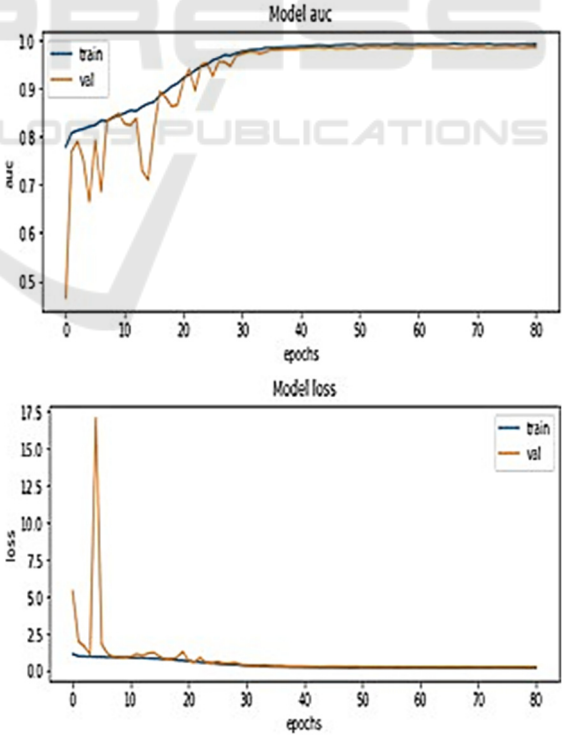

Figure 5: Model AUC and loss metric outcomes.

Though the model evaluation has been done on the validation dataset, we also perform the

experiments on the testing dataset. The testing dataset model AUC curve outcome has presented in Figure 6 and the model achieved a ROC of 83.3% which is considered as an optimal classifier for AD image detection and this value is significantly higher than traditional ML approaches (Battineni, Chintalapudi, en Amenta 2019; A. Khan en Zubair 2020).

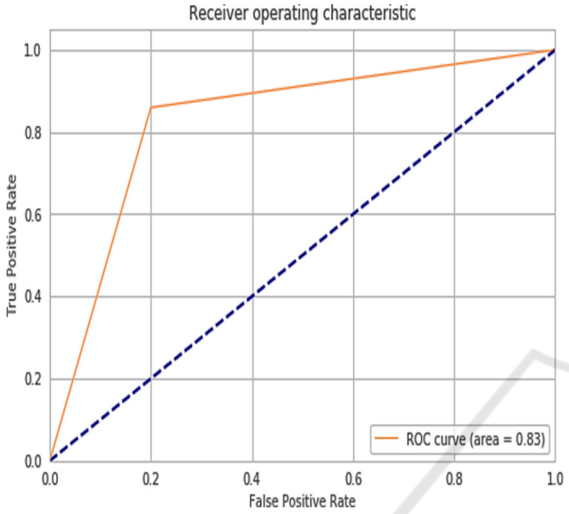

Figure 6: The ROC curve of test data.

4 DISCUSSION

In this work, we presented a novel deep learning type CNN model for the classification of AD subjects. As mentioned above, AD is the most common adult-onset dementia and contributes about 60-70% of worldwide dementia cases (A. Khan en Zubair 2020). Unfortunately, there is no proper medication or cure for AD, and advancements in AD cure have been getting slow. Screening among people of AD risk given electronic health records (EHR) in preclinical stages may prompt early identification of AD pathology and suggest better approaches for complying with the AD beginning. Current biomarkers of AD have required specimen collection (like serum or liquid), MRI image data, or more sophisticated markers that at the present can be identified just in highly specialized centres (Mantzavinos en Alexiou 2017; Hadjichrysanthou et al. 2020).

On the other hand, the EHRs for example medical records in clinical settings, or administrative health information don't require extra time or effort for data collection. Likewise, with the coming of digitalization, the measures of such information have drastically increased (Shao et al. 2019). Since it is omnipresent, enormous, and cost-effective, the digitized medical database might be a significant asset for testing different AD predictive models. Nonetheless, despite its enormous possible value, somehow thought about the degrees to which the enormous scope of EHR data can help in risk of AD prediction (Shao et al. 2019; Mayer et al. 2015). The possible prediction of future AD progression is incredibly significant in clinical practice also, in healthcare research. Advanced neuroimaging techniques like MRI, positron emission tomography (PET) is developed and presented to identify AD-related molecular and structural biomarkers (Hadjichrysanthou et al. 2020).

Computer scientists are recommending applying sophisticated computing techniques like machine learning and deep learning. It is reported that 99.1% of accuracy has been achieved through the application of ensemble learning models for late-life AD detection among 150 patients (Battineni, Chintalapudi, et al. 2020). AD prediction among 123 subjects with Pre-MCI and MCI was done by clinically transmittable ML algorithms and results reported the whole sample accuracy of 96.2% (Grassi et al. 2018). However, most of the outcomes proposed by these algorithms are based on demographic magnetic resonance image (MRI) information. Because of this, researchers believed that deep learning algorithms are the best approaches if brain images were included (Choi en Jin 2018). Most of the works associated with Machine learning in the early prediction of AD occurred with high success. For instance, it is reported that 94.1% of accuracy by 3D convolutional neural networks (CNN) (Esmaeilzadeh et al. 2018).

This work presented a deep CNN with 10-fold cross-validation and achieved more than 80% accuracy. While applying computing methods for diagnosis, a small portion of datasets are presented. Therefore, our model maintained a random image selection of train, test, and validation datasets. The proposed model produced promising results in AD image classification. The most notable outcome for this study is the progressions among predictiveness of AD diseases.

5 CONCLUSIONS

An autonomous AD detection classifier based deep ConvNet framework is presented. We adopted the latest release of the OASIS-3 dataset that contains

different categories of AD datasets. For training, more than 1,500 images model took a bit longer process than expected, but it is faster than mankind process. Deep ConvNets do not need any handcrafted feature selection approach because of having autonomous feature tuning. The main limitation of the study is to adopt only a single classifier for the brain MRI data classification and there are other possibilities to do better improvements in the proposed model architecture. Although attained results of higher 80% accuracy while compared over traditional ML classifiers, many advancements are proposed to enhance the model quality.

CONFLICTS OF INTEREST

No author has produced any conflicts of interest.

REFERENCES

Alzheimer's Disease International (ADI). 2010. "World Alzheimer Report 2010: The Global Economic Impact of Dementia". *Alzheimer's Disease International (ADI)*. https://doi.org/10.1111/j.0963-7214.2004.00293.x.

Basheera, Shaik, en M. Satya Sai Ram. 2019. "Convolution neural network–based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation". *Alzheimer's and Dementia: Translational Research and Clinical Interventions*. https://doi.org/10.1016/j.trci.2019.10.001.

Battineni, Gopi, Nalini Chintalapudi, en Francesco Amenta. 2019. "Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM)". *Informatics in Medicine Unlocked*. https://doi.org/10.1016/j.imu.2019.100200.

Battineni, Gopi, Nalini Chintalapudi, Francesco Amenta, en Enea Traini. 2020. "A Comprehensive Machine-Learning Model Applied to Magnetic Resonance Imaging (MRI) to Predict Alzheimer's Disease (AD) in Older Subjects". *Journal of Clinical Medicine*. https://doi.org/10.3390/jcm9072146.

Battineni, Gopi, Getu Gamo Sagaro, Nalini Chinatalapudi, en Francesco Amenta. 2020. "Applications of machine learning predictive models in the chronic disease diagnosis". *Journal of Personalized Medicine*. https://doi.org/10.3390/jpm10020021.

Choi, Hongyoon, en Kyong Hwan Jin. 2018. "Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging". *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2018.02.017.

Esmaeilzadeh, Soheil, Dimitrios Ioannis Belivanis, Kilian M. Pohl, en Ehsan Adeli. 2018. "End-to-end alzheimer's disease diagnosis and biomarker identification". In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030- 00919-9_39.

Gaugler, Joseph, Bryan James, Tricia Johnson, Ken Scholz, en Jennifer Weuve. 2016. "2016 Alzheimer's disease facts and figures". *Alzheimer's and Dementia*. https://doi.org/10.1016/j.jalz.2016.03.001.

Grassi, Massimiliano, Giampaolo Perna, Daniela Caldirola, Koen Schruers, Ranjan Duara, en David A. Loewenstein. 2018. "A clinically-translatable machine learning algorithm for the prediction of Alzheimer's disease conversion in individuals with mild and premild cognitive impairment". *Journal of Alzheimer's Disease*. https://doi.org/10.3233/JAD- 170547.

Hadjichrysanthou, Christoforos, Stephanie Evans, Sumali Bajaj, Loizos C. Siakallis, Kevin McRae-Mckee, en Frank De Wolf. 2020. "The dynamics of biomarkers across the clinical spectrum of Alzheimer's disease". *Alzheimer's Research and Therapy*. https://doi.org/10.1186/s13195-020-00636-z.

Jaussent, Isabelle, Jean Bouyer, Marie Laure Ancelin, Claudine Berr, Alexandra Foubert-Samier, Karen Ritchie, Maurice M. Ohayon, Alain Besset, en Yves Dauvilliers. 2012. "Excessive sleepiness is predictive of cognitive decline in the elderly". *Sleep*. https://doi.org/10.5665/sleep.2070.

