ALZHEIMER'S DISEASE PREDICTION USING SUPPORT VECTOR ALGORITHM (ML)

[1]Mrs. G. SWAPNA, [2]ALIGETI VIVEK,[3] RADANDLA ROHITH, [4]ODETI MANOHAR [5]HALAVATH JAGAN

[1](Assistant Professor), CSE, J.B. Institute Of Engineering & Technology
[2345]B.Tech Scholor, CSE, J.B. Institute Of Engineering & Technology

ABSTRACT

Alzheimer's disease is a progressive neurodegenerative disorder that affects memory, cognition, and behavior. Early detection of Alzheimer's disease is crucial for successful intervention and therapy. In this project, we focus on predicting the progression of Alzheimer's disease using a Support Vector Machine (SVM) classifier. The first stage of Alzheimer's disease is known as Mild Cognitive Impairment (MCI). Identifying MCI subjects who are at high risk of developing full-blown Alzheimer's over time is essential. To track the disease's progression, we employ automated modeling techniques. Three separate longitudinal data sets are used to train our models. These models are then evaluated using biomarker data from experimental investigations. Finally, the SVM classifier is employed to identify MCI patients at risk of conversion to Alzheimer's in the future. Predicting MCI conversion 1 year and 2 years ahead. Noteworthy findings include the superiority of magnetic resonance image-based features over cognitive points for predicting clinical changes. Additionally, using multiple predictive models outperforms single biomarker models, and neuropsychology programs alone may provide long-term change prediction. The ultimate goal of this research is to enhance the accuracy of Alzheimer's disease prediction using SVM classifier modeling.

1. INTRODUCTION

Alzheimer's Disease (AD) is a neurological disorder that causes the death of nerve cells in the human brain. AD usually begins gradually and its first symptoms may be attributed to the increment of the age or common forgetfulness. As the disease progresses, the patient's cognitive abilities deteriorate, including the ability to make decisions and carry out daily tasks. Currently there is no cure for the disease, only a series of guidelines can be followed to perhaps delay the progress of it. For this reason, an effective diagnosis will be a key factor in order to improve the quality of life of their patients. The motivation for the creation of innovation to support the battle against Alzheimer's disease is evident, not only from an ethical perspective but also due to the continuous proliferation of Alzheimer's cases in our society. Today, 50 million people worldwide live with dementia, where two-thirds of them have Alzheimer's disease. Alzheimer's cases have overtaken cancer ones to become the most feared disease in the United States, with a new case appearing every three seconds in the world. At the moment the diagnosis of this disease is made by combining an analysis of the patient's medical history, different cognitive tests and various clinical tests, such as photographic scans of the brain. But is all this enough given the importance of an early diagnosis in the treatment of the disease? Nowadays, through Machine Learning, it is possible to analyze data on a large scale and with different algorithms, detecting patterns and models in a very short period of time. In this way, there is a significant improvement in diagnostic methods using techniques which are even imperceptible to human experience and reasoning. On the top of that, these days the world of Machine Learning is more advanced than ever before, thanks to the newest deep neural networks. Simply explained, deep neural networks enable the creation of systems which are powerful enough to represent any finite deterministic mapping between any given set of inputs and a set of corresponding outputs.

These networks allow powerful data processing, allowing processes as complex as image identification or natural language processing. In view of all the aforementioned, the aim of this project will be to analyze the possible connection between an improvement in the diagnosis of AD and the latest deep learning techniques. The number of variables that can influence the appearance or not of Alzheimer's disease are numerous and above all uncertain, being the human capacity a bounded resource to detect early cases of the disease with confidence. So, would it be possible to analyze all these variables through different deep learning techniques in order to offer a result that indicates the probability 10 of developing such disease? Perhaps technology can be united once again with the medicine to discover new methods that allows to reveal the most determining parameters in the presence of the disease. Alzheimer's is a hazardous, progressive disease affecting the brain and nervous system. Early diagnosis of Alzheimer's allows for more effective treatments and a powerful and effective treatment strategy [9]. Magnetic resonance imaging (MRI) as a diagnostic tool for identifying plaque and affected regions holds significant value in detecting Alzheimer's disease [10]. One of the research's primary concerns is identifying Alzheimer's diseaseaffected areas using magnetic resonance imaging accurately [11]. The issue can be viewed from two different perspectives. The nature of the classification is the first aspect of the problem, and it must be determined which images have Alzheimer's disease symptoms and which do not. Another aspect of the problem is the character of the zoning, which necessitates the identification of Alzheimer's-affected areas [12]. To identify plaques in the brains of Alzheimer's patients using magnetic resonance imaging, segmentation and classification techniques are required in the images. Zoning techniques aim to separate distinct types of brain tissue and isolate damaged areas from healthy ones

