HINDUSTHAN COLLEGE OF ENGINEERING AND TECHNOLOGY

EARLY DETECTION OF ALZHEIMER’S DISEASE WITH BLOOD PLASMA PROTEINS USING SUPPORT VECTOR MACHINES

PROJECT PHASE II

PROJECT REPORT

Submitted by,

ANEESH ABDUL RAHMAN	**19104017**
AKASH KUMAR	**19104011**
CHARUGNETHRA M	**19104039**
DHESIKA S	**19104047**

In partial fulfilment for the award of the degree

of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

HINDUSTHAN COLLEGE OF ENGINEERING AND TECHNOLOGY

Approved by AICTE, New Delhi, Accredited with ‘A’ Grade by NAAC

(An Autonomous Institution, Affiliated to Anna University, Chennai)

Valley Campus, Pollachi Highway, Coimbatore – 641 032

APRIL 2023

HINDUSTHAN COLLEGE OF ENGINEERING AND TECHNOLOGY

BONAFIDE CERTIFICATE

Certified that this project report "**EARLY DETECTION OF ALZHEIMER'S DISEASE WITH BLOOD PLASMA PROTEINS USING SUPPORT VECTOR MACHINES**" is the bonafide work of **ANEESH ABDUL RAHMAN, AKASH KUMAR, CHARUGNETHRA M, DHESIKA S** who carried out the project work under my supervision**.**

SUPERVISOR

Dr. S. Shankar M.E.,Ph.D.,
Professor & Head
Department of Computer
Science and Engineering,
Hindusthan college of Engineering
& Technology, Coimbatore-641 032

HEAD OF DEPARTMENT

Dr. S. Shankar M.E., Ph.D.,
Professor & Head
Department of Computer
Science and Engineering,
Hindusthan college of Engineering
& Technology, Coimbatore-641 032

Submitted for the autonomous project phase-II examination held on …………

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION

We, hereby jointly declare that the project work entitled **"EARLY DETECTION OF ALZHEIMER'S DISEASE WITH BLOOD PLASMA PROTEINS USING SUPPORT VECTOR MACHINES",** submitted to the Project Viva voce - April 2023 in partial fulfillment for the award of the degree of "**BACHELOR OF ENGINEERING IN COMPUTER SCIENCE AND ENGINEERING"**, is the report of the original project work done by us under the guidance of **Dr.S.Shankar M.E.,Ph.D.,** Professor & Head, Department of Computer Science and Engineering, Hindusthan College of Engineering and Technology, Coimbatore.

NAME	SIGNATURE
ANEESH ABDUL RAHMAN	
AKASH KUMAR	
CHARUGNETHRA M	
DHESIKA S	

I certify that the declarations made by the above candidates are true.

Project Guide

Dr.S.Shankar M.E., PhD.,

Professor & Head,

Department of CSE,

Hindusthan College of Engineering

And Technology, Coimbatore-32

ACKNOWLEDGEMENT

We take this opportunity to express our whole hearted thanks and our profound respect to all those who guided and inspired us in the completion of this project work.

We extend our sincerest thanks to the Managing Trustee of Hindusthan Educational and Charitable Trust Thiru.T.S.R.Khannaiyann and the Secretary Smt. S.R.Sarasuwathi Khannaiyann and the joint Secretary Smt.Priya Satish Prabu for providing essential infrastructure.

We would like to reveal our profound thanks to our respected CEO Dr.K.Karunakaran and our respected Principal Dr.J.Jaya, M.E., ph.D., who happens to be striving force in all endeavors.

We would like to express our gratitude to the Head of the Department Dr. S.Shankar, M.E., Ph.D., for allowing us to do this project.

We would like to express our sincere thanks and deep sense of gratitude to our guide Dr.S.Shankar M.E., Ph.D., Head of The Department, Computer Science and Engineering, for his valuable guidance ,suggestions and constant encouragement which paved way for the successful completion of the project work.

We express our immense pleasure and thankfulness to our Class Advisor and Project Guide, Dr.S.Shankar, M.E.,Ph.D., Professor & Head and all other Faculty members of the Department of Computer Science and engineering, technical staff and friends who stood by us and helped in the successful completion of this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	**Alzheimer’s Disease**
ADRD	**Alzheimer’s disease and related dementias**
AI	**Artificial Intelligence**
DL	**Deep Learning**
MCI	**Mild cognitive impairment**
DLB	**Dementia with Lewy bodies**
FTD	**Frontal temporal dementia**
MRI	**Magnetic Resonance Imaging**
CNN	**Convolutional Neural Network**
DEMNET	**Dementia Network**
AUC	**Area under curve**
ADNI	**Alzheimer’s disease Neuroimaging Initiative**
HMS	**Hilbert marginal spectrum**
WiGMM	**Warped infinite Gaussian mixture model**
EEG	**Electroencephalogram**
GDS	**Global Deterioration Scale**
OR	**Odds Ratio**
CI	**Confidence interval**
NC	**Normal control**
SVM	**Support Vector Machine**
FLOSS	**Free/Libre and Open source software**
GUI	**Graphical User Interface**

ABSTRACT

Alzheimer's is a type of dementia that causes problems with memory, thinking and behaviour. Symptoms usually develop slowly and get worse over time, becoming severe enough to interfere with daily tasks. Dementia is not a specific disease. It’s an overall term that describes a group of symptoms associated with a decline in memory or other thinking skills severe enough to reduce a person’s ability to perform everyday activities. Alzheimer’s disease accounts for 60 to 80 percent of cases. Vascular dementia, which occurs after a stroke, is the second most common dementia type. But there are many other conditions that can cause symptoms of dementia, including some that are reversible, such as thyroid problems and vitamin deficiencies. Dementia is a general term for loss of memory and other mental abilities severe enough to interfere with daily life. It is caused by physical changes in the brain. Alzheimer’s is the most common type of dementia, but there are many kinds. The input data is taken from the dataset repository. In our process, we are take the Alzheimer’s disease dataset as input. The system is developed the machine learning algorithm such as Support vector machine and logistic regression. Our results show that the performance metrics such as accuracy, sensitivity, specificity of our algorithms were high , suggesting that they can be used as reliable tool for the early diagnosis of Alzheimer’s disease. This study highlights the potential of machine learning algorithms in improving the accuracy and efficiency of the diagnosis of Alzheimer’s disease , which can ultimately lead to better treatment and care for patients.

CHAPTER 1
INTRODUCTION

1.1 General Introduction:

ALZHEIMER's disease (AD) is the leading cause of dementia and poses a significant social and economic challenge. It is responsible for more than half of all cases of dementia. Over 50 million individuals currently suffer from dementia worldwide with a projected increase to 152 million by 2050. No cure for AD has been discovered, but there is intense effort to develop new clinical interventions that may slow or halt the disease. Such interventions are aimed at early (including preclinical and prodromal) stages of the disease prior to extensive cell damage, when it is thought treatment is more likely to be effective. Alzheimer's disease and related dementias (ADRD) have become a major public health concern in the United States. An estimated 5.6 million Americans aged 65 and older (10% of the US population) were living with ADRD in 2019, and this number is expected to grow dramatically as the population continues to age. By 2025, the number of Americans aged 65 or older with ADRD is expected to reach 7.1 million, nearly a 27% increase from 2019, and by 2050, this population is projected to be 13.8 million, with the highest growth among those in ADRD's advanced stage. Persons with ADRD require progressively extensive assistance in their daily lives, the majority of which is provided by family members, friends, and other unpaid caregivers. It is estimated that in 2018, American caregivers of persons with ADRD provided 18.5 billion hours of informal unpaid assistance, valued at $233.9 billion. Family caregivers (hereafter "caregivers") of persons with ADRD are expected to make important care decisions for their family members with ADRD on a daily basis. However, these caregivers report being unprepared for their roles and responsibilities, uninformed about care options, and unsupported by professionals in their decision making. Caregiving for persons with ADRD is stressful, and it can severely affect the caregiver's own health and well-being. There is an urgent need to better prepare caregivers to manage their daily lives and those of their family members with ADRD, yet there are critical knowledge gaps regarding the types and amounts of information that caregivers may want to have in order to better manage ADRD. To provide patient-centered care for people with ADRD and enhance caregivers' quality of life, we must address those gaps. Artificial intelligence (AI) is showing great promise in areas of health care—in precision treatments, patient education, virtual assistance, and cost reduction. Some attempts have been made to apply AI for persons with ADRD and their caregivers in order to improve patients' daily functioning, quality of life, and well-being, as well as reduce caregiver burden (e.g., social robots to facilitate social interaction and engagement, assistive robots to facilitate daily activities such as hand washing, tea making, or dressing). To date, however, there has been little systematic review to identify research on AI for ADRD management by caregivers and gaps that remain in our understanding of AI for ADRD management. We have conducted this systematic review to identify and examine literature on AI that provides information to facilitate ADRD management by caregivers of individuals diagnosed with ADRD and to identify gaps in the literature that suggest future directions for research. The accurate diagnosis of Alzheimer's disease (AD) plays an important role in patient treatment, especially at the disease's early stages, because risk awareness allows the patients to undergo preventive measures even before the occurrence of irreversible brain damage. Although many recent studies have used computers to diagnose AD, most machine detection methods are limited by congenital observations. AD can be diagnosed-but not predicted-at its early stages, as prediction is only applicable before the disease manifests itself. Deep Learning (DL) has become a common technique for the early diagnosis of AD. AD is the most common stages of dementia that requires extensive medical care. For initiation of clinical progress and efficient patient treatment, early and precise analysis of AD prediction is necessary. AD is a chronic, neurobiological brain disorder that steadily kills brain cells induces memory and thinking capacity deficits, and eventually accelerates the loss of ability to perform even the most basic tasks. In the early stages of AD, doctors use neuroimaging and computer-aided diagnostic approaches to classify the disease. A summary of the most recent census by the World Alzheimer's Association reports that over 4.7 million individuals aged over 65 years have survived this disease in the United States. In the next fifty years, they estimated 60 million people may be affected by AD. Of all forms of dementia globally, Alzheimer's disease accounts for around 60-80%. Every three seconds, one person affected by dementia out of it, 60% is due to AD. Dementia with Alzheimer's is approximately divided into the following: - Mild Cognitive Impairment: Commonly affected by lack of memory to many individuals as they become older, whereas, for others, it leads to the problem of dementia. - Mild Dementia: Cognitive impairments that sometimes affect their daily lives are encountered by people with moderate dementia. Symptoms include lack of memory, uncertainty, changes in personality, being lost, and difficulties in executing routine tasks. - Moderate Dementia: The everyday lifestyle becomes much complex, where the patient requires extra care and support. Symptoms are equivalent to mild yet elevated dementia. People may need more help even to comb their hair. They can also exhibit significant personality changes; for example, they become paranoid or irritated for no reason. Sleep disorders are likely to occur as well. - Severe Dementia: The symptoms may become deteriorated during this stage. These patients may lack the capacity to communicate, and full-time treatment may be required for the person. One's bladder control may be lost, and even small activities are impossible for them to perform actions like keeping their head up in a normal position and sitting in a chair. Early detection of this disorder is being researched to slow down the abnormal degeneration of the brain, reduce medical care cost reduction, and ensure improved treatment. The recent failures in Alzheimer's disease research studies may suggest that early intervention and diagnosis could be crucial to the effectiveness of treatment. A wide variety of neuroimaging methods are becoming increasingly dependent on the diagnosis of dementia, and this is reflected in many new diagnostic criteria. Neuroimaging increases diagnosis accuracy for various subtypes of dementia using machine learning. Specific pre-processing steps are needed to implement machine learning algorithms. Extraction and selection of features, reduction of feature dimensionality and classifier algorithm are all phases of the machine learning-based classification process. Such techniques need advanced knowledge and several optimization steps, which can be time-consuming. The history of AD, as presented in this section, is consolidation of finding from AD publications searched in Google Scholar. Only the latest publications were considered, and only the papers published between 2008 and 2019 were selected. Our research focused on datasets used to examine AD and mild cognitive impairment MCI), the forerunner of AD. The processes and techniques used by previous researchers were studied. AD is currently ranked as the sixth leading cause of death in the US. Recent estimates indicate also that the disorder may even rank third (after heart disease and cancer) as the leading cause of the death for elderly. Clearly, predicting the progression of AD at its early stages and preventing the disease from progressing are of great importance. The diagnosis of AD requires various medical tests and enormous multivariate heterogeneous data. However, manual comparison, visualisation, and analysis of data are difficult and tedious due to the heterogeneous nature of medical tests.