Johnson, Piers, Luke Vandewater, William Wilson, Paul Maruff, Greg Savage, Petra Graham, Lance S. Macaulay, et al. 2014. "Genetic algorithm with logistic regression for prediction of progression to Alzheimer's disease". *BMC Bioinformatics*. https://doi.org/10.1186/1471-2105-15-S16-S11.

Khan, Afreen, en Swaleha Zubair. 2020. "An Improved Multi-Modal based Machine Learning Approach for the Prognosis of Alzheimer's disease". *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2020.04.004.

Khan, Mehshan Ahmed, Muhammad Attique Khan, Fawad Ahmed, Mamta Mittal, Lalit Mohan Goyal, D. Jude Hemanth, en Suresh Chandra Satapathy. 2020. "Gastrointestinal diseases segmentation and classification based on duo-deep architectures". *Pattern Recognition Letters*. https://doi.org/10.1016/j.patrec.2019.12.024.

Krizhevsky, Alex, Ilya Sutskever, en Geoffrey E. Hinton. 2017. "ImageNet classification with deep convolutional neural networks". *Communications of the ACM*. https://doi.org/10.1145/3065386.

Lin, Weiming, Tong Tong, Qinquan Gao, Di Guo, Xiaofeng Du, Yonggui Yang, Gang Guo, Min Xiao, Min Du, en Xiaobo Qu. 2018. "Convolutional neural networks-based MRI image analysis for the Alzheimer's disease prediction from mild cognitive impairment". *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2018.00777.

Mantzavinos, Vasileios, en Athanasios Alexiou. 2017. "Biomarkers for Alzheimer's Disease Diagnosis".

Current Alzheimer Research. https://doi.org/10.2174/1567205014666170203125942.

Mayer, Miguel A., Laura I. Furlong, Pilar Torre, Ignasi Planas, Francesc Cots, Elisabet Izquierdo, Jordi Portabella, Javier Rovira, Alba Gutierrez-Sacristan, en Ferran Sanz. 2015. "Reuse of EHRs to Support Clinical Research in a Hospital of Reference". In *Studies in Health Technology and Informatics*. https://doi.org/10.3233/978-1-61499-512-8-224.

Mehmood, Atif, Muazzam Maqsood, Muzaffar Bashir, en Yang Shuyuan. 2020. "A deep siamese convolution neural network for multi-class classification of alzheimer disease". *Brain Sciences*. https://doi.org/10.3390/brainsci10020084.

Mittal, Mamta, Lalit Mohan Goyal, Sumit Kaur, Iqbaldeep Kaur, Amit Verma, en D. Jude Hemanth. 2019. "Deep learning based enhanced tumor segmentation approach for MR brain images". *Applied Soft Computing Journal*. https://doi.org/10.1016/j.asoc.2019.02.036.

Palumbo, L., P. Bosco, M. E. Fantacci, E. Ferrari, P. Oliva, G. Spera, en A. Retico. 2019. "Evaluation of the intra- and inter-method agreement of brain MRI segmentation software packages: A comparison between SPM12 and FreeSurfer v6.0". *Physica Medica*. https://doi.org/10.1016/j.ejmp.2019.07.016.

Pan, Dan, An Zeng, Longfei Jia, Yin Huang, Tory Frizzell, en Xiaowei Song. 2020. "Early Detection of Alzheimer's Disease Using Magnetic Resonance Imaging: A Novel Approach Combining Convolutional Neural Networks and Ensemble Learning". *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2020.00259.

Rathore, Saima, Mohamad Habes, Muhammad Aksam Iftikhar, Amanda Shacklett, en Christos Davatzikos. 2017. "A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages". *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2017.03.057.

Shao, Yijun, Qing T. Zeng, Kathryn K. Chen, Andrew Shutes-David, Stephen M. Thielke, en Debby W. Tsuang. 2019. "Detection of probable dementia cases in undiagnosed patients using structured and unstructured electronic health records". *BMC Medical Informatics and Decision Making*. https://doi.org/10.1186/s12911-019-0846-4.

Sweeney, Elizabeth M., Russell T. Shinohara, Navid Shiee, Farrah J. Mateen, Avni A. Chudgar, Jennifer L. Cuzzocreo, Peter A. Calabresi, Dzung L. Pham, Daniel S. Reich, en Ciprian M. Crainiceanu. 2013. "OASIS is Automated Statistical Inference for Segmentation, with applications to multiple sclerosis lesion segmentation in MRI". *NeuroImage: Clinical*. https://doi.org/10.1016/j.nicl.2013.03.002.

diagnostics

MDPI

Article

Improved Alzheimer's Disease Detection by MRI Using Multimodal Machine Learning Algorithms

Gopi Battineni [1,*], Mohmmad Amran Hossain [1], Nalini Chintalapudi [1], Enea Traini [1], Venkata Rao Dhulipalla [2], Mariappan Ramasamy [2] and Francesco Amenta [1]

[1] Telemedicine and Telepharmacy Centre, School of Medicinal and Health Products Sciences, University of Camerino, 62032 Camerino, Italy; mohammadamra.hossain@studenti.unicam.it (M.A.H.); nalini.chintalapudi@unicam.it (N.C.); enea.traini@unicam.it (E.T.); francesco.amenta@unicam.it (F.A.)

[2] The Research Centre of the ECE Department, V.R. Siddhartha Engineering College, Vijayawada 521002, Andhra Pradesh, India; hodece@vrsiddhartha.ac.in (V.R.D.); mariappan@vrsiddhartha.ac.in (M.R.)

* Correspondence: gopi.battineni@unicam.it; Tel.: +39-3331728206

Abstract: Adult-onset dementia disorders represent a challenge for modern medicine. Alzheimer's disease (AD) represents the most diffused form of adult-onset dementias. For half a century, the diagnosis of AD was based on clinical and exclusion criteria, with an accuracy of 85%, which did not allow for a definitive diagnosis, which could only be confirmed by post-mortem evaluation. Machine learning research applied to Magnetic Resonance Imaging (MRI) techniques can contribute to a faster diagnosis of AD and may contribute to predicting the evolution of the disease. It was also possible to predict individual dementia of older adults with AD screening data and ML classifiers. To predict the AD subject status, the MRI demographic information and pre-existing conditions of the patient can help to enhance the classifier performance. In this work, we proposed a framework based on supervised learning classifiers in the dementia subject categorization as either AD or non-AD based on longitudinal brain MRI features. Six different supervised classifiers are incorporated for the classification of AD subjects and results mentioned that the gradient boosting algorithm outperforms other models with 97.58% of accuracy.

Keywords: Dementia; Alzheimer's disease; machine learning; prediction; performance; AUROC

Citation: Battineni, G.; Hossain, M.A.; Chintalapudi, N.; Traini, E.; Dhulipalla, V.R.; Ramasamy, M.; Amenta, F. Improved Alzheimer's Disease Detection by MRI Using Multimodal Machine Learning Algorithms. *Diagnostics* **2021**, *11*, 2103. https://doi.org/10.3390/diagnostics11112103

Academic Editor: Semen A. Kurkin

Received: 1 September 2021
Accepted: 11 November 2021
Published: 13 November 2021

1. Introduction

Alzheimer's disease (AD) is an adult-onset cognitive disorder (AOCD) which represents the sixth leading cause of mortality and the third most common disease after cardiovascular diseases and cancer [1]. AD is mainly characterized by nerve cell widespread loss, neuro-fibrillary tangles, and senile plaques occurring primarily in the hippocampus, entorhinal cortex, neocortex, and other brain regions [2]. It is hypothesized that there are 44.4 million people experiencing dementia in the world and this number will probably increase to 75.6 million in 2030 and 135.5 million in 2050 [3]. For half a century, the diagnosis of AOCD was based on clinical and exclusion criteria (neuropsychological tests, laboratory, neurological assessments, and imaging findings). The clinical criteria have an accuracy of 85% and do not allow a definitive diagnosis, which could only be confirmed by post-mortem evaluation. Clinical diagnosis has been associated with time with instrumental examinations, such as analysis of the liquoral levels of specific proteins and demonstration of cerebral atrophy with neuroimaging [4]. Further evolution of neuroimaging techniques is associated with quantitative assessment.

Various neuroimaging approaches, such as the AD neuroimaging initiative (ADNI) [4], were developed to identify early stages of dementia. The early diagnosis and possible prediction of AD progression are relevant in clinical practice. Advanced neuroimaging techniques, such as magnetic resonance imaging (MRI), have been developed and presented

to identify AD-related molecular and structural biomarkers [5]. Clinical studies have shown that neuroimaging modalities such as MRI can improve diagnostic accuracy [6]. In particular, MRI can detect brain morphology abnormalities associated with mild cognitive impairment (MCI) and has been proposed to predict the shift of MCI into AD accurately at an early stage.