OVERVIEW

Alzheimer's Disease (AD), a debilitating neurological disorder characterized by the progressive loss of nerve cells in the brain. With no current cure, early diagnosis becomes crucial for implementing effective treatment strategies and improving patient outcomes. Traditional diagnostic methods involve analyzing medical history, cognitive tests, and clinical scans, but the question arises whether these methods suffice given the disease's significance. The text highlights the potential of Machine Learning (ML) techniques, particularly deep learning, in enhancing AD diagnosis. ML algorithms can analyze large-scale data sets and detect patterns imperceptible to human observation, offering a promising avenue for improving diagnostic accuracy. Specifically, Magnetic Resonance Imaging (MRI) emerges as a valuable tool for identifying AD-related changes in the brain, such as plaque formation. The research focus narrows down to two primary concerns: classification and zoning of AD-affected areas in MRI scans. Classification involves distinguishing between images with AD symptoms and those without, while zoning aims to identify specific regions within the brain affected by the disease. To achieve this, segmentation and classification techniques are employed to isolate damaged areas from healthy brain tissue.

MOTIVATION

In the current landscape of healthcare, there often exists a dissonance between human intuition and conventional measurements when it comes to diagnosing and treating diseases. This incongruity poses a significant challenge in providing optimal healthcare solutions. To bridge this gap effectively, we must turn to innovative approaches, particularly those rooted in advanced technology such as machine learning (ML). ML techniques, characterized by their computational intensity and departure from traditional methodologies, offer a promising avenue for revolutionizing disease prediction and treatment customization. By leveraging vast amounts of data, ML algorithms can discern intricate patterns and correlations that might elude human observation. This capability holds immense potential in healthcare, particularly in the realm of disease prediction and visualization. One of the primary advantages of ML in healthcare lies in its ability to offer prescient and personalized prescriptions. By analyzing a patient's medical history, genetic makeup, lifestyle factors, and other pertinent data points, ML algorithms can generate tailored treatment plans that account for individual variability and optimize therapeutic outcomes.

This personalized approach not only enhances patients' quality of life but also empowers physicians in making more informed treatment decisions. Moreover, the integration of ML into healthcare systems facilitates more comprehensive analyses for health economists and policymakers. By synthesizing large-scale healthcare data, ML algorithms can uncover nuanced insights into disease trends, treatment efficacy, and healthcare resource allocation. This wealth of information enables stakeholders to devise more effective strategies for improving public health outcomes and optimizing healthcare delivery. However, despite its transformative potential, the adoption of ML in healthcare is not without challenges. Integrating ML frameworks with electronic health record (EHR) derived data poses technical and ethical considerations. Ensuring data privacy, maintaining data quality, and addressing biases inherent in datasets are critical concerns that must be addressed to harness the full benefits of ML in healthcare.

2. LITERATURE SURVEY

The Tosa straw RS-fMRI measures spontaneous activities consolidated into several distinct resting state networks (RSNs) that exhibit comparable temporal characteristics. SMRI-derived images have been utilized in previous studies. Structural adaptations in SMRI images pertain to the collective arrangement of neurons or neural elements, physically or through synaptic networks. Researchers have been focusing on using structural images to diagnose Alzheimer’s disease throughout history. The utilization of said images has presented challenges for scholars because of the restricted extraction attributes and the resultant precision of prognostication. Various approaches have been implemented to identify the indications and manifestations of advertising. In advertisements, sophisticated equipment has been widely used to evaluate the disease’s severity. The authors provided a retrospective analysis of the illness and its mechanisms. This work delves into the study and provision of unique techniques for photograph processing, enhancement, and segmentation.

The methodology employs magnetic resonance (MR) brain imaging to ascertain the interconnectivity between distinct brain regions. A novel artificial intelligence-based approach has been developed to detect alterations in individuals with a heightened susceptibility to Alzheimer’s disease a decade before the manifestation of the ailment. The approach utilizes brain magnetic resonance imaging (MRI) scans to determine the interconnections among diverse brain regions. Alzheimer’s disease is the predominant among elderly individuals and is characterized by neurological dysfunction that results in compromised memory and cognitive functions. The development of an effective protocol for the timely detection of the ailment is of paramount importance. Although a cure for the disease has not yet been discovered, the prompt administration of drugs currently in development will likely enhance their efficacy.