1.2 Objectives:

The main objective of our project is,

- To predict or detect the Alzheimer’s dementia disease.
- To implement the machine learning algorithm.
- To enhance the overall performance analysis.

1.3 Problem Statement:

Memory loss is often the first and main symptom in early Alzheimer's disease. It is also seen, although less often, in early vascular dementia and dementia with Lewy bodies (DLB). Memory loss is not common in early front temporal dementia (FTD).it is difficult to detect the Alzheimer’s disease dementia.so to overcome this problem we are using machine learning algorithms like support vector machines and logistic regression.

CHAPTER 2
SYSTEM PROPOSAL

2.1 EXISTING SYSTEM:

Our method is mainly based on machine learning (ML) techniques (support vector machines in particular) because of their ability to create multivariable models by learning patterns from complex data. Using novel feature selection and evaluation modalities, we identified 5 novel panels of non-amyloid proteins with the potential to serve as biomarkers of early AD. In particular, we found that the combination of A2M, ApoE, BNP, Eot3, RAGE and SGOT may be a key biomarker profile of early disease. Disease detection models based on the identified panels achieved sensitivity (SN) > 80%, specificity (SP) > 70%, and area under receiver operating curve (AUC) of at least 0.80 at prodromal stage (with higher performance at later stages) of the disease. Existing ML models performed poorly in comparison at this stage of the disease, suggesting that the underlying protein panels may not be suitable for early disease detection. Our results demonstrate the feasibility of early detection of AD using non-amyloid based biomarkers.

2.1.1 DISADVANTAGES:

- The results is low when compared with proposed algorithm
- It is not efficient for large volumes of data.
- Theoretical limits.

2.2 PROPOSED SYSTEM:

In this system, the Alzheimer’s dataset was taken as input. The input data was taken from the dataset repository. Then, we have to implement the data preprocessing step.in this step, we have to handle the missing values to avoid wrong prediction. If any missing values are present in our input data, we have to replace the missing values by zero or null values. Then, we have to use label encoding, to encode the label for input data. To encode the columns into numeric values. Next, we have to implement the data splitting. In this step, we have to split the data into test and train. Then, we have to implement the machine learning algorithms such as Support Vector Machine (SVM) and Logistic regression (LR). Finally, the experimental results shows that the performance metrics such as accuracy, precision, recall, sensitivity and confusion matrix. Our experimental results show that both SVM and LR algorithms performed well in classifying individuals with Alzheimer’s disease. The performance metrics, such as accuracy ,precision, recall, sensitivity and confusion matrix , indicated that the models were effective in predicting Alzheimer’s disease. Additionally, we conducted comparative analysis of the performance metrics of SVM and LR algorithms, and found that SVM outperformed LR in terms of accuracy and sensitivity. Overall , our study highlights the potential of machine learning algorithms in improving the diagnosis of Alzheimer’s disease and providing better treatment and care for patients.

2.2.1 ADVANTAGES:

- It is efficient for large number of datasets.
- The experimental result is high when compared with the existing system.
- To increase the performance metrics of results.

2.3 LITERATURE SURVEY:

2.3.1 Predicting Prodromal Dementia Using Linguistic Patterns and Deficits 2020

Author: ahmed h. alkenani 1, 2, yuefeng li 1, yue xu 1 and qing zhang2

Published in: IEEE Access (Volume: 8)

Date of Publication: 09 October 2020

Electronic ISSN: 2169-3536

INSPEC Accession Number: 20129116

DOI: 10.1109/ACCESS.2020.3029907

***Methodology*:**

Data was derived from the cookie theft picture corpus of Dementia Bank, from which all language samples of the identified aetiologies were used, with a random subsampling technique that handles the skewness of the classes. Several original lexical and syntactic (i.e., lexicosyntactic) features were introduced and used alongside previously established lexicosyntactics to train machine learning (ML) classifiers against these aetiologies'. Further, a statistical analysis was conducted to uncover the deficiency across these aetiologies'. Our models resulted in benchmarks for differentiating all the identified classes with accuracies ranging between 95 to 98% and corresponding F1 values falling between 94 and 98%. The statistical analysis of our lexicosyntactic biomarkers shows that linguistic deviations are associated with prodromal as well as advanced neurodegenerative pathologies, being greatly impacted as cognitive decline increases and suggesting that language biomarkers may aid the early diagnosis of these pathologies.

***Advantage*:**

- The advantage of n-grams features as being easily computed without requiring manual annotation, which suggests that our models could be extended to other clinically recommended pictures for the same purpose.
- It introduces original lexicosyntactic features and investigates their representations, in conjunction with other well-known lexicosyntactics, across different dementia etiologies.

2.3.2 DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia from MR Images, 2021

***Author*:** Suriya murugan 1, chandran venkatesan 2 , m. g. sumithra 2 , (senior member, ieee), xiao-zhi gao 3 , b. elakkiya 4 , m. akila 5 , and s. Manoharan

Published in: IEEE Access (Volume: 9)

Page(s): 90319 - 90329

Date of Publication: 18 June 2021

Electronic ISSN: 2169-3536

INSPEC Accession Number: 21082758

DOI: 10.1109/ACCESS.2021.3090474

***Methodology*:**

Alzheimer's disease (AD) is the most common cause of dementia globally. It steadily worsens from mild to severe, impairing one's ability to complete any work without assistance. It begins to outstrip due to the population ages and diagnosis timeline. For classifying cases, existing approaches incorporate medical history, neuropsychological testing, and Magnetic Resonance Imaging (MRI), but efficient procedures remain inconsistent due to lack of sensitivity and precision. The

Convolutional Neural Network (CNN) is utilized to create a framework that can be used to detect specific Alzheimer's disease characteristics from MRI images. By considering four stages of dementia and conducting a particular diagnosis, the proposed model generates high-resolution disease probability maps from the local brain structure to a multilayer perceptron and provides accurate, intuitive visualizations of individual Alzheimer's disease risk. To avoid the problem of class imbalance, the samples should be evenly distributed among the classes. The obtained MRI image dataset from Kaggle has a major class imbalance problem. A DEMentia NETwork (DEMNET) is proposed to detect the dementia stages from MRI. The DEMNET achieves an accuracy of 95.23%, Area under Curve (AUC) of 97% and Cohen's Kappa value of 0.93 from the Kaggle dataset, which is superior to existing methods. We also used the Alzheimer's disease Neuroimaging Initiative (ADNI) dataset to predict AD classes in order to assess the efficacy of the proposed model.

Advantage:

- The high model parameter and class imbalance in the multiclass AD classification is still an issue.

2.3.3 Alzheimer’s Diseases Detection by Using Deep Learning Algorithms: A Mini-Review, 2020

***Author*:** suhad al-shoukry 1,2, taha h. rassem 1 , (senior member, ieee), and nasrin m. makbo

Published in: IEEE Access (Volume: 8)

Page(s): 77131 - 77141

Date of Publication: 21 April 2020

Electronic ISSN: 2169-3536

INSPEC Accession Number: 19581889

DOI: 10.1109/ACCESS.2020.2989396

***Methodology*:**

The accurate diagnosis of Alzheimer’s disease (AD) plays an important role in patient treatment, especially at the disease’s early stages, because risk awareness allows the patients to undergo preventive measures even before the occurrence of irreversible brain damage. Although many recent studies have used computers to diagnose AD, most machine detection methods are limited by congenital observations. AD can be diagnosed-but not predicted-at its early stages, as prediction is only applicable before the disease manifests itself. Deep Learning (DL) has become a common technique for the early diagnosis of AD. Here, we briefly review some of the important literature on AD and explore how DL can help researchers diagnose the disease at its early stages. From a computational perspective, this recent advancement has spawned the development of tools that incorporate several patient-specific observations into predictions and improve the clinical outcomes of patients suffering from such disorders.

Advantage:

- No expertise was required, as no image segmentation was involved in preprocessing the data. This feature generally serves as the advantage of this approach over the other methods.

2.3.4 Artificial Intelligence for Caregivers of Persons with Alzheimer’s disease and Related Dementias: Systematic Literature Review, 2020

Author: Bo xie,Cui Tao,JuanLi,Robin C Hilsabeck

Published on 20.8.2020 **in Vol 8 , No 8 (2020) :August**

Publisher: JMIR Publications

ISSN: 2291-9694

***Methodology*:**

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for conducting systematic literature reviews, during August and September 2019, we performed 3 rounds of selection. First, we searched predetermined keywords in PubMed, Cumulative Index to Nursing and Allied Health Literature Plus with Full Text, PsycINFO, IEEE Xplore Digital Library, and the ACM Digital Library. This step generated 113 nonduplicate results. Next, we screened the titles and abstracts of the 113 papers according to inclusion and exclusion criteria, after which 52 papers were excluded and 61 remained. Finally, we screened the full text of the remaining papers to ensure that they met the inclusion or exclusion criteria; 31 papers were excluded, leaving a final sample of 30 papers for analysis.

Advantage:

- To identify and examine literature on AI that provides information to facilitate ADRD management by caregivers of individuals diagnosed with ADRD and identify gaps in the literature that suggest future directions for research.