A further suggested approach is the analysis of the so-called multimodal biomarkers that can play a relevant role in the early diagnosis of AD. Studies of Gaubert and coworkers trained the machine learning (ML) classifier using features such as EEG, APOE4 genotype, demographic, neuropsychological, and MRI data of 304 subjects [7]. The model is trained to predict amyloid, neurodegeneration, and prodromal AD. It has been reported that EEG can predict neurodegenerative disorders and demographic and MRI data are able to predict amyloid deposition and prodromal at five years, respectively. In line with the above investigations, ML techniques were considered useful to predict AD. This helps in quick decision making [8]. Different supervised ML models were developed and tested their performance in AD classification [9]. However, it is said that boosting models [10] such as the generalized boosting model (GLM Boost) and gradient boosting machines (GBM) outperform other models in terms of classification accuracy and specificity.

Dementia can also be predicted via integrating ML knowledge with the patient's clinical history. A gradient boosting model (light GBM) to predict the onset of dementia using two years AD patient records was proposed as well [11]. This obtained 87% of accuracy. Another approach using Recurrent Neural Networks (RNN) was presented for the AD progression modeling [12]. This network was compared with another existing RNN modeling with data assertion and regression method. This resulted in a 74% of accuracy even with unlabeled data. At the same time, MRI demographic data can also help to predict AD by learning the intradata relationships. It has been reported that with this approach random forest (RF) models outperform other classification algorithms such as SVM [13]. In particular, deep learning models produced promising results in predicting the shift of MCI into overt AD and in early AD detection [14]. Deep learning models used unlabeled data during pre-processing and are well suited for imbalanced datasets and achieving a knowledge base. It has been suggested that deep learning could be a promising solution in AD identification and symptom detection [15]. An effective and comprehensive deep learning model can help to an early AD prediction, and consequently, to provide timely treatment to the suffering patients.

Discretization of MRI data efficiently handles the outliers and thereby improves the accuracy of ML classifiers. It is reported that the successful classification of dementia subjects can be done by supervised models associated with feature selection [16]. In another study, patient classification was accomplished via multifactor affiliation analysis with the inter feature relationships [17]. This technique helps in getting better patient classification and produce higher performance compared with classification trees and generic-distribution zones [17]. The above approaches did not highlight the importance of data-centric ML techniques and the adoption of model boosting knowledge, which can transform weak learners into strong learners and improve model performance.

In this study, we have applied the datacentric ML classification techniques by involving both supervised and boosting models and comparing performance in the detection of the best model. To achieve this, we proposed an ML framework for the classification of AD and non-AD patients, and the classifier performance was assessed and validated with cross-validation techniques. This work has developed the presentation and comparison of the classification models efficiently on smaller datasets. The main purpose of this investigation was to present the list of classification accuracies along with other performance metrics, such as precision and recall. The most notable outcome for this research study is the analysis of the progression among prediction and classification of AD detection.

2. Methods

2.1. Subjects

The dataset was retrieved from the Open Access Series of Imaging Studies (OASIS) of neurology. Patients in the age group between 60 and 96 years of age were chosen from a bigger dataset of people who had taken an interest in MRI studies at Washington University. The dataset is based on the accessibility of something like two separate visits in which clinical and MRI data were recorded, three or more gained T1-weighted images per imaging session and right-hand strength. The patient database was acquired from the longitudinal pool of Alzheimer Disease Research Centre (ADRC) at Washington University [18]. The controlled group and psychologically disabled patients' group were enlisted in the ADRC, especially through media offers, among which 80% of people by direct contact with the center and the rest of people by doctors' referral. All the patients took part as per the rules of the Human Studies Committee, Washington University. Endorsement for public sharing of the anonymized data was also explicitly obtained. The subject demographic information is shown in Table 1.

Table 1. Demographic characteristics of the subjects investigated (D: demented; ND: nondemented; SD: standard deviation).

Subjects	78 D	72 ND
Male	40 D	22 ND
Female	38 D	50 ND
Age range (years)	60–96	
Median	77.0	
Mean ± SD	77.01 ± 7.3	

2.2. Clinical Assessment

Dementia status was assessed by the Clinical Dementia Rating (CDR) scale. The classification of dementia or non-dementia control groups was based on clinical criteria, without reference to psychometric execution, and any likely reasons for dementia (known neurological, clinical, or mental issues), which would not lead to dementia. The diagnosis of AD was made based on clinical data (obtained basically from an insurance source). The subjects experienced a slow, gradual decrease in memory and other psychological and functional impairments. In particular, the CDR is a dementia scale, which rates patients for the level of impedance in every one of six areas: memory, orientation, judgment and critical thinking, work in the community, home and hobbies, and individual care. Based on the reliable source and subject meeting, the global CDR score is obtained from singular evaluations in each domain. The global CDR of 0 indicates no dementia and a CDR of 0.5, 1, 2, and 3 indicate extremely mild, mild, moderate, and severe dementia, respectively [19]. The proposed techniques here take into account the clinical finding of AD in people with a CDR of 0.5 or more prominent based on standard criteria based on histopathological assessment in 93% of the people [20]. Those in the earliest or mildest cognitive decline (CDR of 0.5) of AD might be considered as MCI. The diagnostic characteristics of different age groups considered are presented in Table 2.

2.3. Image Acquisition

For each subject, three or four individual T1-weighted magnetizations prepared rapid gradient-echo (MP-RAGE) images were acquired on a 1.5-T Vision scanner (Siemens, Erlangen, Germany) in a single imaging session. Head movement was minimized by padding and utilizing a thermoplastic face mask. Each image presents 14 independent features each corresponding to classify the dependent value of the subject group. The binary classifier subject group defines each individual either as non-demented (0) or demented (1). Table 3 presents the description of each independent feature allowing the classification of the subject group.

Table 2. Age and characteristics of the individuals investigated on the first clinical visit [18].

	Non-Demented						Demented				
Age	**N**	**n**	**Mean**	**Male**	**Female**	**Convert**	**n**	**Mean**	**Male**	**Female**	**CDR (0.5/1)**
60s	34	23	65.71	6	17	3	11	65.67	8	3	8/3
70s	71	35	74.91	11	24	4	36	73.97	20	16	29/7
80s	41	26	84.30	9	17	7	15	82.33	7	8	13/2
90s	4	2	92.50	0	2	0	2	93.00	1	1	1/1
Total	150	86	75.82	26	59	14	64	74.95	36	29	52/13

Table 3. Dataset feature description.

Features	Description
Subject ID	Subject identification number
MRI ID	Image identification number of an individual subject
Visit	Number of subject visits
Gender	Male/Female
Hand	Right/Left-handed
EDUC	Subject education level (in years)
SES	Socioeconomic status
MMSE	Mini-mental state examination score
CDR	Clinical dementia rating score
e-TIV	Estimated total intracranial volume result
n-WBV	Normalized whole brain volume result
ASF	Atlas scaling factor
Age	Subject age while scanning
Group	Demented/Nondemented/Converted
MR delay	Magnetic resonance (MR) delay is the delay time that is before the image procurement

2.4. Experimental Setup

The experimental setup was introduced for the classification of AD patients and included

- ➢ A learning model that can effectively predict and segregate true AD subjects from a given population.
- ➢ The development of a novel ML classifier and validate its performance.

To achieve this, OASIS longitudinal MRI data of 150 subjects were used. The ML model pipeline approach was applied in the diagnosis of AD, to classify true dementia subjects. The proposed ML framework can learn data by the provided classifiers and categorize them as true and non-AD subjects. The Jupiter platform with Python libraries was used for an experimental setup; this platform is well known by developers for processing, assessment, and model building. Python is a high-level programming language with dynamic semantics. Figure 1 shows the proposed method to evaluate a high-performance model in AD patient classification.

2.4.1. Data Pre-Processing

(a) Missing Data Handling

The real-world data contain missing values and noise, also in a raw format that cannot be directly involved in the development of ML models. To convert such noisy data into a machine-understandable format, data pre-processing steps are needed, such as data cleaning and data formatting. The first step in data pre-processing was the handling of missing data. In this, we identified that the SES (1–5) feature had 19 missing values and MMSE (0–30) had 2 missing values. For handling these two features, we replaced missing data points with the values that occurred the most (for SES this was 2 and for MMSE this was 30) [21].