The timely identification of the ailment may also enable individuals to adopt modifications in their lifestyle that can impede the advancement of the condition. The researchers initially used a dataset of 67 MRI images to train the algorithm. Among these images, 38 were obtained from individuals who had been diagnosed with Alzheimer’s disease, while the remaining 29 were acquired from healthy subjects. The goal of the study was to teach the algorithm to distinguish between brains that are suffering from a pathological condition and those that are in a normal physiological state.

The researchers partitioned the brain scan images into smaller segments. The optimal size of the divided pieces was determined by manipulating their size across multiple trials. The algorithm underwent evaluation by utilizing neuroimaging data derived from a separate cohort of 148 individuals. The study’s sample comprised 100 participants, with 52 categorized exhibiting normal health and 48 diagnosed with Alzheimer’s disease. A cohort of 48 participants exhibiting mild cognitive impairment (MCI) were observed to have progressed to a diagnosis of Alzheimer’s disease within 2.5 to 9 years. With a precision rate of 86%, the AI algorithm demonstrates a keen ability to distinguish between a brain in a healthy state and a brain affected with Alzheimer’s disease. The study exhibited a significant capacity to differentiate between a healthy brain and one with mild cognitive impairment, achieving an accuracy rate of 84%. According to the algorithm exhibits good usefulness, mainly when prophylactic interventions for Alzheimer’s disease are present. Initially, the researchers trained the algorithm on a dataset comprising 67 MRI images, of which 38 were obtained from individuals diagnosed with Alzheimer’s disease and 29 were acquired from healthy subjects. The objective was to instruct the algorithm to discriminate between diseased and healthy brains. The brain scan images were partitioned into smaller segments by the researchers. Through the manipulation of the size of the divided pieces in various trials, the optimal piece size was determined. 14 The study in reference utilized a hybrid approach involving SMRI scans, cognitive criteria, and age to diagnose individuals with Alzheimer’s disease.

The level of precision in forecasting is 82%. The present study, conducted by Ms. Moradi and her colleagues, aimed to eliminate low-resolution features by selecting features from SMRI scans. In a manner akin to Ms. Moradi’s publication, the authors of the article [98] employed cognitive criteria in conjunction with the features of SMRI and PET images to forecast the onset of Alzheimer’s disease. This article reports on a study investigating the ability to predict Alzheimer’s disease in patients for 24 months. The results indicate a predictive accuracy of 78%. Furthermore, the Association for Alzheimer’s Disease and Related Disorders (NINCDS) (ADRDA) developed criteria for diagnosing Alzheimer’s disease. This group included the patient’s medical history, clinical examination, and neuropsychological and laboratory evaluations.

3. SYSTEM DESIGN

3.1 System Architecture

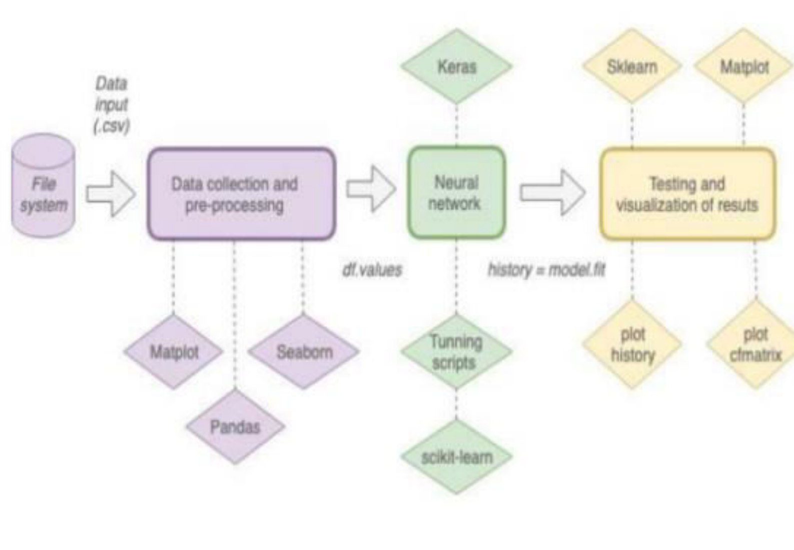
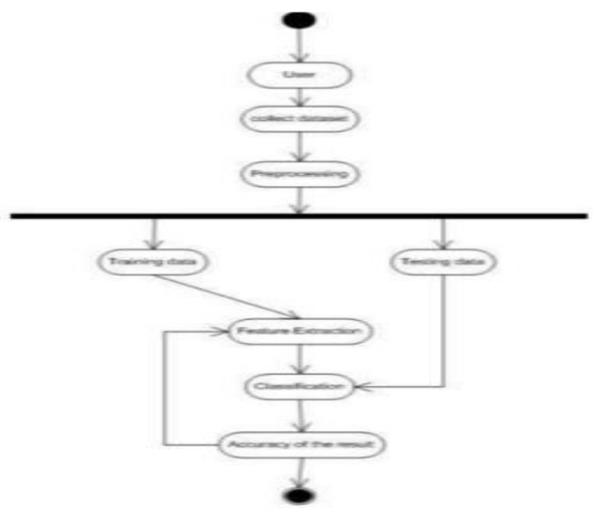