2.3.5 Detecting Alzheimer's Dementia Degree, 2020

Author: Edmond Q. Wu, Xian-Yong Peng, Sheng-Di Chen, Xiao-Yan Zhao and Zhi-Ri Tang

Page(s): 116 - 125

Date of Publication: 12 August 2020

ISSN Information:

Print ISSN: 2379-8920

Electronic ISSN: 2379-8939

INSPEC Accession Number: 21682058

DOI: 10.1109/TCDS.2020.3015131

Publisher: IEEE

Methodology:

The diagnosis of Alzheimer's disease (AD) faces two important issues. They are how to extract the features of the rhythms of patients with AD, and how to label them and reveal the degree of dementia in patients. This study defines 14 instantaneous power indicators of dementia judgment through Hilbert marginal spectrum (HMS) from rhythm waves. A warped infinite Gaussian mixture model (WiGMM) is proposed to learn the latent variables of these indicators to detect the degree of dementia. The experimental results show that HMSbased indicators are able to reflect the cognitive function of AD patients. This proposed method has the ability to detect brain cognitive status through a warped transform and Dirichlet process parameter prior inference. According to the cholinergic injury theory of the pathogenesis of AD, the slowing of electroencephalogram (EEG) signals in AD patients is associated with loss of cholinergic neurons in the basal ganglia, hippocampus, and neocortex .The characteristic pathological change that AD first appeared was the neurofibrillary tangles of the temporal lobe in brain**.**

Advantage:

- The warped infinite Gaussian mixture model can easily capture local information and provide higher resolution.
- The power of θ wave in AD group is higher in the frontal lobe than that in control group.

2.3.6 Depression as a Risk Factor for Dementia and Alzheimer's disease, 2020

Author**:** Vanesa Cantón-Habas, Manuel Rich-Ruiz, Manuel Romero-Saldaña

Biomedicines **2020**, *8*(11), 457; **https://doi.org/10.3390/biomedicines8110457**

Received: 24 September 2020 / Accepted: 26 October 2020 / Published: 28 October 2020

(This article belongs to the Special Issue **Crosstalk between Depression, Anxiety, and Dementia: Comorbidity in Behavioral Neurology and Neuropsychiatry**)

Methodology:

Preventing the onset of dementia and Alzheimer's disease (AD), improving the diagnosis, and slowing the progression of these diseases remain a challenge. The aim of this study was to elucidate the association between depression and dementia/AD and to identify possible relationships between these diseases and different sociodemographic and clinical features. In this regard, a case-control study was conducted in Spain in 2018–2019. The definition of a case was: A person ≥ 65 years old with dementia and/or AD and a score of 5–7 on the Global Deterioration Scale (GDS). The sample consisted of 125 controls; among the cases, 96 had dementia and 74 had AD. The predictor variables were depression, dyslipidemia, type 2 diabetes mellitus, and hypertension. The results showed that depression, diabetes mellitus, and older age were associated with an increased likelihood of developing AD, with an Odds Ratio (OR) of 12.9 (95% confidence interval (CI): 4.3–39.9), 2.8 (95% CI: 1.1–7.1) and 1.15 (95% CI: 1.1–1.2), respectively. Those subjects with treated dyslipidaemia were less likely to develop AD (OR 0.47, 95% CI: 0.22–1.1). Therefore, depression and diabetes mellitus increase the risk of dementia, whereas treated dyslipidaemia has been shown to reduce this risk.

Disadvantage:

- The difficulty in identifying all articles that are related to this study: This problem is identified and was considered to be a key problem of SLR.

2.3.7 Automatic detection of linguistic indicators as a means of early detection of Alzheimer's disease and of related dementias: A computational linguistics analysis, 2017

Author**:** Eva Danasi, Dimitra Arfani, Katerina Fragkopoulou, Spyridoula Varlokosta

Published in: 2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)

Date of Conference: 11-14 September 2017

Date Added to IEEE *Xplore*: 25 January 2018

ISBN Information:

Electronic ISBN:978-1-5386-1264-4

Print on Demand(PoD) ISBN:978-1-5386-1265-1

INSPEC Accession Number: 17525390

DOI: 10.1109/CogInfoCom.2017.8268212

Publisher: IEEE

Methodology:

In the present study, we analyzed written samples obtained from Greek native speakers diagnosed with Alzheimer’s in mild and moderate stages and from age matched cognitively normal controls (NC). We adopted a computational approach for the comparison of morph syntactic complexity and lexical variety in the samples. We used text classification approaches to assign the samples to one of the two groups. The classifiers were tested using various features: morph-syntactic and lexical characteristics. Degenerative conditions, such as Alzheimer’s disease (henceforth AD) are commonly associated with deficits across a range of subcomponents of linguistic competence. Although both AD and other types of dementia are associated with changes in spoken and written language, these changes have not been extensively examined or compared. Memory impairment implies that the vocabulary of patients with dementia is poorer and simpler than that of healthy subjects and more incoherent. Language expression shows that neuronal brain activity in the area of Broca and Wernicke is reduced, so that words often do not make sense and complement the information lost with the damaged regions neuronal cells the proposed method excels in discerning AD patients in mild and moderate stages from NC leading to the in-depth understanding of language deficits.

Advantage:

- A related index is Brunet’s W, lower values of which imply a higher number of distinct word types, and thus a richer vocabulary.
- The high accuracies achieved in both comparisons imply that the classifiers’ performance was high in all the 10 fold classifications tasks.