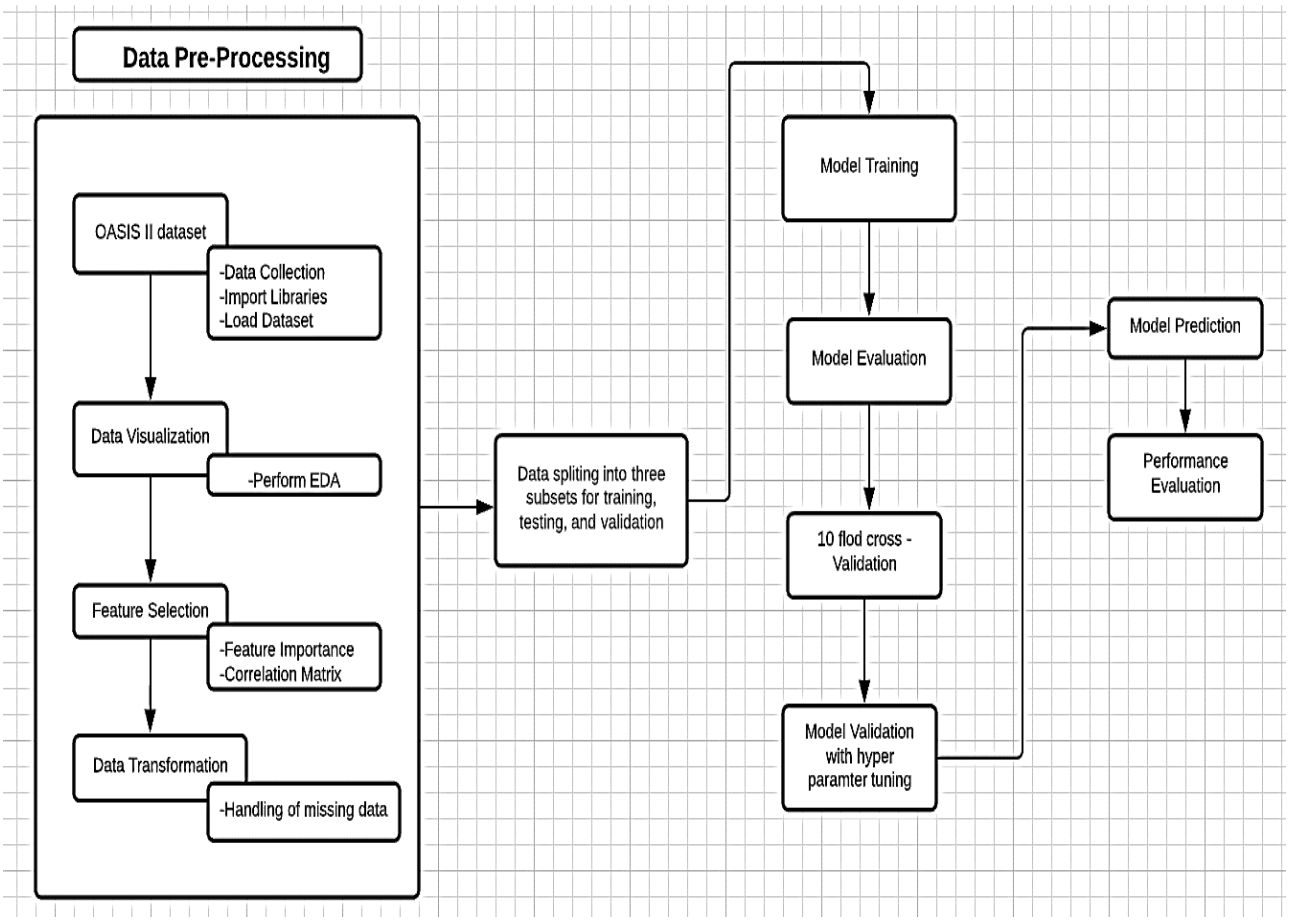

Figure 1. Experimental setup for the proposed model.

(b) Data Visualization

In this step, we perform an exploratory data analysis (EDA) technique that incorporates different methods and tools employed to advance the statistical insight and graphical data representation. Figure 2 represents the value distribution of different MRI features in the prediction of the target AD group value.

The identification of a relationship between different MRI features helps in the detection of highly correlated features with the target group. To do that correlation, a matrix was developed to understand the relationship among given features and targeted outcomes. The features with at least 50% of correlation with the target group are included. The outcome of the correlation matrix heatmap can be visualized in Figure 3. Similarly, an outlier can be a data point that varies significantly from other parameters. The dataset outliers present the quantitative distribution of data in a way that helps in the comparison of given features. The box plots of outliers with 50% of correlation were reported as e-TIV, age, n-WBV, ASF, MR delay, and VISIT are presented in Figure 4.

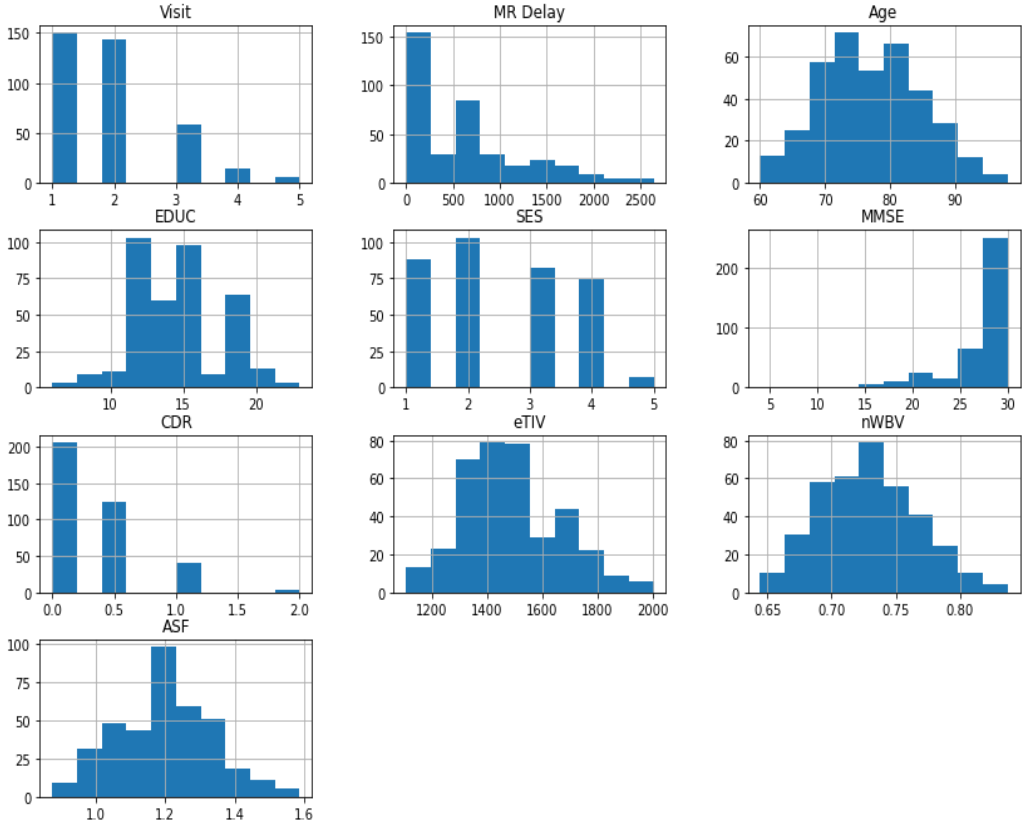

Figure 2. Histogram representation of the feature value distribution.

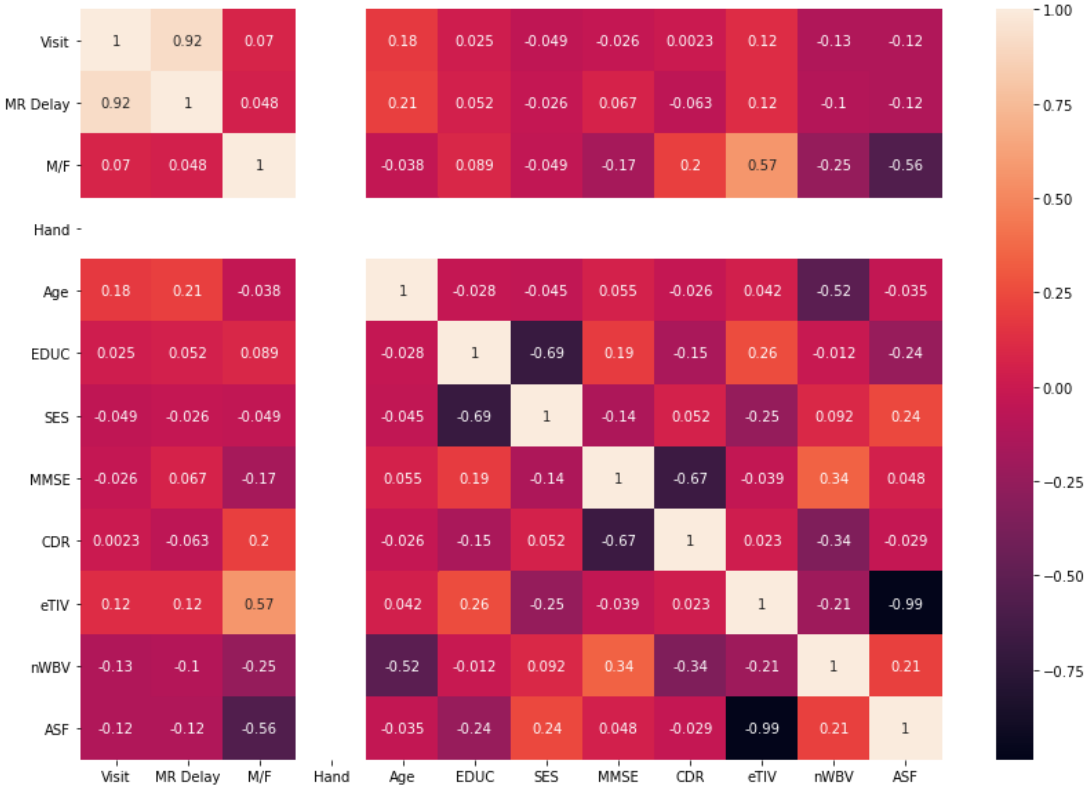

Figure 3. Correlation matrix heatmap after processing of missing values.