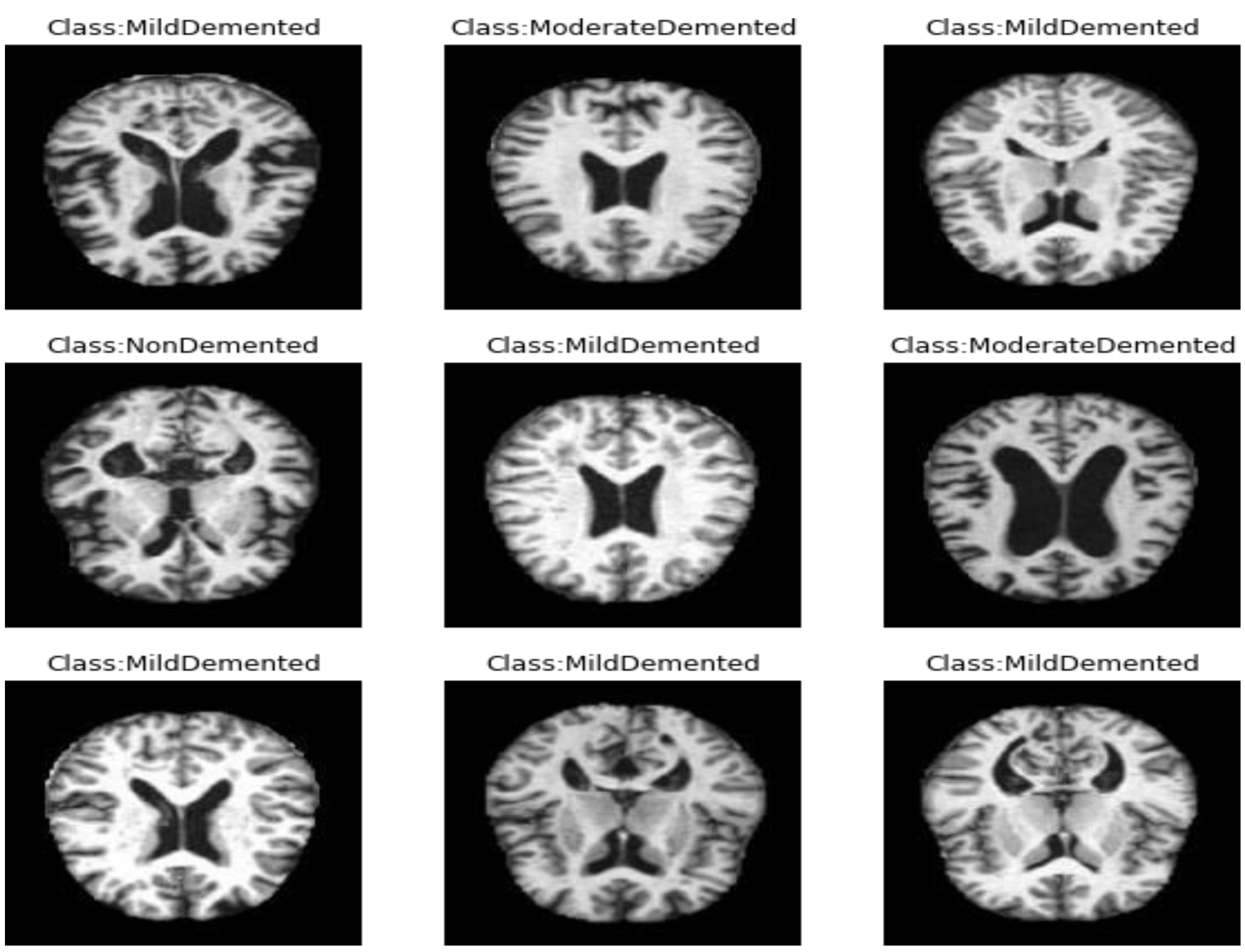
FIGURE 5.1. Conceptual System Architecture