CHAPTER 3

SYSTEM DIAGRAMS

3.1 SYSTEM ARCHITECTURE:

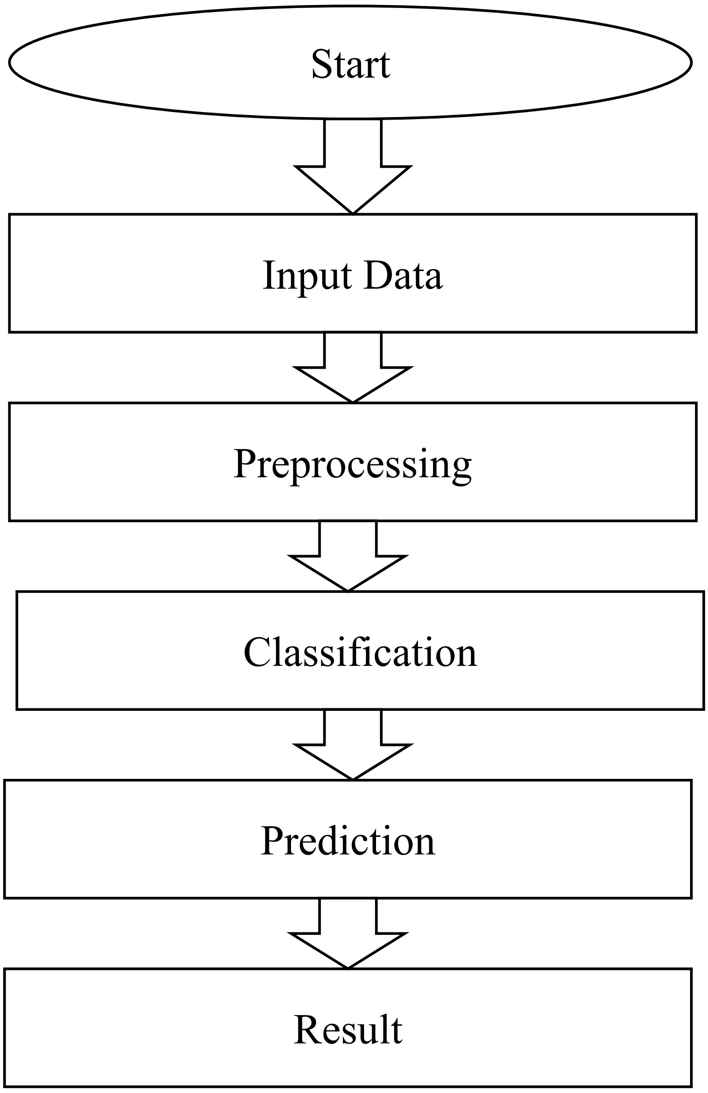


FIGURE 3.1: SYSTEM ARCHITECTURE

3.2 FLOW DIAGRAM

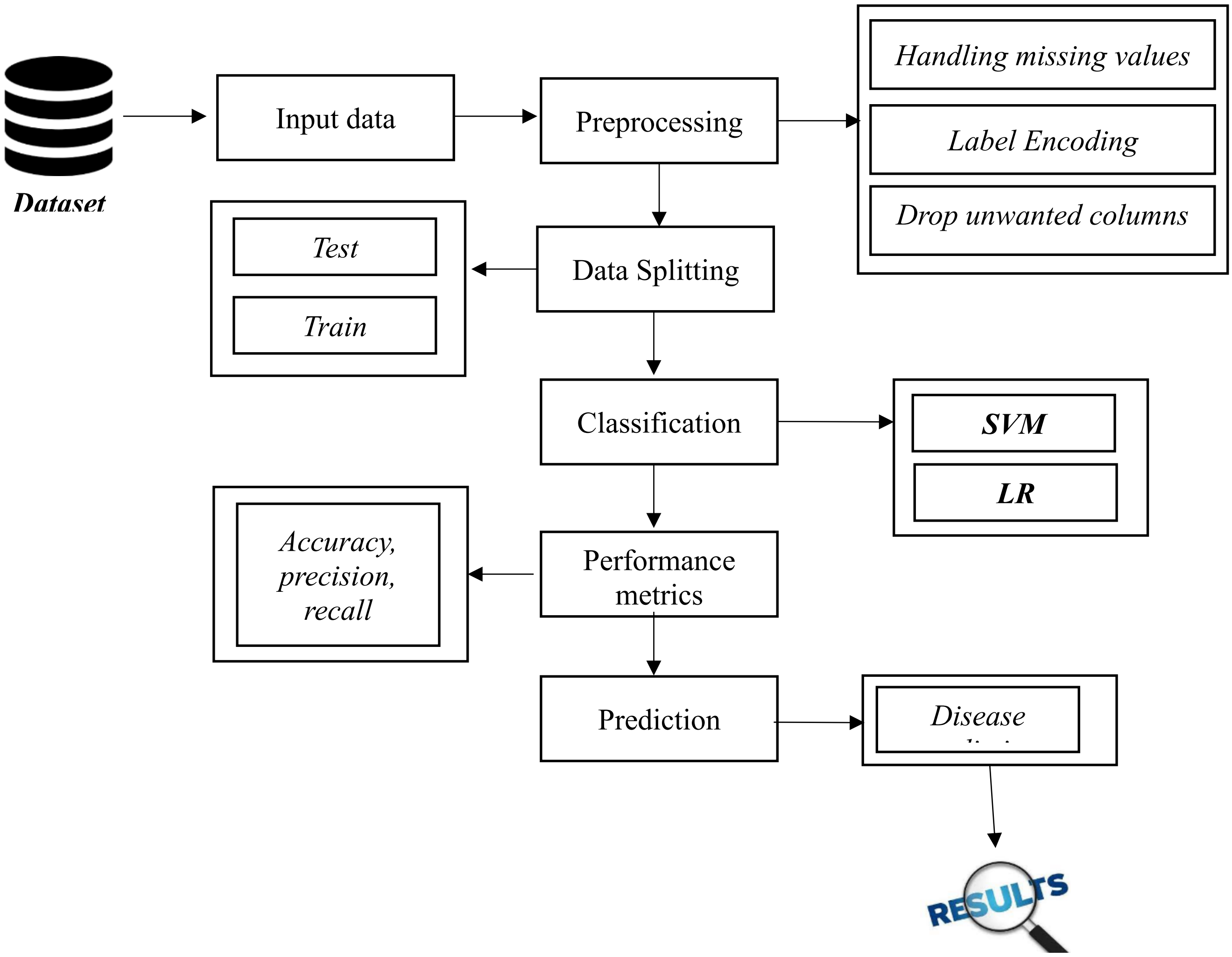


FIGURE 3.2: FLOW DIAGRAM

3.3 UML DIAGRAMS:

3.3.1 USE CASE DIAGRAM:

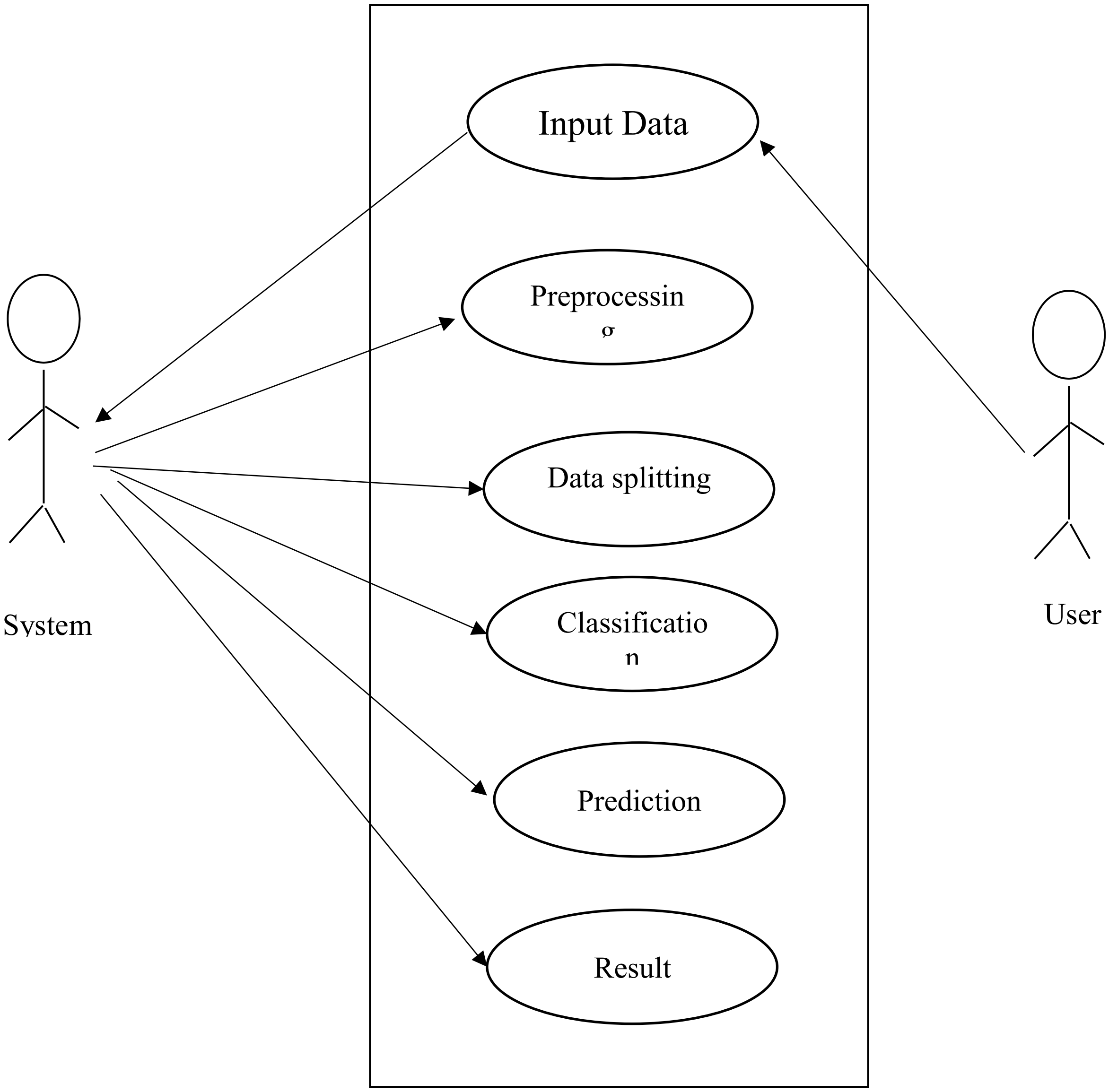


FIGURE 3.3.1: USE CASE DIAGRAM

3.3.2 ACTIVITY DIAGRAM:

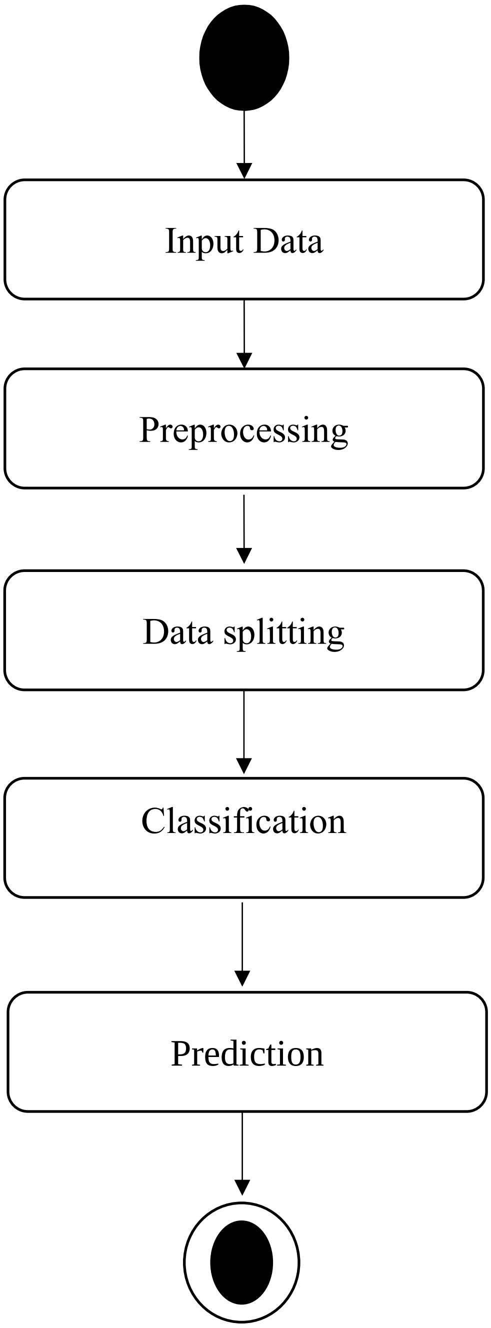


FIGURE 3.3.2: ACTIVITY DIAGRAM

3.3.3 SEQUENCE DIAGRAM:

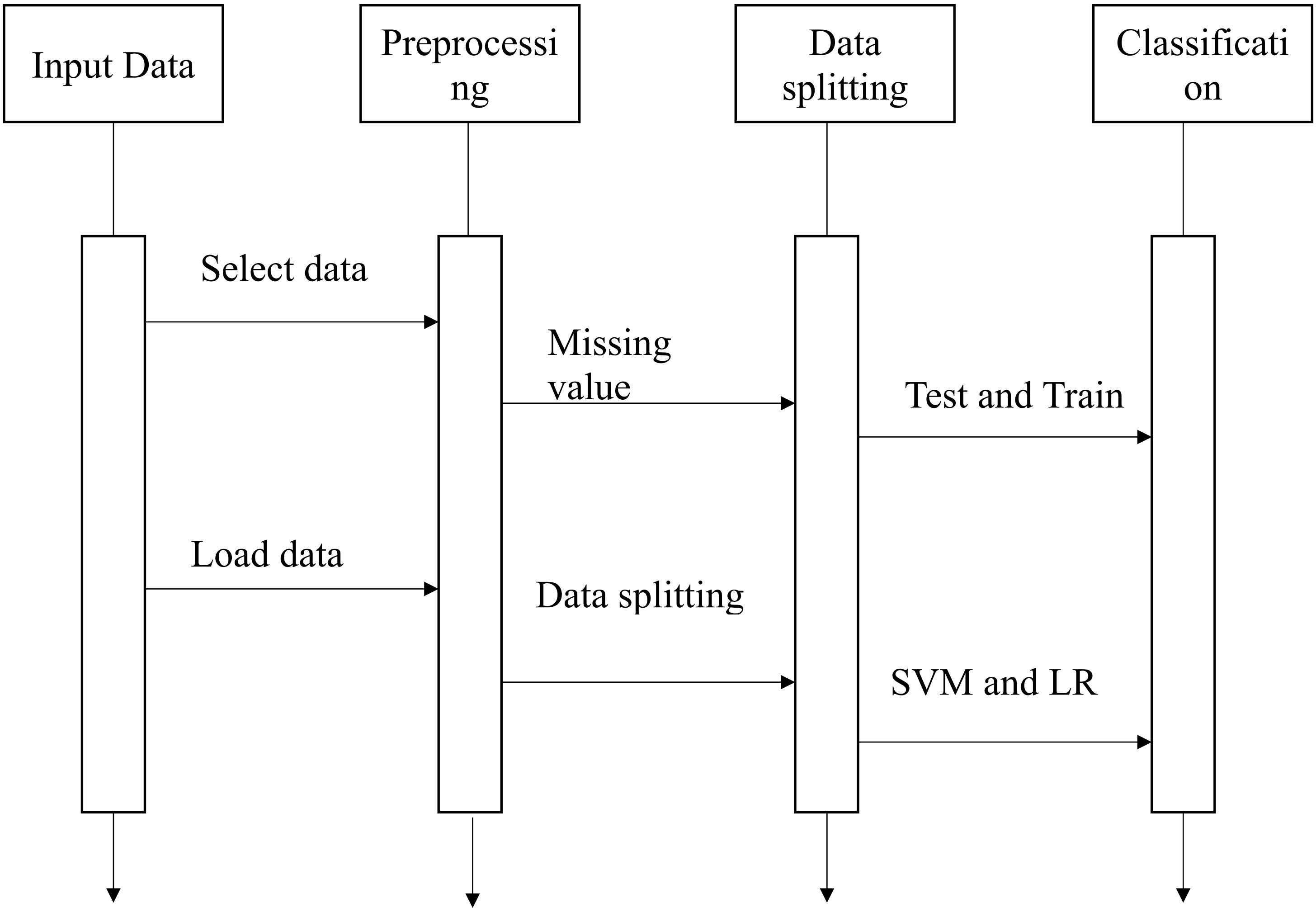


FIGURE 3.3.3: SEQUENCE DIAGRAM

3.3.4 ER DIAGRAM:

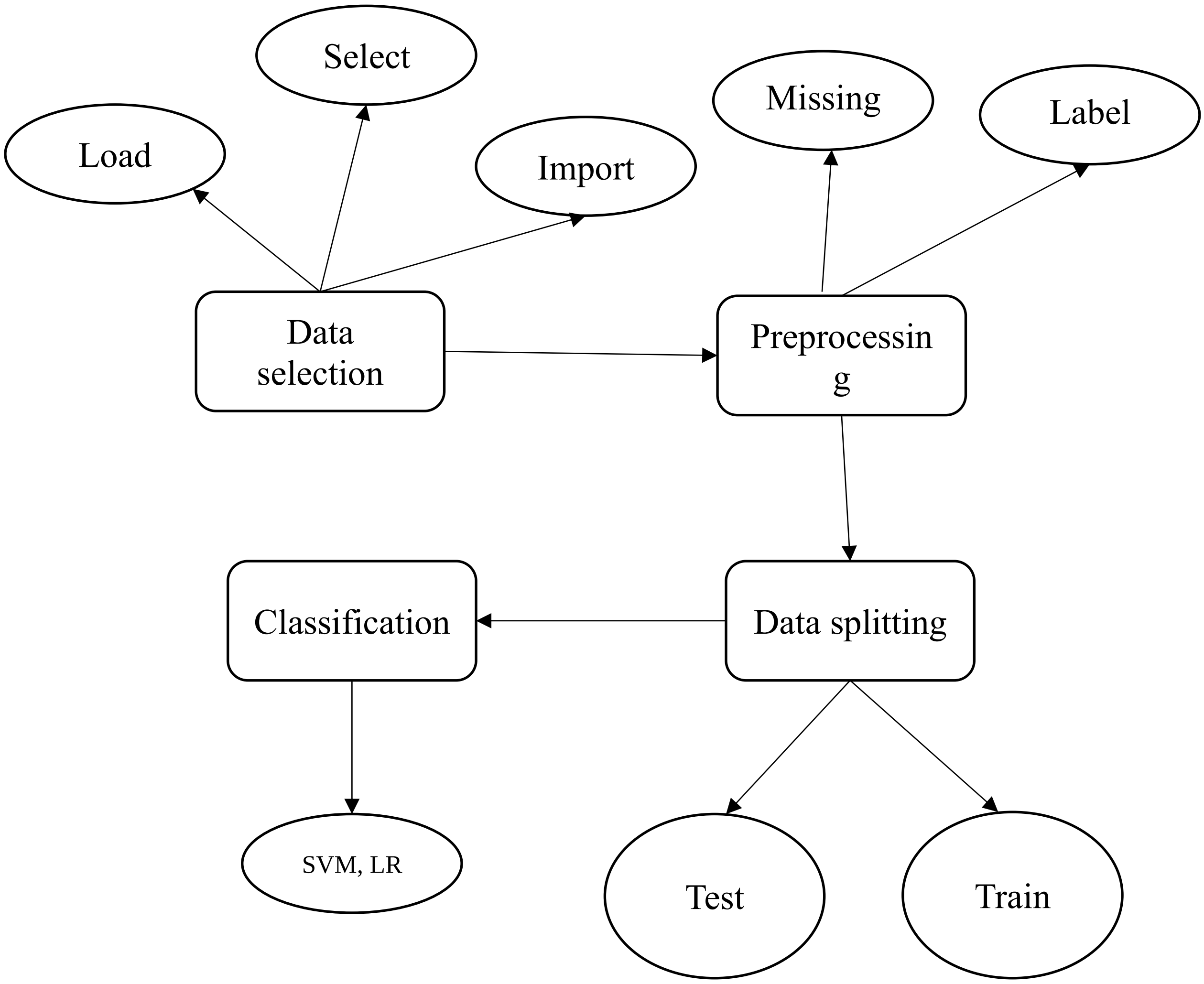


FIGURE 3.3.4: ER DIAGRAM

3.3.6 CLASS DIAGRAM:

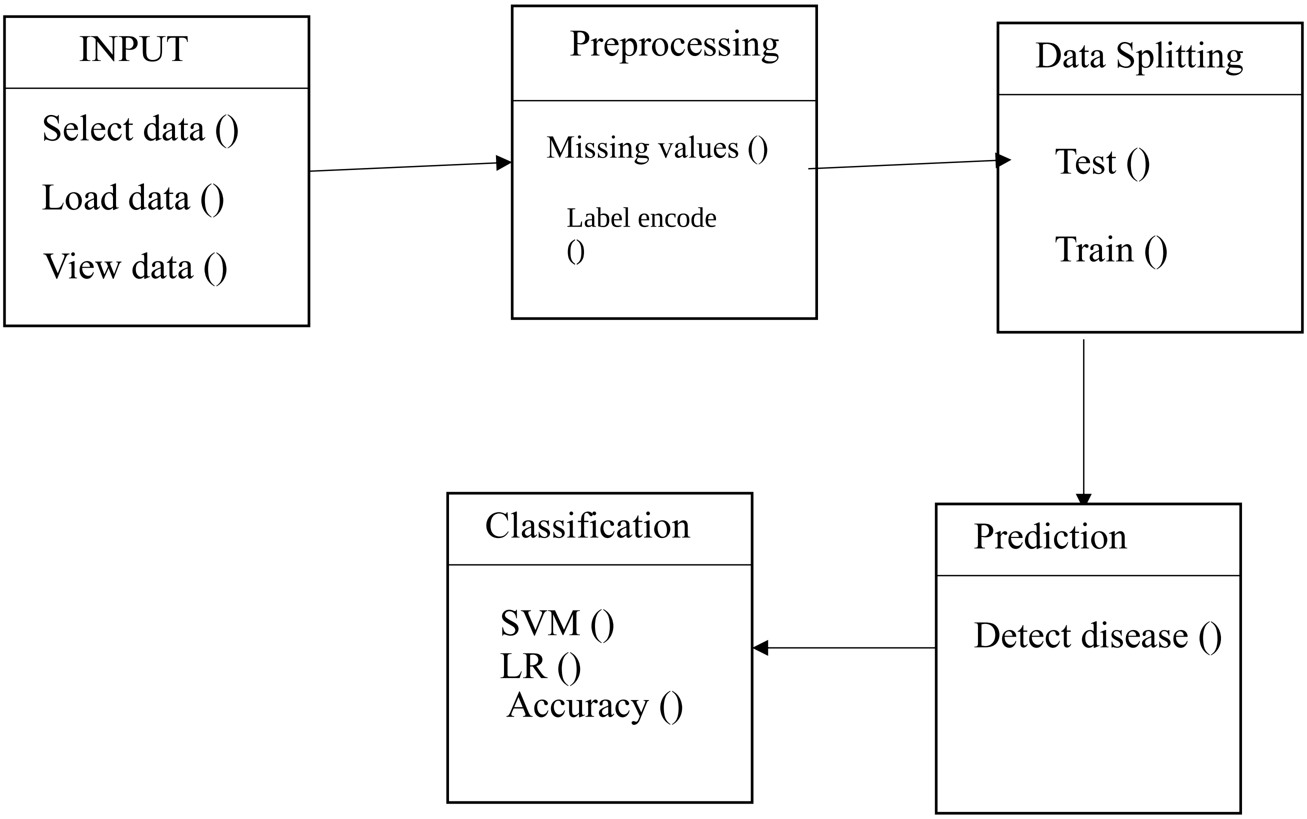


FIGURE 3.3.5: CLASS DIAGRAM

CHAPTER 4
IMPLEMENTATION

4.1 MODULES:

- Data selection
- Data preprocessing
- Data splitting
- Classification
- Result Generation

4.2 MODULES DESCRIPTION:

4.2.1: DATA SELECTION:

- The input data was collected from dataset repository.
- In our process, the Alzheimer’s disease dataset is used.
- This set consists of a longitudinal collection of 150 subjects aged 60 to 96.
- Each subject was scanned on two or more visits, separated by at least one year for a total of 373 imaging sessions.
- For each subject, 3 or 4 individual T1-weighted MRI scans obtained in single scan sessions are included.
- The subjects are all right-handed and include both men and women.
- 72 of the subjects were characterized as non-demented throughout the study.
- 64 of the included subjects were characterized as demented at the time of their initial visits and remained so for subsequent scans, including 51 individuals with mild to moderate Alzheimer’s disease.
- Another 14 subjects were characterized as non-demented at the time of their initial visit and were subsequently characterized as demented at a later visit.

4.2.2: DATA PREPROCESSING:

- Data pre-processing is the process of removing the unwanted data from the dataset.
- Pre-processing data transformation operations are used to transform the dataset into a structure suitable for machine learning.
- This step also includes cleaning the dataset by removing irrelevant or corrupted data that can affect the accuracy of the dataset, which makes it more efficient.
- Missing data removal
- Encoding Categorical data
- Missing data removal: In this process, the null values such as missing values and Nan values are replaced by 0.
- Missing and duplicate values were removed and data was cleaned of any abnormalities.
- Encoding Categorical data: That categorical data is defined as variables with a finite set of label values.
- That most machine learning algorithms require numerical input and output variables.
- In our process, we have to remove the missing values from our input dataset attributes such as MMSE and SES.

4.2.3: DATA SPLITTING:

- During the machine learning process, data are needed so that learning can take place.