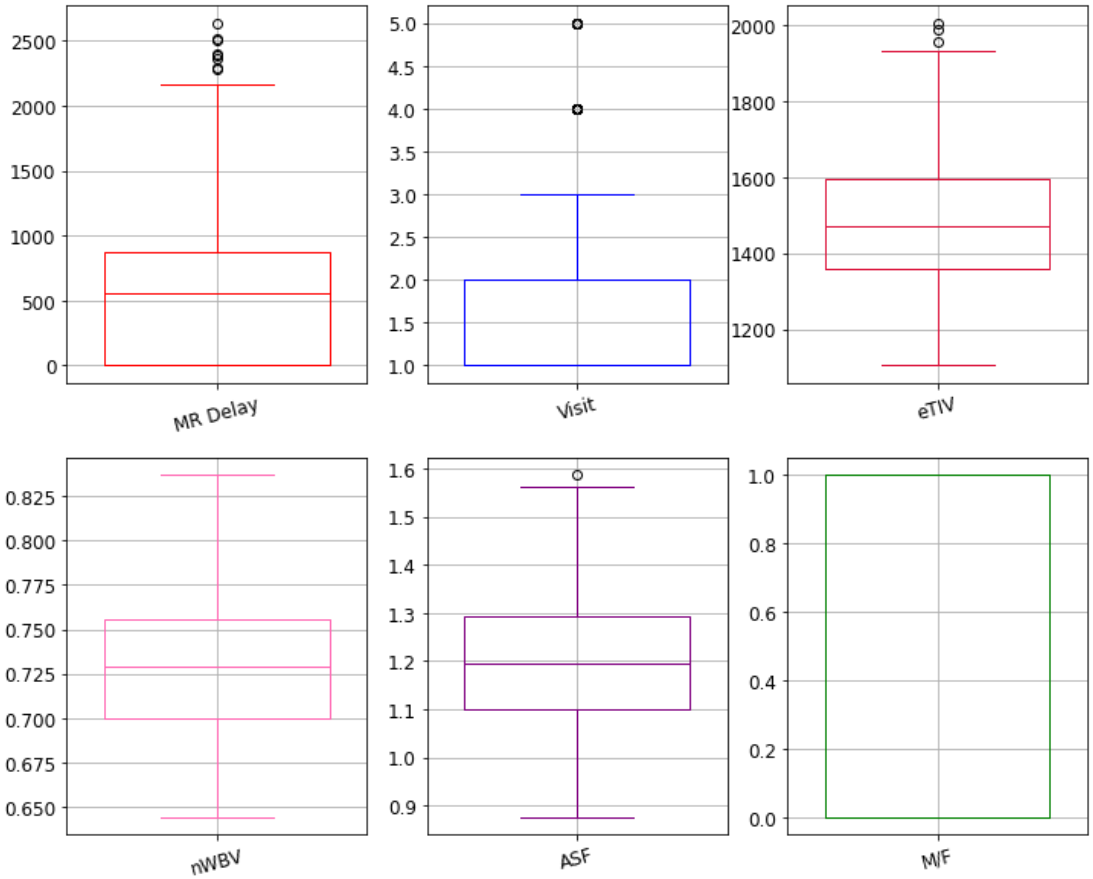

Figure 4. Box plot representation of the features with high correlation.

2.4.2. Data Splitting

In this method, we divided the dataset into three subsets for cross-validation purposes. One subset is used for model prediction (i.e., test data) and the other two sets (i.e., training and validation) are used to assess model performance by training against new data. After data preprocessing, we randomly split the whole dataset into an 80:20 ratio, where 80% was used for training and 20% was used for testing. This will enable the machine to create new combinations every time to run the model and make it possible to predict it with the highest accuracy.

After model training, the training dataset was split into two subsets for training and validation (Figure 5). The validation dataset helps to choose hyper tuning parameters, such as regularization and learning rate. These hyper tuning parameters can limit the model overfitting and improve accuracy. After a model has been performed efficiently with a validation subset, the model stops training itself at a particular epoch to avoid repeating the same experiment.

2.4.3. Training of ML Classifiers

The training of the ML classifier depends on the trained data for the prediction of the subject group across the given features. The classifier will then be well-tuned and validated on holdout data. Firstly, model training involves a process that ML can pass with the trained data and the classifier uncovers the train data patterns. Therefore, the parameters are inputted to the target variables. As mentioned, we aimed to propose an ML classifier for an explicit work of classifying AD and non-AD patients with the highest accuracy. To predict the AD patient status given to a set of independent features, we applied different supervised and ensemble learning models to propose an optimized ML classifier in AD subject categorization. Four supervised algorithms, namely Random Forest (RF), Support

Vector Machines (SVM), Naive Bayes (NB), and Logistic Regression (LR), and ensemble learning models such as gradient boosting and Adaboosting, are employed to conduct model training. A brief description of each model is given below.

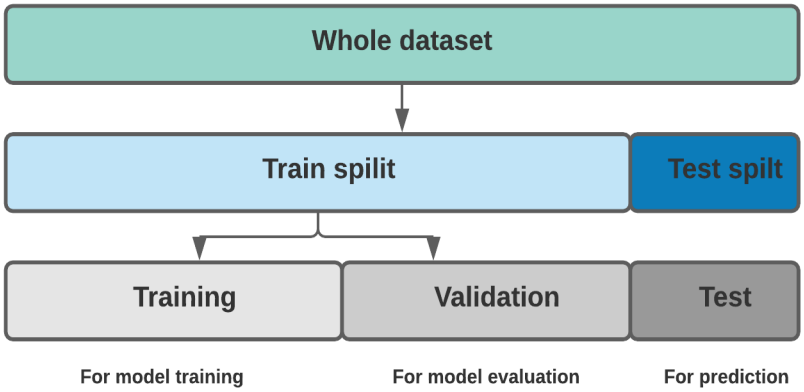

Figure 5. Schematic representation of the data splitting stage.

❖ Random Forest (RF)

The RF model is a bootstrap aggregating (bagging) model, which is implemented using a set of randomly generated decision trees or applying the divide and conquer method with random sampling, and calculates a weighted average of nodes reached [22]. For each sample taken in the training dataset, a decision tree is formed and then trained followed by grid search using 10-fold cross-validation with different parameters combinations. The classifier performance of the RF model is studied using the Gini criterion.

❖ Support Vector Machines (SVM)

The SVM is a non-linear ML classifier, which finds a hyperplane that separates the data points and classifies them into multi-dimensional space depending on the number of features [23]. It can be used for classification and regression analysis but is most often used for classification. To divide data into different classes, SVM generates the best line or decision boundary known as the hyperplane. The extreme points or vectors chosen by SVM to draw the hyperplane are known as support vectors. This hyperplane was crucial in improving the SVM model's performance. This model is implemented initially without fine-tuning, just taking the regularization parameter, C = 1, and radial basis function as the kernel. Then, fine-tuning is done as with grid search and different combinations of 'C' values and kernel functions, followed by 10-fold cross-validation. Finally, its classification or prediction performance is studied with the help of a confusion matrix.

❖ Gaussian Naive Bayes (GNB)

The GNB classifier uses the Bayes theorem and is implemented using mutually independent variables [24]. An NB classifier is a probabilistic machine learning model that uses the Bayes theorem to perform classification:

$$p(\mathrm{A}|\mathrm{B}) = \frac{p(\mathrm{B}|\mathrm{A})\ p(\mathrm{A})}{p(\mathrm{B})}$$

We calculate the probability of A occurring when features B occurred using Bayes' Theorem. The prediction or assumption is based on a strong assumption of feature independence. The predictors or features are self-contained and unrelated to one another. Because of its predictability, this model is famous in the ML environment. The GNB model is applied as a selective classifier for dementia, which calculates the set of probabilities by counting the frequency and combination of values in a given dataset. After training the GNB model, a 10-fold cross-validation was performed.

❖ Logistic Regression (LR)

The LR classifier is a linear type that is implemented similar to the SVM with dependent and independent variables, but with a greater number of values for regularization parameter 'C' [25]. This model will use the 'sigmoid function' for the prediction probability and classifier decision boundaries.

❖ Gradient Boosting

The Gradient boosting (GB) model is an ensemble ML algorithm, which utilizes a gradient boosting structure and is built on basis of the decision tree [26]. When it is implemented for structured data, decision tree-based algorithms are performing best, whereas ensemble learning algorithms outperform other algorithms, in prediction or classification problems involving unstructured data. Here, we implement the gradient boosting machine (GBM) model to classify dementias and predict the shift of MCI to AD.

❖ AdaBoost

AdaBoosting is one of the ensemble boosting classifiers, which was proposed by Yoav Freund and Robert Schapire [27]. It is an iterative ensemble learning system, which incorporates a sequential combination of several base/weak classifiers, resulting in an efficient classifier with improved accuracy. The main concept of the AdaBoost algorithm is to set the weights of classifiers and train the sample data in each iteration to predict the unusual observations accurately with minimal error.