The system's architecture is divided into three main modules: Data collection and pre-processing, Neural Network, and Testing, each handling specific functionalities. Modules interact to process data sequentially, with the final result emerging from the last module. Libraries are utilized to provide predefined functions, with sub-modules developed for specific functions if necessary. Input data, meeting non-functional requirements and is presented as CSV files and feeds into the Data collection and pre-processing module, responsible for cleaning and preparing data, fulfilling functional requirements #1 and #2 and non-functional requirement #7. 23 The Neural Network module, representing the system's core, designs the network based on pre-processed data, meeting functional requirements #3 and non-functional requirements #9, #13, and #10. The Testing module, powered by results from the Neural Network module, evaluates system performance, addressing requirements #11 and #12 to improve the solution and detect issues.

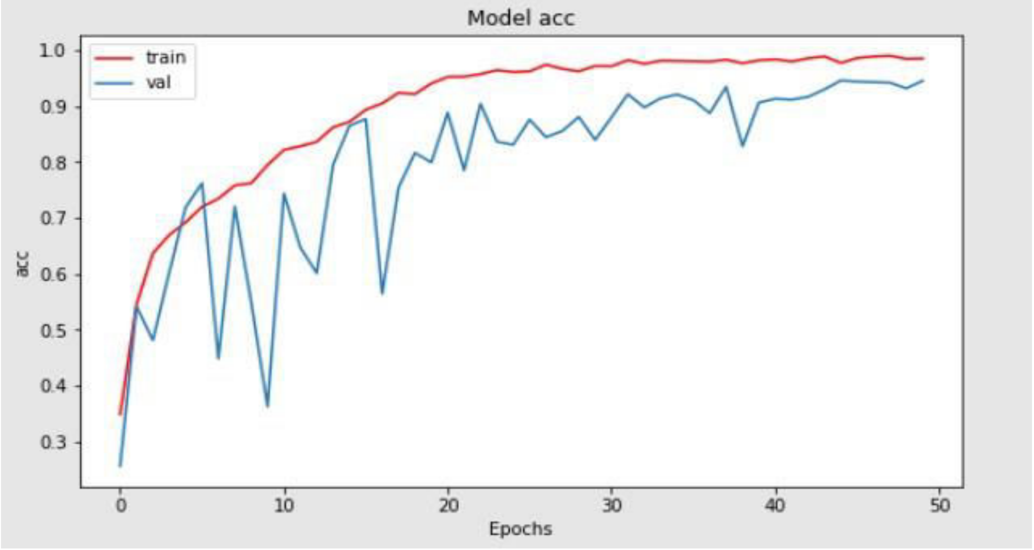
ACTIVITY DIAGRAM



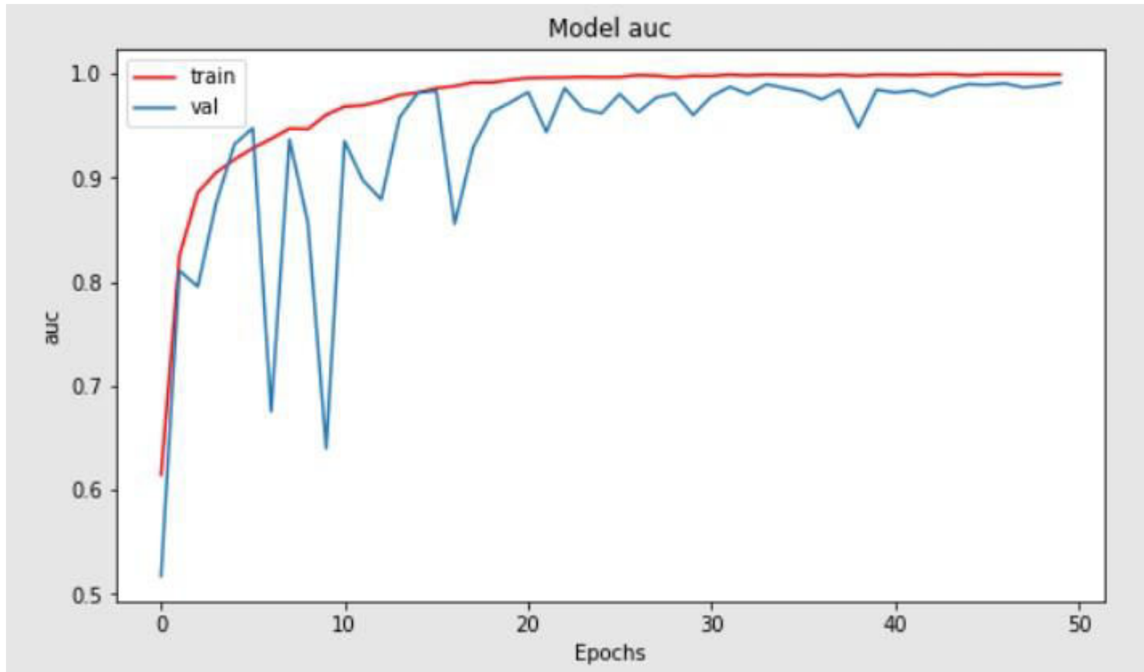
4. OUTPUT SCREENS



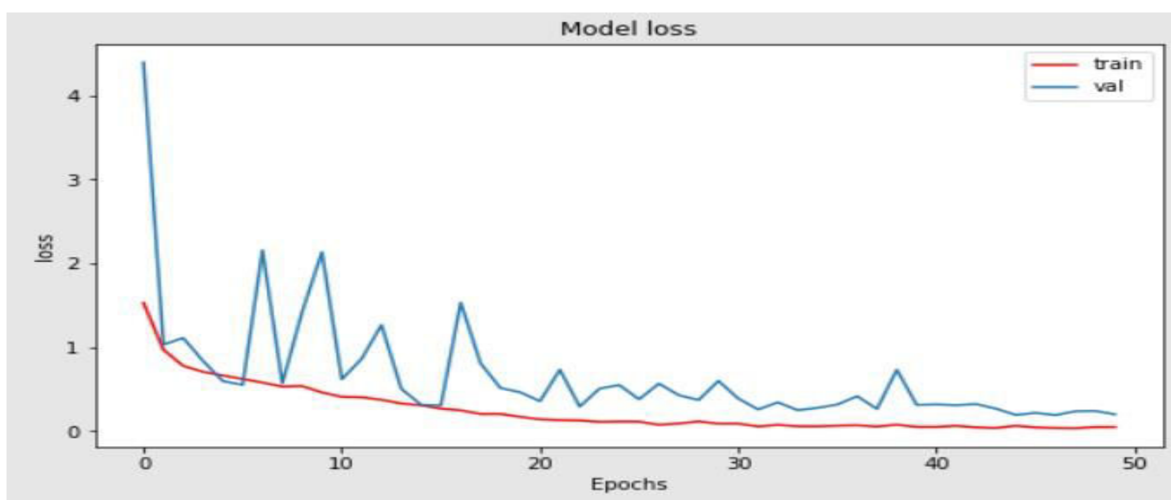
a) CNN Model Layers



b) **Training Model Accuracy**



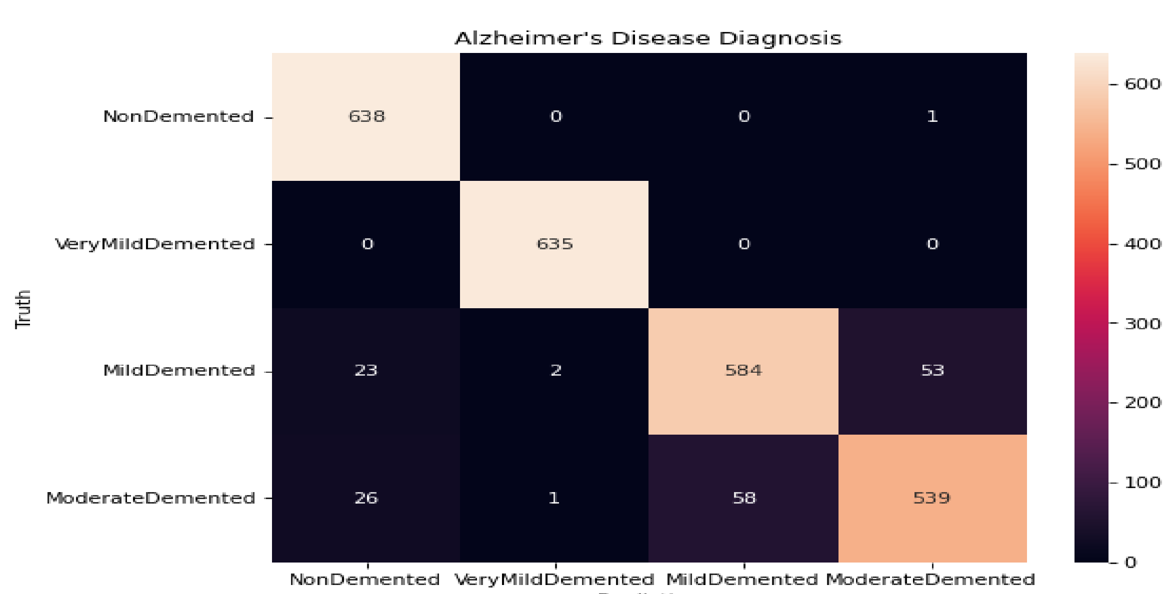
c) **Training Model AUC**



d) **Training Model Loss**

	precision	recall	f1-score	support
NonDemented	0.93	1.00	0.96	639
VeryMildDemented	1.00	1.00	1.00	635
MildDemented	0.91	0.88	0.90	662
ModerateDemented	0.91	0.86	0.89	624
micro avg	0.94	0.94	0.94	2560
macro avg	0.94	0.94	0.94	2560
weighted avg	0.94	0.94	0.94	2560
samples avg	0.94	0.94	0.94	2560

e) **Detailed Final Model Analysis of Each Category**



5. CONCLUSION