- In addition to the data required for training, test data are needed to evaluate the performance of the algorithm in order to see how well it works.

- In our process, we considered 70% of the Alzheimer’s disease dataset to be the training data and the remaining 30% to be the testing data.
- Data splitting is the act of partitioning available data into two portions, usually for cross-validator purposes.
- One Portion of the data is used to develop a predictive model and the other to evaluate the model's performance.
- Separating data into training and testing sets is an important part of evaluating data mining models.
- Typically, when you separate a data set into a training set and testing set, most of the data is used for training, and a smaller portion of the data is used for testing.

4.2.4: CLASSIFICATION:

- In machine learning, classification refers to a predictive modelling problem where a class label is predicted for a given example of input data.
- Classification is the task of predicting a discrete class label. Regression is the task of predicting a continuous quantity.
- In machine learning, classification is a supervised learning concept which basically categorizes a set of data into classes.
- Before classification, we should have split the data into test and train.
- Most of data’s are used for training and smaller portion of the data’s are used for testing.
- Training data is used for evaluate the model and testing data is used for predictive the model.
- After data splitting, we have to implement the classification algorithm.
- In our process, we have to use, support vector machine (SVM).
- SVM It is basically a representation of different classes in a hyper plane in multidimensional space.

- The hyper plane will be generated in an iterative manner by SVM so that the error can be minimized the goal of SVM is to divide the datasets into classes to find a maximum marginal hyper plane.
- Support Vector Machine" (SVM) is a supervised machine learning algorithm that can be used for both classification and regression challenges.
- Support Vectors are simply the coordinates of individual observation. The SVM classifier is a frontier that best segregates the two classes (hyper-plane/ line).
- The choice of SVM for the model development task was informed by the fact that it is robust even with limited training data, and not prone to local extremum.
- SVM classifies training instances belonging to either of two classes by fitting a separation boundary (hyper plane) between the classes such that the margin between the boundary and either class is maximized.
- Logistic Regression is a Machine Learning algorithm which is used for the classification problems, it is a predictive analysis algorithm and based on the concept of probability.
- The hypothesis of logistic regression tends it to limit the cost function between 0 and 1.