2.4.4. Model Validation

Model validation is the practice of identifying an optimal model through skipping the train and test on the same data and helps to reduce complex overfitting issues. To overcome such an issue, we performed the cross-validation (CV) method to train the model and thereafter to calculate the accuracy [28]. It is always a challenge to validate the model with a trained dataset, and to ensure the model is noise-free, computer scientists use CV techniques. In this work, we applied the CV technique because it is a popular ML technique and produces low bias models. CV technique is also known as a k-fold approach that segregates the entire dataset into k divisions with equal size. For each iteration, the model is trained with the remaining k-1 divisions [29]. Ultimately, performance is evaluated by the mean of all k-folds for estimating the ability of the classifier problem. Usually, for the imbalanced dataset, the best value for k is 5 or 10. For this work, we applied the 10-fold CV technique, which means that model was trained and tested 10 times.

2.5. Performance Metrics

Once the ML model is created, the performance of each model can be defined in terms of different metrics such as accuracy, sensitivity, F1-score, and area under the receiver operating characteristic (AUROC) curve values. To do that, the confusion matrix can help to identify misclassification in tabular form. When the subject is classified as demented (1) is considered as a true positive, when it is classified as non-demented, (0) is considered a true negative. The confusion matrix representation of a given dataset is shown in Table 4.

Table 4. Confusion matrix of demented subjects.

Classification	1	0
D = 1	TP	FN
ND = 0	FP	TN

D: demented; ND: nondemented; TP: true-positive; TN: true-negative; FP: false-positive; FN: false-negative.

The performance measures are defined by the confusion matrix explained below.

Accuracy: The percentage of the total accurately classified outcomes from the total outcomes. Mathematically, it is written as:

$$\text{Acc } (\%) = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100$$

Precision: This is calculated as the number of true positives divided by the sum of true positives and false positives:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall (Sensitivity): This is the ratio of true positives to the sum of true positives and false negatives:

$$\text{Sensitivity } (\%) = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

AU-ROC: In medical diagnosis, the classification of true positives (i.e., true demented subjects) is vital, as leaving true subjects can lead to disease severity. In such cases, accuracy is not the only metric to evaluate model performance; therefore, in most medical diagnosis procedures, an ROC tool can help to visualize binary classification.

3. Results

After cross-validation, the classifiers were tested on a test data subset to understand how they accurately predicted the status of the AD subject. The performance of each classifier was assessed by the visualization of the confusion matrix. The confusion matrices were used to check the ML classifiers were predicting target variables correctly or not. In the confusion matrix, virtual labels present actual subjects and horizontal labels present predicted values. Figure 6 depicts the confusion matrix outcomes of six algorithms and the performance comparison of given AD classification models are presented in Table 5.

Table 5. Performance results of binary classification of each classifier.

N	Classifier	Accuracy	Precision	Recall	F-Score	AUROC
1.	Gradient boosting	97.58	0.98	0.96	0.97	0.981
2.	SVM	96.77	0.98	0.95	0.96	0.968
3.	LR	96.77	0.98	0.95	0.96	0.977
4.	RF	96.77	0.96	0.96	0.96	0.983
5.	AdaBoosting	96.77	0.96	0.96	0.96	0.971
6.	NB	95.96	0.96	0.95	0.95	0.980

As can be seen from Table 5, all given classifiers produced better accuracy in the classification of AD subjects, but gradient boosting outperforms all the adopted classifiers. The highest classification accuracy was achieved by the accusation of missing data with the most occurring values and features with high correlation values. It resulted in a high classification accuracy of 97.58% against 95.96% of NB classifiers with low accuracy among them. We can also observe that SVM, LR, RF, and Adaboosting have the same accuracy of 96.77%. As mentioned by [30], for imbalanced datasets, we cannot justify model performance through accuracy metrics; therefore, by creating ROC plots, conclusions can be drawn by the reliability of classification performance. Figure 7 presents the AUROC curves of the given algorithms.

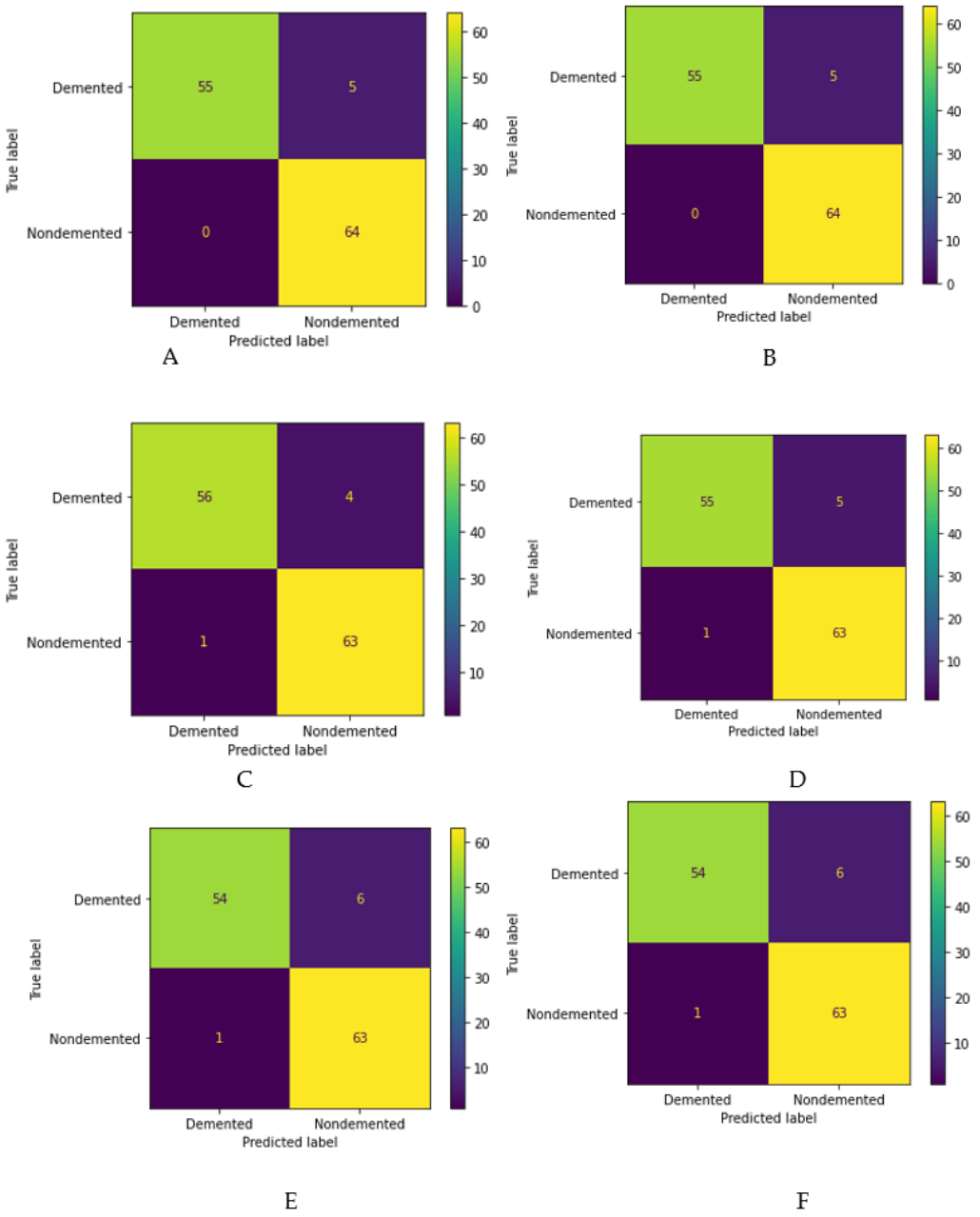

Figure 6. The confusion matrix outcomes of (**A**) Support vector machines (**B**) Logistic regression (**C**) Random Forest (**D**) Naïve Bayes (**E**) AdaBoosting (**F**) Gradient boosting.

The RF classifier had the highest AUC value of 0.983, which was followed by the values of gradient boosting (0.981) and NB classifier (0.980), and the lowest AUC value (0.968) was generated by SVM classifiers. LR and AdaBoosting presented AUC scores of 0.977 and 0.971, respectively. These observations indicate that boosting techniques outperformed the supervised models; in particular, the gradient boosting technique has a large capability in the classification of true AD subjects.