The objective of the present project was to answer the research question presented on Section 1.0. For this purpose, an analysis of the possible improvement on the diagnostic process of AD by means of deep learning techniques has been addressed through the implementation of the project. It is for this reason that, during this last chapter, an answer to this question will be exposed explaining its conclusions and future improvements of the system. As a resume, the result of the analysis was to build a classifier able to discriminate data from patients with AD or not. To do so, the first step was to collect the data needed to train the classifier. But before that the data can be used, the data had to be cleaned and pre-processed. Once the data was ready to use to train a model, an appropriate model architecture had to be selected to build the classifier. Thus, a deep neural network was designed. In order to do so, a MLP architecture was selected. Different design considerations were tested, as the proper selection of the model hyperparameters or the analysis of the learning process. Finally,

the performance of the model has been evaluated through his accuracy, confusion matrix and ROC curve based on previously unseen data. In order to answer properly the research question, firstly, its sub-questions have to be addressed. • Which data will be necessary in order to train the system successfully? The NACC dataset has been the one selected. It has been proved as being relevant, variate and big enough obtaining an accuracy of 82.61%. • Which is the most suitable architecture and parameters to achieve an accurate result? But not only a good dataset is enough to achieve accurate results, as it can be seen through the analysis and testing chapter the proper selection of the parameters of the model will be a key aspect to consider.