4.2.5: RESULT GENERATION:

The Final Result will get generated based on the overall classification and prediction. The performance of this proposed approach is evaluated using some measures like,

- **Accuracy**

Accuracy of classifier refers to the ability of classifier. It predicts the class label correctly and the accuracy of the predictor refers to how well a given predictor can guess the value of predicted attribute for a new data.

$$AC= (TP+TN)/ (TP+TN+FP+FN)$$

- **Precision**

Precision is defined as the number of true positives divided by the number of true positives plus the number of false positives.

Precision=TP/ (TP+FP)

- **Recall**

Recall is the number of correct results divided by the number of results that should have been returned. In binary classification, recall is called sensitivity. It can be viewed as the probability that a relevant document is retrieved by the query.

Recall=TP/ (TP+FN)

CHAPTER 5

SYSTEM REQUIREMENTS

5.1 HARDWARE REQUIREMENTS:

• System	:	intel core i7-11700k: 3.6 GHz
• Hard Disk	:	500 GB
• Mouse	:	Logitech.
• Keyboard	:	110 keys enhanced
• Ram	:	8GB

5.2 SOFTWARE REQUIREMENTS:

• O/S	:	Windows 10
• Language	:	Python
• Front End	:	Anaconda Navigator – Spyder

5.3 SOFTWARE DESCRIPTION:

5.3.1 Python

Python is one of those rare languages which can claim to be both *simple* and powerful. You will find yourself pleasantly surprised to see how easy it is to concentrate on the solution to the problem rather than the syntax and structure of the language you are programming in. The official introduction to Python is Python is an easy to learn, powerful programming language. It has efficient high-level data structures and a simple but effective approach to object-oriented programming. Python's elegant syntax and dynamic typing, together with its interpreted nature, make it an ideal language for scripting and rapid application development in many areas on most platforms. I will discuss most of these features in more detail in the next section.

5.3.2 Features of Python

- **Simple**

Python is a simple and minimalistic language. Reading a good Python program feels almost like reading English, although very strict English! This pseudo-code nature of Python is one of its greatest strengths. It allows you to concentrate on the solution to the problem rather than the language itself.

- **Easy to Learn**

As you will see, Python is extremely easy to get started with. Python has an extraordinarily simple syntax, as already mentioned.

- **Free and Open Source**

Python is an example of a *FLOSS* (Free/Libré and Open Source Software). In simple terms, you can freely distribute copies of this software, read its source code, make changes to it, and use pieces of it in new free programs. FLOSS is based on the concept of a community which shares knowledge. This is one of the reasons why Python is so good - it has been created and is constantly improved by a community who just want to see a better Python.

- **High-level Language**

When you write programs in Python, you never need to bother about the low-level details such as managing the memory used by your program, etc.

- **Portable**

Due to its open-source nature, Python has been ported to (i.e. changed to make it work on) many platforms. All your Python programs can work on any of these platforms without requiring any changes at all if you are careful enough to avoid any system-dependent features.

You can use Python on GNU/Linux, Windows, FreeBSD, Macintosh, Solaris, OS/2, Amiga, AROS, AS/400, BeOS, OS/390, z/OS, Palm OS, QNX, VMS, Psion, Acorn RISC OS, VxWorks, PlayStation, Sharp Zaurus, Windows CE and PocketPC!

You can even use a platform like Kivy to create games for your computer *and* for iPhone, iPad, and Android.

- **Interpreted**

This requires a bit of explanation.

A program written in a compiled language like C or C++ is converted from the source language i.e. C or C++ into a language that is spoken by your computer (binary code i.e. 0s and 1s) using a compiler with various flags and options. When you run the program, the linker/loader software copies the program from hard disk to memory and starts running it.

Python, on the other hand, does not need compilation to binary. You just *run* the program directly from the source code. Internally, Python converts the source code into an intermediate form called bytecodes and then translates this into the native language of your computer and then runs it. All this, actually, makes using Python much easier since you don't have to worry about compiling the program, making sure that the proper libraries are linked and loaded, etc. This also makes your Python programs much more portable, since you can just copy your Python program onto another computer and it just works!

- **Object Oriented**

Python supports procedure-oriented programming as well as object-oriented programming. In *procedure-oriented* languages, the program is built around procedures or functions which are nothing but reusable pieces of programs. In *object-oriented* languages, the program is built around objects which combine data and functionality. Python has a very powerful but simplistic way of doing OOP, especially when compared to big languages like C++ or Java.

- **Extensible**

If you need a critical piece of code to run very fast or want to have some piece of algorithm not to be open, you can code that part of your program in C or C++ and then use it from your Python program.

- **Embeddable**

You can embed Python within your C/C++ programs to give *scripting* capabilities for your program's users.

- **Extensive Libraries**

The Python Standard Library is huge indeed. It can help you do various things involving regular expressions, documentation generation, unit testing, threading, databases, web browsers, CGI, FTP, email, XML, XML-RPC, HTML, WAV files, cryptography, GUI (graphical user interfaces), and other system-dependent stuff. Remember, all this is always available wherever Python is installed. This is called the *Batteries Included* philosophy of Python.

Besides the standard library, there are various other high-quality libraries which you can find at the Python Package Index.

5.4 TESTING PRODUCTS:

System testing is the stage of implementation, which aimed at ensuring that system works accurately and efficiently before the live operation commence. Testing is the process of executing a program with the intent of finding an error. A good test case is one that has a high probability of finding an error. A successful test is one that answers a yet undiscovered error.

Testing is vital to the success of the system. System testing makes a logical assumption that if all parts of the system are correct, the goal will be successfully achieved. . A series of tests are performed before the system is ready for the user acceptance testing. Any engineered product can be tested in one of the following ways. Knowing the specified function that a product has been designed to from, test can be conducted to demonstrate each function is fully operational. Knowing the internal working of a product, tests can be conducted to ensure that “al gears mesh”, that is the internal operation of the product performs according to the specification and all internal components have been adequately exercised.

5.4.1 UNIT TESTING:

Unit testing is the testing of each module and the integration of the overall system is done. Unit testing becomes verification efforts on the smallest unit of software design in the module. This is also known as ‘module testing’. The modules of the system are tested separately. This testing is carried out during the programming itself. In this testing step, each model is found to be working satisfactorily as regard to the expected output from the module. There are some validation checks for the fields. For example, the validation check is done for verifying the data given by the user where both format and validity of the data entered is included. It is very easy to find error and debug the system.

5.4.2 INTEGRATION TESTING:

Data can be lost across an interface, one module can have an adverse effect on the other sub function, when combined, may not produce the desired major function. Integrated testing is systematic testing that can be done with sample data. The need for the integrated test is to find the overall system performance. There are two types of integration testing. They are:

i) Top-down integration testing.

ii) Bottom-up integration testing.

5.4.3 TESTING TECHNIQUES/STRATEGIES:

- **WHITE BOX TESTING:**

White Box testing is a test case design method that uses the control structure of the procedural design to drive cases. Using the white box testing methods, we

Derived test cases that guarantee that all independent paths within a module have been exercised at least once.

- **BLACK BOX TESTING:**

1. Black box testing is done to find incorrect or missing function
2. Interface error
3. Errors in external database access
4. Performance errors.
5. Initialization and termination errors

In ‘functional testing’, is performed to validate an application conforms to its specifications of correctly performs all its required functions. So this testing is also called ‘black box testing’. It tests the external behaviour of the system. Here the engineered product can be tested knowing the specified function that a product has been designed to perform, tests can be conducted to demonstrate that each function is fully operational.

5.4.4 SOFTWARE TESTING STRATEGIES

VALIDATION TESTING:

After the culmination of black box testing, software is completed assembly as a package, interfacing errors have been uncovered and corrected and final series of software validation tests begin validation testing can be defined as many, But a single definition is that validation succeeds when the software functions in a manner that can be reasonably expected by the customer

USER ACCEPTANCE TESTING:

User acceptance of the system is the key factor for the success of the system. The system under consideration is tested for user acceptance by constantly keeping in touch with prospective system at the time of developing changes whenever required.

OUTPUT TESTING:

After performing the validation testing, the next step is output asking the user about the format required testing of the proposed system, since no system could be useful if it does not produce the required output in the specific format. The output displayed or generated by the system under consideration. Here the output format is considered in two ways. One is screen and the other is printed format. The output format on the screen is found to be correct as the format was designed in the system phase according to the user needs. For the hard copy also output comes out as the specified requirements by the user. Hence the output testing does not result in any connection in the system.

CHAPTER 6

CONCLUSION

The proposed system for early AD detection in IoT networks using non-amyloid protein profiles and machine learning algorithms has the potential to revolutionize the diagnosis and treatment of this devastating disease. By identifying metabolic processes that are associated with or precede the onset of AD, doctors and researchers can gain valuable insights into the underlying mechanisms of the disease and develop new therapeutic strategies.

The experimental results analysis showed that the proposed method using SVM and LR achieved better performance results on average compared to existing methods. This demonstrates the potential of machine learning in facilitating the development of accurate and efficient diagnostic tools for AD and other complex diseases.

However, further research is needed to validate the proposed system and optimize its performance in diverse populations. Additionally, ethical considerations and data privacy concerns need to be addressed to ensure the responsible and equitable use of IoT technologies in healthcare.

CHAPTER 7

FUTURE ENHANCEMENT

As a future work, a growing understanding of how the disease disrupts the brain has led to potential Alzheimer's treatments that short-circuit basic disease processes. Future Alzheimer's treatments may include a combination of medications, similar to how treatments for many cancers or HIV/AIDS include more than a single drug.