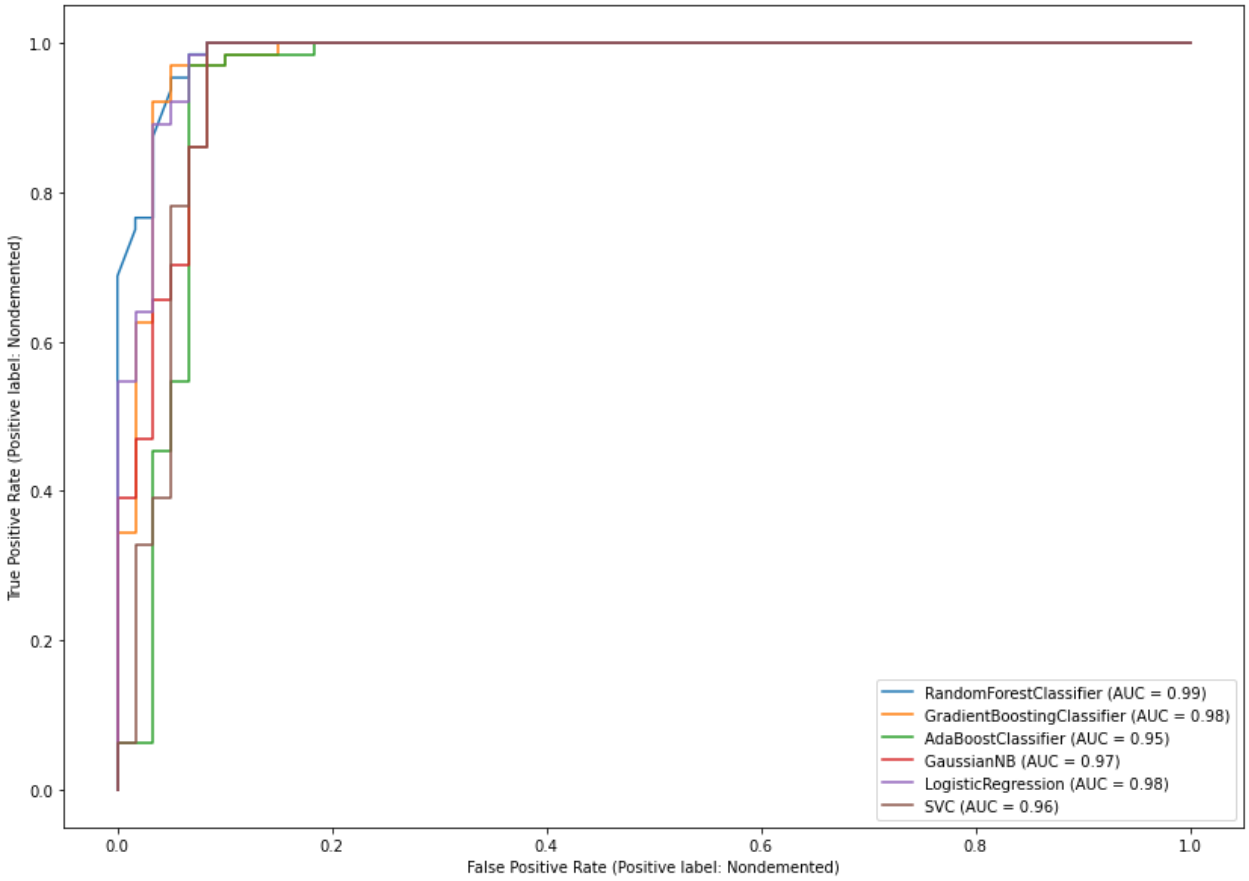

Figure 7. The area under the curve (AUC) of the classification performance of each algorithm.

4. Discussion

Adult-onset dementia disorders have serious effects on the lifestyles of people due to the loss of cognitive functions and the progression of brain atrophy. AD is the most common form of dementia and contributes to about 60–70% of adult-onset dementia cases worldwide. Unfortunately, as already mentioned in the introduction, diagnosis of AD was based on clinical and exclusion criteria which have an accuracy of 85% and do not allow a definitive diagnosis, which could only be confirmed by post-mortem evaluation. On the other hand, an early and accurate diagnosis of AD is important for timely brain health interventions. Screening among people of AD risk in preclinical stages may prompt early identification of AD pathology and suggest better remedial procedures for complying with the AD beginning. Current biomarkers of AD have required specimen collection (such as serum or liquid) or MRI data.

Finding a computational approach for early detection of AD patients who do not exhibit any clinical signs of AD at the time of the test is an open challenge [31]. As the disease's prevalence increases, several symptoms are found in cognitive abilities, such as language, memory, psychology, etc. Consequently, there is a need for precise and early diagnosis of dementia for helping health professionals to treat the disease at an aborning stage. There are a few techniques currently available for diagnosing adult-onset dementia. These include CSF (cerebrospinal fluid) measures, CT (Computer-based Tomography), MRI (Magnetic Resonance Imaging) assessments, ultrasound, and PET (photon outflow tomography) as a blend of neurological and psychological tests [32]. These approaches are expensive, could be partially invasive, and require time and dedicated resources. Hence, discovering effective strategies to identify dementia and finding sub-types is a challenging issue today.

Previous studies on dementia detection using conventional approaches, such as laboratory tests, patient medical history, or behavioral changes, produced less accuracy in AD

diagnosis. Subsequently, computer researchers incorporated ML technologies in neurological diagnostic procedures. ML modeling was used to predict the conversion of MCI to dementia patients with a focus on cognitive reserve among 169 subjects [33]. The outcomes showed that the gradient boosting algorithm generated the highest accuracy, i.e., 93%, and also proved that cognitive reserve can play an important role in the conversion of MCI to dementia patients. It is reported that ML models can help to distinguish age-related cognitive decline (ARCD) from different dementia types (including AD, MCI, and VD) using neurocognitive tests [34].

Most health informatics researchers prefer data-centric machine learning approaches in the diagnosis of early-stage AD [35–37]. In data-centric approaches, data are systematically changed or preprocessed for the datasets for enhancing the performance of ML models. This is generally ignored and data collection is considered as one of the tasks. It is all about the data quality which helps to accurate data labelling [38]. The era of data-driven approaches in dementia assessment is generated with the capacity to alternate the healthcare systems with more efficiency, transparency, and personalized services for AD. The main reason behind the "dirty" AD clinical data is because there is no standardization in pathways of dementia care. For example, the dementia-related data in Northern Ireland is properly retrieved and analyzed based on the social and healthcare organizations, but the set of datasets of dementia assessments across diverse practice sites can be different. Similarly, doctors in England are also followed similar non-standardization approaches in dementia assessments [39]. This research was further validated by proposing data-driven approaches based on deep learning models from data dementia patients for calculating the agitation rates [37].

The studies with the experimental setup of ML-based data-centric methods with preprocessed MRI data can help efficient screening of MCI levels. For instance, some authors that adopted kernel density estimations for extracting texture information from the MR images and reported linear discriminant analysis (LDA) and SVM achieved high detection accuracy with limited features [40]. AD diagnosis through data preprocessing-based recursive feature elimination is proposed in [41], and results produced the highest AD subjects classification accuracy with different levels of dementia.

There is a scarcity of works that proposed data-centric ML models on demographic MRI features; rather, most of them focused on the image related datasets. Therefore, the present work strives to achieve comprehensive performance analysis in the classification of AD patients and proposed data-driven ML methodologies which utilize the data of longitudinal MRI features. Handling of missing data was done by replacing the highest occurrence value followed by normalization and standardization. With the adoption of EDA techniques, we present the feature dataset distribution and inclusion of the highest correlated features along with outliers helped us achieve the highest classification accuracy. Thereafter, we trained six different ML classifiers without reducing the dimensions of the data.

The data-driven ML classifiers were used to successfully classify the true dementia subjects and these studies were carried out by applying a combination of supervised and boosting algorithms. The advantage of conducting these types of studies can help the early identification of AD and as a result reduce medical expenses and contribute to undertaking therapeutic measures. Despite generating the highest classification accuracy, this study has some limitations, namely the small sample size involved in the final dementia subject classification. The OASIS datasets are very popular in brain studies. However, incorporation of external MRI data cannot guarantee the data quality and this can affect the study significance.

5. Conclusions

ML research associated with neurological studies can offer a more precise analysis of AD. We proposed a framework based on supervised learning models in the classification of AD patients into two categories, i.e., either AD or non-AD, based on longitudinal

brain MRI features. It was also possible to predict individual dementia of older adults with a screening of AD data by ML classifiers. To predict the AD subject status, the MRI demographic information and pre-existing conditions of the patient can help to enhance the classification accuracy. Three classifiers (RF, NB, and Gradient boosting) produced the highest average AUC scores of 0.98. However, by considering both classification accuracy metric and AUC, the gradient boosting technique can seem a better potential classifier than others. In this study, we suggested a simple and efficient method of dementia subject identification technique by using ML classifiers. More sophisticated prediction models with detailed subject data and clinical features around the world should be investigated in future studies.

Author Contributions: Conceptualization, G.B. and M.A.H.; methodology, G.B.; software, M.A.H.; validation, G.B., M.A.H. and N.C.; formal analysis, G.B.; investigation, M.A.H.; resources, N.C.; data curation, G.B.; writing—original draft preparation, G.B., M.A.H., V.R.D., and M.R.; writing—review and editing, F.A.; visualization, E.T.; supervision, F.A.; project administration, F.A.; funding acquisition, F.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by institutional grant of the University of Camerino.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest during the preparation and publication of the manuscript.