The best architecture will differ from case to case, but given the dataset selected and the problem to be solved, the most suitable architecture has result ina multilayer percepton with one hidden layer and 25 neurons. The sum of model 30 parameters which provides best accuracy results was Adamax as optimizer, Tanh asactivation function, 5000 as batch size and 1600 as number of epochs. • What level of accuracy can be achieved? As it has been seen on Chapter 6, the system achieved 82.61% of accuracy, being the percentage of predictions that the model made correctly with respect to the total number of predictions to be made. This result can be defined as successful, since the accuracy values obtained by similar projects on the field are between 60% and 90%. • What framework could be used to implement and test the selected model? The programming language used to implement the system was Python, making use of Keras, Pandas, Matplot and Seaborn libraries.

These libraries offer the best trade-off between performance and easy to use, being the best option for the present project. To conclude, it has been demonstrated that it is possible to measure the probability of havingAD by means of combining clinical history, lifestyle habits and cognitive examinations presentingan accuracy result of 82.61%. Moreover, it has been demonstrated that even though this type ofdata has been medically claimed to be insufficient to confidently diagnose the disease, adequateresults can be obtained. Across the report it has be shown that this can be achieved with intensivedata cleansing and the selection and design of the right model. Moving back to the research question presented at the beginning of this project: How could automatic learning techniques be used to improve the diagnosis of Alzheimer's disease? His answer can be summarized in the implementation of a deep neuronal network that will be basedon a correct selection of input data and proper design of the parameters and architecture of themodel, achieving a precise discrimination, 82.61% accurate, between patients with AD or not.