CHAPTER 8

SAMPLE CODE

```
#====================== IMPORT
PACKAGES===============================
import pandas as pd
from sklearn import preprocessing
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import confusion_matrix
from sklearn import linear_model
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings("ignore")
import numpy as np

#===================== READ A INPUT DATA
==============================


dataframe=pd.read_csv("dataset2.csv")
print("*******************************************************")
print()
print("                    Data Selection                    ")
print()
print("******************************************************")
print()
print(dataframe.head(10))
print()
```

```
#========================== PRE PROCESSING
=================================

#=== ckecking missing values ===

print("**************************************")
print()
print("    Before Handling Missing Values    ")
print()
print("**************************************")
print()
print(dataframe.isnull().sum())
print()

#=== replace the missing values by 0 ===

median = dataframe['MMSE'].median()
dataframe['MMSE'].fillna(median, inplace=True)
print("****************************************************")
print()
print("     After Handling Missing Values    ")
print()
print("****************************************************")
print()
print("---- 1.Remove missing values in MMSE ----")
print()
print(dataframe.isnull().sum())
print()
```

```
median = dataframe['SES'].median()
dataframe['SES'].fillna(median, inplace=True)
print()
print("---- 2.Remove missing values in SES ----")
print()
print(dataframe.isnull().sum())
print()

#=== label encoding ===

print("**************************************************")
print()
print("             Before Label Encoding                ")
print()
print("**************************************************")
print()
print(dataframe['Group'].head(10))
label_encoder = preprocessing.LabelEncoder()
print("**************************************************")
print()
print("             After Label Encoding                ")
print()
print("**************************************************")
print()
dataframe['Group']= label_encoder.fit_transform(dataframe['Group'])
print(dataframe['Group'].head(10))
dataframe['M/F']= label_encoder.fit_transform(dataframe['M/F'])
dataframe['Hand'] = label_encoder.fit_transform(dataframe['Hand'])
```

```
#=========================== DATA
SPLITTING  ===========================

feature_col_names = ["M/F", "Age", "EDUC", "SES", "MMSE", "eTIV",
"nWBV", "ASF"]
predicted_class_names = ['Group']

X = dataframe[feature_col_names].values
y = dataframe[predicted_class_names].values

#spliting the x and y into test and train
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25,
random_state=2)

#===================== CLASSIFICATION
====================================

#=== SUPPORT VECTOR MACHINE ===

#initialize the model
svm = SVC(kernel="linear", C=0.1,random_state=0)

#fitting the model
svm.fit(X_train, y_train.ravel())

#predict the model
pred_svm = svm.predict(X_test)
```

```
#=================== PERFORMANCE ANALYSIS
=============================

#=== confusion matrix ===
print("***************************************************")
print()
print("          Performance Metrics for SVM                 ")
print()
print("***************************************************")
cm_svm=confusion_matrix(y_test,pred_svm)
print()
print("1.Confusion Matrix",cm_svm)
print()

#find the performance metrics
TP = cm_svm[0][2]
FP = cm_svm[0][1]
FN = cm_svm[1][0]
TN = cm_svm[1][1]

#Total TP,TN,FP,FN
Total=TP+FP+FN+TN

#Accuracy Calculation
accuracy1=((TP+TN)/Total)*100
print("2.Accuracy",accuracy1,'%')
print()

#Precision Calculation
```

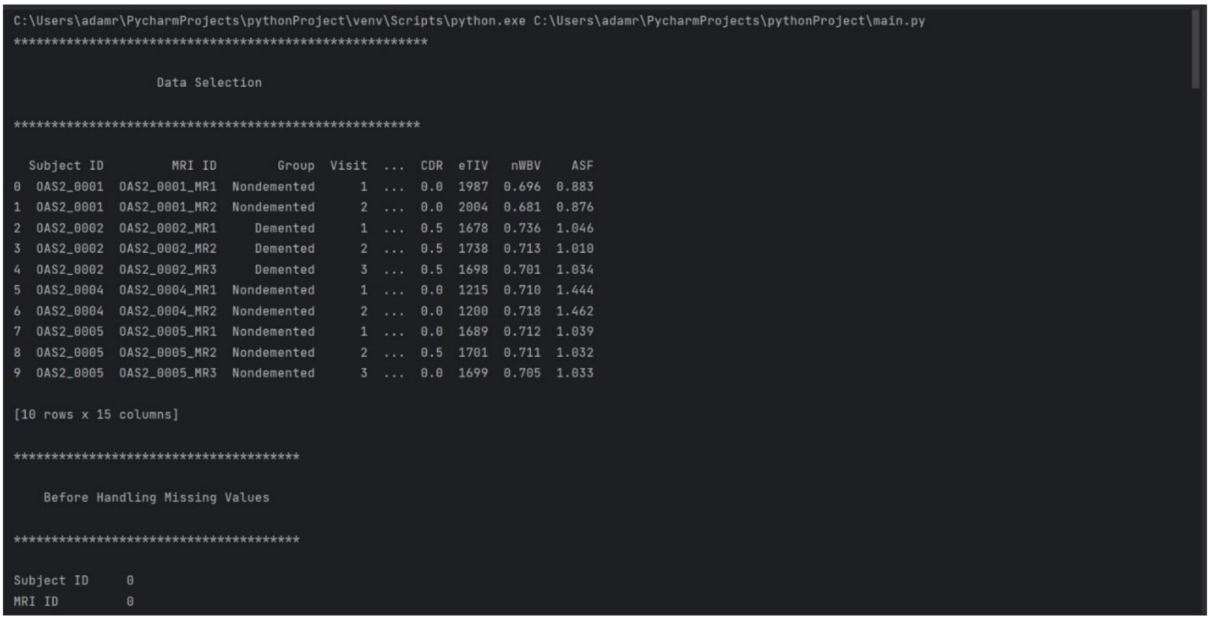
```
precision=TP/(TP+FP)*100
print("3.Precision",precision,'%')
print()

#Sensitivity Calculation
Sensitivity=TP/(TP+FN)*100
print("4.Sensitivity",Sensitivity,'%')
print()

#specificity Calculation
specificity = (TN / (TN+FP))*100
print("5.specificity",specificity,'%')
print()

#========================== LOGISTIC REGRESSION
==================================

#initialize the model
lr = linear_model.LogisticRegression()
#fitting the model
lr.fit(X_train, y_train.ravel())

#predict the model
pred_lr = lr.predict(X_test)

#==================== PERFORMANCE ANALYSIS
===============================

#=== confusion matrix ===
print("****************************************************")
print()
print("          Performance Metrics for LR                 ")
print()
print("****************************************************")
print()
cm_lr=confusion_matrix(y_test,pred_lr)
print()
print("1.Confusion Matrix",cm_lr)
print()

#find the performance metrics
TP = cm_lr[0][2]
FP = cm_lr[0][1]
FN = cm_lr[1][0]
TN = cm_lr[1][1]

#Total TP,TN,FP,FN
Total=TP+FP+FN+TN

#Accuracy Calculation
accuracy2=((TP+TN)/Total)*100
print("2.Accuracy",accuracy2,'%')
print()

#Precision Calculation
precision=TP/(TP+FP)*100
print("3.Precision",precision,'%')
print()
```

```
#Sensitivity Calculation
Sensitivity=TP/(TP+FN)*100
print("4.Sensitivity",Sensitivity,'%')
print()

#specificity Calculation
specificity = (TN / (TN+FP))*100
print("5.specificity",specificity,'%')
print()

#==================== PREDICTION
==================================

#disease prection
for i in range(1,10):
    if pred_lr[i]== 2:
        print("**************************")
        print()
        print([i],' Demented ')
        print()
        print("**************************")
        print()
    else:
        print("**************************")
        print()
        print([i],'Non Demented ')
        print()
        print("*************************")
```

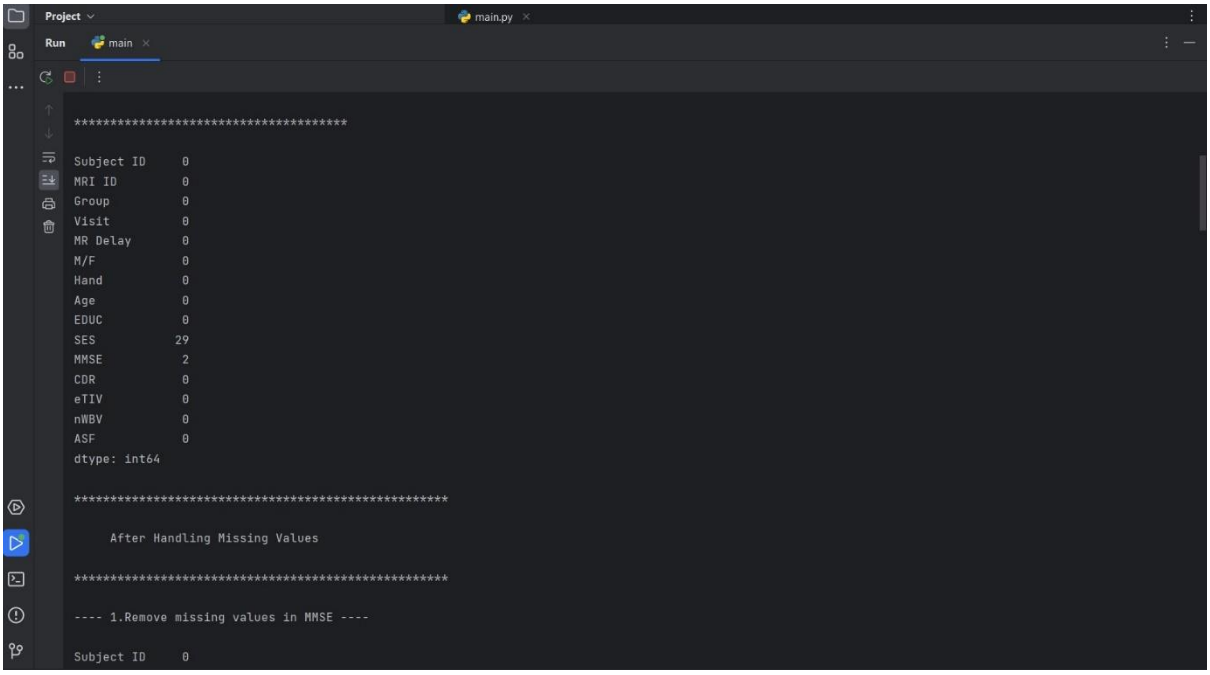
```
      print()

#====================== ALGORITHM COMPARISON
==================================

#algorithm comparion
if(accuracy1>accuracy2):
   print("*****************************************************")
   print()
   print("   Support Vector Machine algorithm is efficient     ")
   print()
   print("*****************************************************")
else:
   print("*****************************************************")
   print()
   print("        Logistic regression is efficient             ")
   print()
   print("*****************************************************")
#====================== VISUALIZATION
==================================
objects = ('SVM', 'LR')
y_pos = np.arange(len(objects))
performance = [accuracy1,accuracy2]

plt.bar(y_pos, performance, align='center', alpha=0.5)
plt.xticks(y_pos, objects)
plt.ylabel('Accuracy')
plt.title('Algorithm comparison')
plt.show()
```