References

1. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. *Alzheimer's Dement.* **2016**, *12*, 459–509. [CrossRef] [PubMed]
2. Serrano-Pozo, A.; Frosch, M.P.; Masliah, E.; Hyman, B.T. Neuropathological Alterations in Alzheimer Disease. *Cold Spring Harb. Perspect. Med.* **2011**, *1*, a006189. [CrossRef]
3. Vradenburg, G. A pivotal moment in Alzheimer's disease and dementia: How global unity of purpose and action can beat the disease by 2025. *Expert Rev. Neurother.* **2015**, *15*, 73–82. [CrossRef] [PubMed]
4. Mueller, S.G.; Weiner, M.W.; Thal, L.J.; Petersen, R.C.; Jack, C.R.; Jagust, W.; Trojanowski, J.Q.; Toga, A.W.; Beckett, L. Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimer's Dement.* **2005**, *1*, 55–66. [CrossRef]
5. Risacher, S.L.; Saykin, A.J. Neuroimaging and Other Biomarkers for Alzheimer's Disease: The Changing Landscape of Early Detection. *Annu. Rev. Clin. Psychol.* **2013**, *9*, 621–648. [CrossRef]
6. Wu, C.-L.; Lin, T.-J.; Chiou, G.-L.; Lee, C.-Y.; Luan, H.; Tsai, M.-J.; Potvin, P.; Tsai, C.-C. A Systematic Review of MRI Neuroimaging for Education Research. *Front. Psychol.* **2021**, *12*, 1763. [CrossRef]
7. Gaubert, S.; Houot, M.; Raimondo, F.; Ansart, M.; Corsi, M.-C.; Naccache, L.; Sitt, J.D.; Habert, M.-O.; Dubois, B.; Fallani, F.D.V.; et al. A machine learning approach to screen for preclinical Alzheimer's disease. *Neurobiol. Aging* **2021**, *105*, 205–216. [CrossRef] [PubMed]
8. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 1–21. [CrossRef] [PubMed]
9. Battineni, G.; Sagaro, G.G.; Chinatalapudi, N.; Amenta, F. Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis. *J. Pers. Med.* **2020**, *10*, 21. [CrossRef] [PubMed]
10. Spooner, A.; Chen, E.; Sowmya, A.; Sachdev, P.; Kochan, N.A.; Trollor, J.; Brodaty, H. A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. *Sci. Rep.* **2020**, *10*, 20410. [CrossRef]
11. Nori, V.S.; Hane, C.A.; Crown, W.H.; Au, R.; Burke, W.J.; Sanghavi, D.M.; Bleicher, P. Machine learning models to predict onset of dementia: A label learning approach. *Alzheimer's Dement. Transl. Res. Clin. Interv.* **2019**, *5*, 918–925. [CrossRef]
12. Ghazi, M.M.; Nielsen, M.; Pai, A.; Cardoso, M.J.; Modat, M.; Ourselin, S.; Sørensen, L. Training recurrent neural networks robust to incomplete data: Application to Alzheimer's disease progression modeling. *Med. Image Anal.* **2019**, *53*, 39–46. [CrossRef]
13. Moore, P.J.; Lyons, T.J.; Gallacher, J.; Initiative, F.T.A.D.N. Random forest prediction of Alzheimer's disease using pairwise selection from time series data. *PLoS ONE* **2019**, *14*, e0211558. [CrossRef] [PubMed]
14. Jo, T.; Nho, K.; Saykin, A.J. Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data. *Front. Aging Neurosci.* **2019**, *11*, 220. [CrossRef]
15. Lundervold, A.; Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. *Z. Med. Phys.* **2018**, *29*, 102–127. [CrossRef]

16. Battineni, G.; Chintalapudi, N.; Amenta, F. Comparative Machine Learning Approach in Dementia Patient Classification using Principal Component Analysis. In Proceedings of the 12th International Conference on Agents and Artificial Intelligence, Valletta, Malta, 22–24 February 2020. [CrossRef]
17. Aditya, C.; Pande, M.S. Devising an interpretable calibrated scale to quantitatively assess the dementia stage of subjects with alzheimer's disease: A machine learning approach. *Inform. Med. Unlocked* **2017**, *6*, 28–35. [CrossRef]
18. Marcus, D.S.; Fotenos, A.F.; Csernansky, J.G.; Morris, J.C.; Buckner, R.L. Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults. *J. Cogn. Neurosci.* **2010**, *22*, 2677–2684. [CrossRef] [PubMed]
19. Atri, A. The Alzheimer's Disease Clinical Spectrum: Diagnosis and Management. *Med. Clin. N. Am.* **2019**, *103*, 263–293. [CrossRef] [PubMed]
20. Balota, D.A.; Tse, C.-S.; Hutchison, K.A.; Spieler, D.H.; Duchek, J.M.; Morris, J.C. Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in stroop color naming. *Psychol. Aging* **2010**, *25*, 208–218. [CrossRef]
21. Towards Data Science. All about Missing Data Handling. Missing Data Is a Every Day Problem . . . , by Baijayanta Roy. Available online: https://towardsdatascience.com/all-about-missing-data-handling-b94b8b5d2184 (accessed on 5 October 2021).
22. Yang, B.-S.; Di, X.; Han, T. Random forests classifier for machine fault diagnosis. *J. Mech. Sci. Technol.* **2008**, *22*, 1716–1725. [CrossRef]
23. Suykens, J. Support Vector Machines: A Nonlinear Modelling and Control Perspective. *Eur. J. Control.* **2001**, *7*, 311–327. [CrossRef]
24. Raizada, R.D.S.; Lee, Y.-S. Smoothness without Smoothing: Why Gaussian Naive Bayes Is Not Naive for Multi-Subject Searchlight Studies. *PLoS ONE* **2013**, *8*, e69566. [CrossRef]
25. Maroco, J.; Silva, D.; Rodrigues, A.; Guerreiro, M.; Santana, I.; De Mendonça, A. Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Res. Notes* **2011**, *4*, 299. [CrossRef]
26. Singh, U.; Rizwan, M.; Alaraj, M.; Alsaidan, I. A Machine Learning-Based Gradient Boosting Regression Approach for Wind Power Production Forecasting: A Step towards Smart Grid Environments. *Energies* **2021**, *14*, 5196. [CrossRef]
27. Boosting and AdaBoost for Machine Learning. Available online: https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/ (accessed on 27 August 2021).
28. Watanabe, S. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *J. Mach. Learn. Res.* **2010**, *11*, 3571. Available online: http://jmlr.org/papers/v11/watanabe10a.html (accessed on 7 October 2021).
29. Berrar, D. Cross-Validation. *Encycl. Bioinform. Comput. Biol.* **2019**, *1*, 542–545. [CrossRef]
30. Saito, T.; Rehmsmeier, M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. *PLoS ONE* **2015**, *10*, e0118432. [CrossRef] [PubMed]
31. Ding, X.; Bucholc, M.; Wang, H.; Glass, D.H.; Wang, H.; Clarke, D.H.; Bjourson, A.J.; Dowey, L.R.C.; O'Kane, M.; Prasad, G.; et al. A hybrid computational approach for efficient Alzheimer's disease classification based on heterogeneous data. *Sci. Rep.* **2018**, *8*, 9774. [CrossRef]
32. Turner, R.; Stubbs, T.; Davies, D.A.; Albensi, B.C. Potential New Approaches for Diagnosis of Alzheimer's Disease and Related Dementias. *Front. Neurol.* **2020**, *11*, 496. [CrossRef] [PubMed]
33. Facal, D.; Valladares-Rodriguez, S.; Lojo-Seoane, C.; Pereiro, A.X.; Anido-Rifon, L.; Juncos-Rabadán, O. Machine learning approaches to studying the role of cognitive reserve in conversion from mild cognitive impairment to dementia. *Int. J. Geriatr. Psychiatry* **2019**, *34*, 941–949. [CrossRef] [PubMed]
34. Er, F.; Iscen, P.; Sahin, S.; Çinar, N.; Karsidag, S.; Goularas, D. Distinguishing age-related cognitive decline from dementias: A study based on machine learning algorithms. *J. Clin. Neurosci.* **2017**, *42*, 186–192. [CrossRef] [PubMed]
35. Wong-Lin, K.; McClean, P.L.; McCombe, N.; Kaur, D.; Sanchez-Bornot, J.M.; Gillespie, P.; Todd, S.; Finn, D.P.; Joshi, A.; Kane, J.; et al. Shaping a data-driven era in dementia care pathway through computational neurology approaches. *BMC Med.* **2020**, *18*, 398. [CrossRef] [PubMed]
36. Weiss, J.; Puterman, E.; Prather, A.A.; Ware, E.B.; Rehkopf, D.H. A data-driven prospective study of dementia among older adults in the United States. *PLoS ONE* **2020**, *15*, e0239994. [CrossRef] [PubMed]
37. Goins, H.; HekmatiAthar, S.; Byfield, G.; Samuel, R.; Anwar, M. Toward Data-Driven Assessment of Caregiver's Burden for Persons with Dementia using Machine Learning Models. In Proceedings of the 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), Las Vegas, NV, USA, 11–13 August 2020; pp. 379–384. [CrossRef]
38. Anik, A.I.; Bunt, A. Data-Centric Explanations: Explaining Training Data of Machine Learning Systems to Promote Transparency. In Proceedings of the CHI'21: CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021. [CrossRef]
39. Koch, T.; Iliffe, S. Implementing the National Dementia Strategy in England: Evaluating innovative practices using a case study methodology. *Dementia* **2011**, *10*, 487–498. [CrossRef]
40. Veluppal, A.; Sadhukhan, D.; Gopinath, V.; Swaminathan, R. Detection of Mild Cognitive Impairment using Kernel Density Estimation based texture analysis of the Corpus Callosum in brain MR images. *IRBM* **2021**, in press. [CrossRef]
41. Richhariya, B.; Tanveer, M.; Rashid, A. Diagnosis of Alzheimer's disease using universum support vector machine based recursive feature elimination (USVM-RFE). *Biomed. Signal Process. Control.* **2020**, *59*, 101903. [CrossRef]