6. FUTURE PERSPECTIVES

What has been presented in this project is a proof of concept. There are further improvements needed to create a real-time classifier that could be used as tool to help the diagnosis of AD patients. Here, some suggestions are presented that could be beneficial to the future system: 1. The different influence of each of the features on the model accuracy can be analyzed, optimizing the data input. 2. A more precise parameter selection can be done, considering the analysis of extra hyper parameters as the learning rate. 3. In order to be closer to a real medical diagnosis procedure, the data that feeds the systemcan be complemented with MRI images or other bio-markers. 4. To increase the complexity of the solution, and to be able to process images as data input,a future implementation of a CNN should be considered. 5. In addition, the inclusion of data from different sources will improve the robustness of themodel. In this manner, a robust transfer learning can be achieved, in which the model will transfer not only his performance but also robustness from a source model to a target domain. 6. The business value of the system should be analyzed, evaluating different cases of use and adapting the solution to the final user needs. At the same time, aspects as the security, reliability or performance of the system must be considered as a priority in the case of a real use of the system with a medical purpose.

7.REFERENCES

[1] Dementia statistics - Alzheimer's diasease international. url: https://www.alz. co.uk/research/statistics.

[2] Christina Patterson. World Alzheimer Report 2018 The state of the art of dementia research: New frontiers. Tech. rep. Alzheimer's Disease International (ADI), London., 2018. doi: 10.1111/j.0033- 0124.1950.24{_}14.x. url: https://www.alz.co. uk/research/WorldAlzheimerReport2018.pdf?2.

[3] J. Manyika et al. Big data: The next frontier for innovation, competition and productivity. Tech. rep. May. 2011. url: https://www.mckinsey.com/~/media/McKinsey/ Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_exec_summary. ashx.

[4] Warren W. Barker et al. "Relative Frequencies of Alzheimer Disease, Lewy Body, Vascular and Frontotemporal Dementia, and Hippocampal Sclerosis in the State of Florida Brain Bank". In: Alzheimer Disease & Associated Disorders 16.4 (Oct. 2002), pp. 203–212. issn: 0893-0341. doi: 10.1097/00002093-200210000-00001. url: https://insights.ovid.com/crossref?an=00002093- 200210000-00001.

[5] Melonie Heron. "National Vital Statistics Reports Volume 67, Number 5 July 26, 2018". In: National Vital Statistics Reports 67.6 (2018), pp. 1–15. issn: 1551-8922. url: https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_06.pdf.

[6] Joseph Gaugler et al. "Alzheimer's Association Report". In: Alzheimer's and Dementia 11.3 (2015), pp. 332–384. issn: 15525279. doi: 10.1016/j.jalz.2016. 03.001. url: http://dx.doi.org/10.1016/j.jalz.2015.02.003%20https: //ac.elscdn.com/S1552526015000588/1-s2.0-S1552526015000588-main. pdf?_tid=d7c35d3a03bb-4b42-a93a-11fa491d3599&acdnat=1550351168_ f7ed67e075417bdf021d44889920839c.

[7] James C. Vickers et al. "The cause of neuronal degeneration in Alzheimer's disease". In: Progress in Neurobiology 60.2 (2000), pp. 139–165. issn: 03010082. doi: 10.1016/ 65 S0301- 0082(99)00023-4. url: https://ac.els-cdn.com/S0301008299000234/ 1-s2.0-S0301008299000234-main.pdf?_tid=7722ae13-f4d5- 451fa6e598ed91385196&acdnat=1550351534_af1b85fcb126558384dfef49b5240cbe.

[8] Richard J. O'Brien and Philip C. Wong. "Pathological Cascade". In: Archives of Medical Research 12.3 (2012), pp. 173–189. issn: 01406736. doi: 10.1146/annurevneuro-061010- 113613.Amyloid. url: http://dx.doi.org/10.1016/j.arcmed. 3 2012.11.008%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC3174086/ %5Cnhttp://dx.doi.org/10.1016/S0140-6736(12)62145-X.

[9] J. P. Brion. "Neurofibrillary tangles and Alzheimer's disease". In: European Neurology 40.3 (1998), pp. 130–140. issn: 00143022. doi: 10.1159/000007969.

[10] Clifford R. Jack et al. "Serial PIB and MRI in normal, mild cognitive impairment and Alzheimers disease: Implications for sequence of pathological events in Alzheimers disease". In: Brain 132.5 (2009), pp. 1355–1365. issn: 14602156. doi: 10.1093/brain/ awp062. url: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2677798/pdf/ awp062.pdf