CHAPTER 9

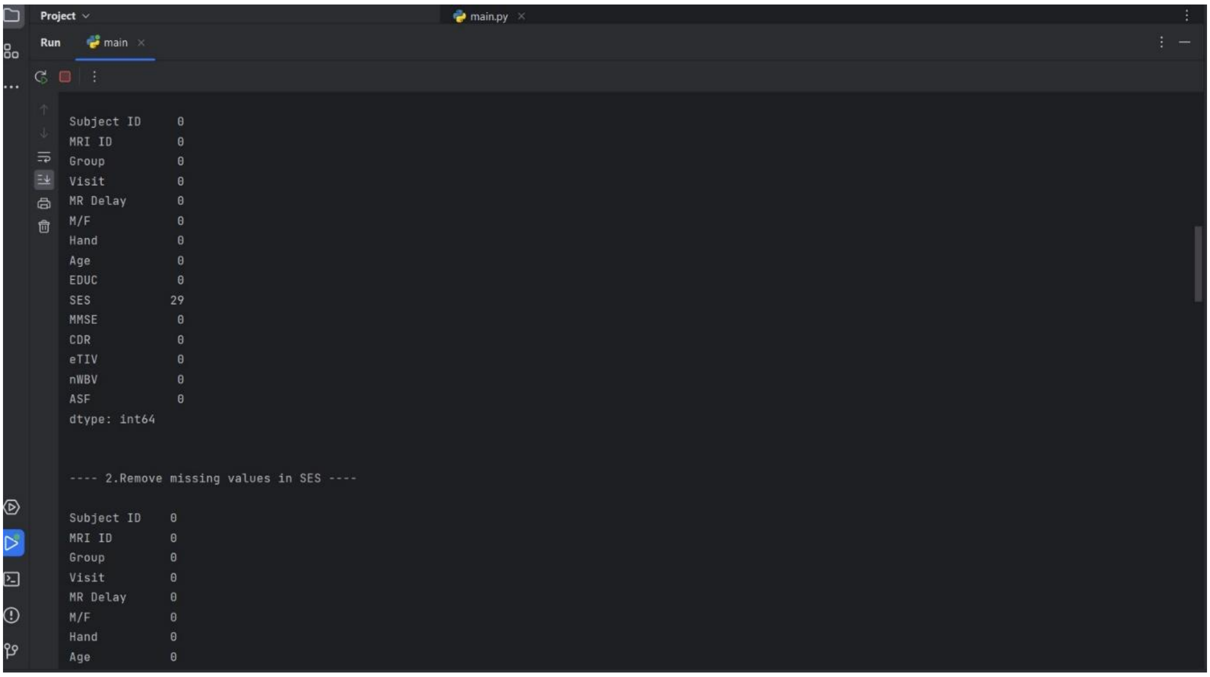
SAMPLE SCREENSHOTS



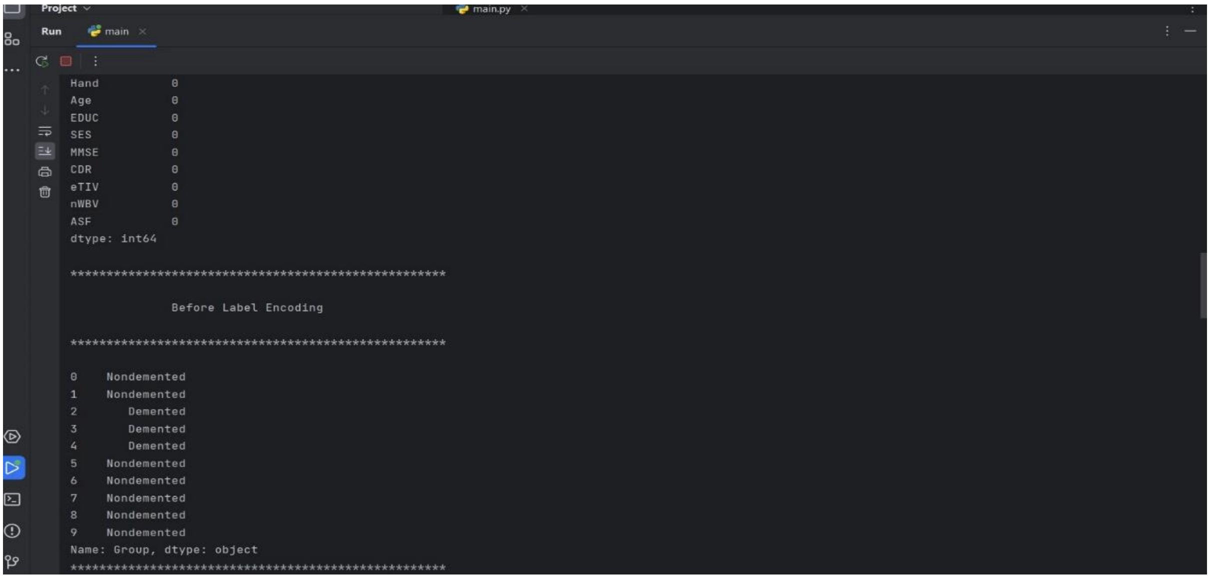
```
C:\Users\adamr\PycharmProjects\pythonProject\venv\Scripts\python.exe C:\Users\adamr\PycharmProjects\pythonProject\main.py
*******************************************************

                   Data Selection

******************************************************

  Subject ID         MRI ID        Group  Visit  ...  CDR  eTIV   nWBV    ASF
0  OAS2_0001  OAS2_0001_MR1  Nondemented      1  ...  0.0  1987  0.696  0.883
1  OAS2_0001  OAS2_0001_MR2  Nondemented      2  ...  0.0  2004  0.681  0.876
2  OAS2_0002  OAS2_0002_MR1     Demented      1  ...  0.5  1678  0.736  1.046
3  OAS2_0002  OAS2_0002_MR2     Demented      2  ...  0.5  1738  0.713  1.010
4  OAS2_0002  OAS2_0002_MR3     Demented      3  ...  0.5  1698  0.701  1.034
5  OAS2_0004  OAS2_0004_MR1  Nondemented      1  ...  0.0  1215  0.710  1.444
6  OAS2_0004  OAS2_0004_MR2  Nondemented      2  ...  0.0  1200  0.718  1.462
7  OAS2_0005  OAS2_0005_MR1  Nondemented      1  ...  0.0  1689  0.712  1.039
8  OAS2_0005  OAS2_0005_MR2  Nondemented      2  ...  0.5  1701  0.711  1.032
9  OAS2_0005  OAS2_0005_MR3  Nondemented      3  ...  0.0  1699  0.705  1.033

[10 rows x 15 columns]

**************************************

    Before Handling Missing Values

**************************************

Subject ID     0
MRI ID         0
```



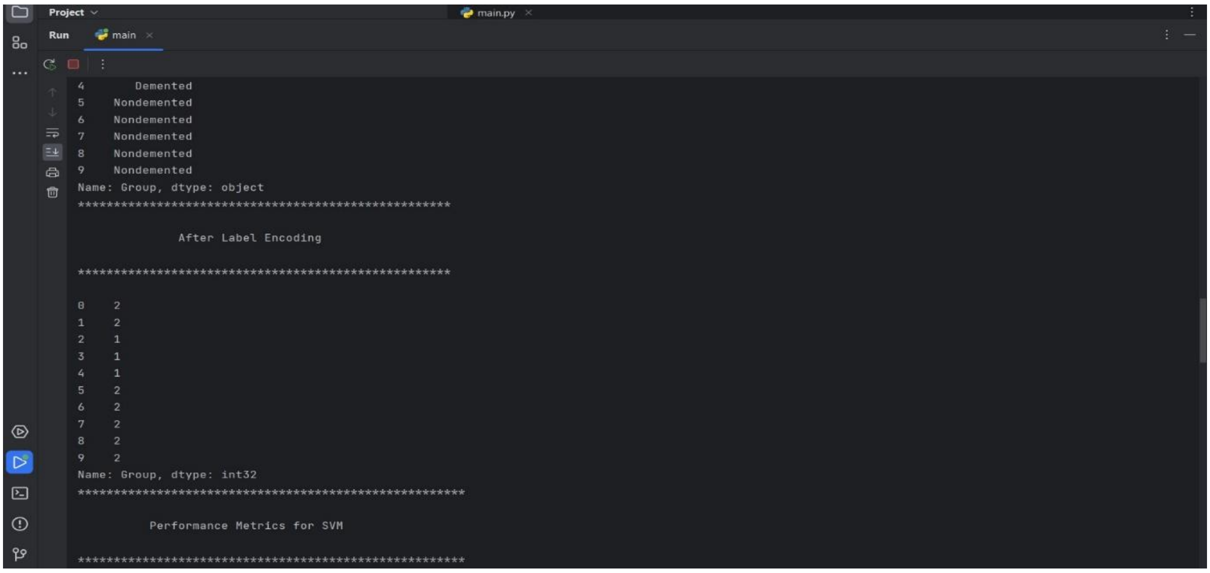
```
**************************************

Subject ID     0
MRI ID         0
Group          0
Visit          0
MR Delay       0
M/F            0
Hand           0
Age            0
EDUC           0
SES           29
MMSE           2
CDR            0
eTIV           0
nWBV           0
ASF            0
dtype: int64

****************************************************

     After Handling Missing Values

****************************************************

---- 1.Remove missing values in MMSE ----

Subject ID     0
```



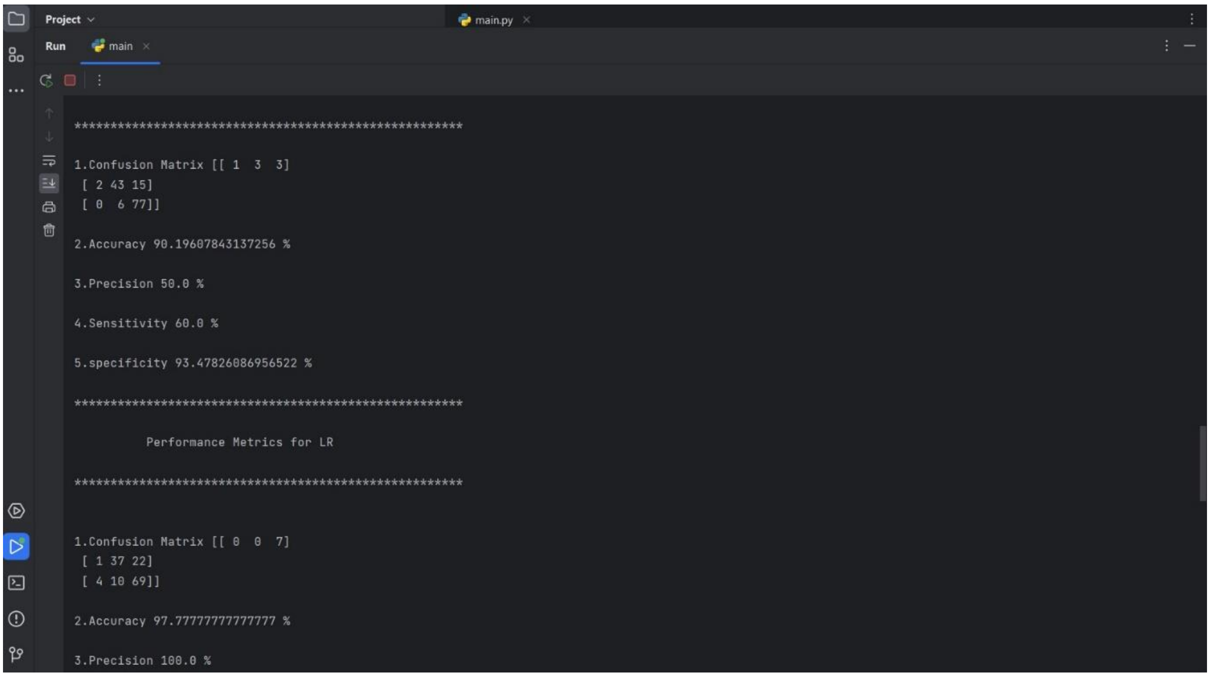
```
Subject ID     0
MRI ID         0
Group          0
Visit          0
MR Delay       0
M/F            0
Hand           0
Age            0
EDUC           0
SES           29
MMSE           0
CDR            0
eTIV           0
nWBV           0
ASF            0
dtype: int64


---- 2.Remove missing values in SES ----

Subject ID    0
MRI ID        0
Group         0
Visit         0
MR Delay      0
M/F           0
Hand          0
Age           0
```



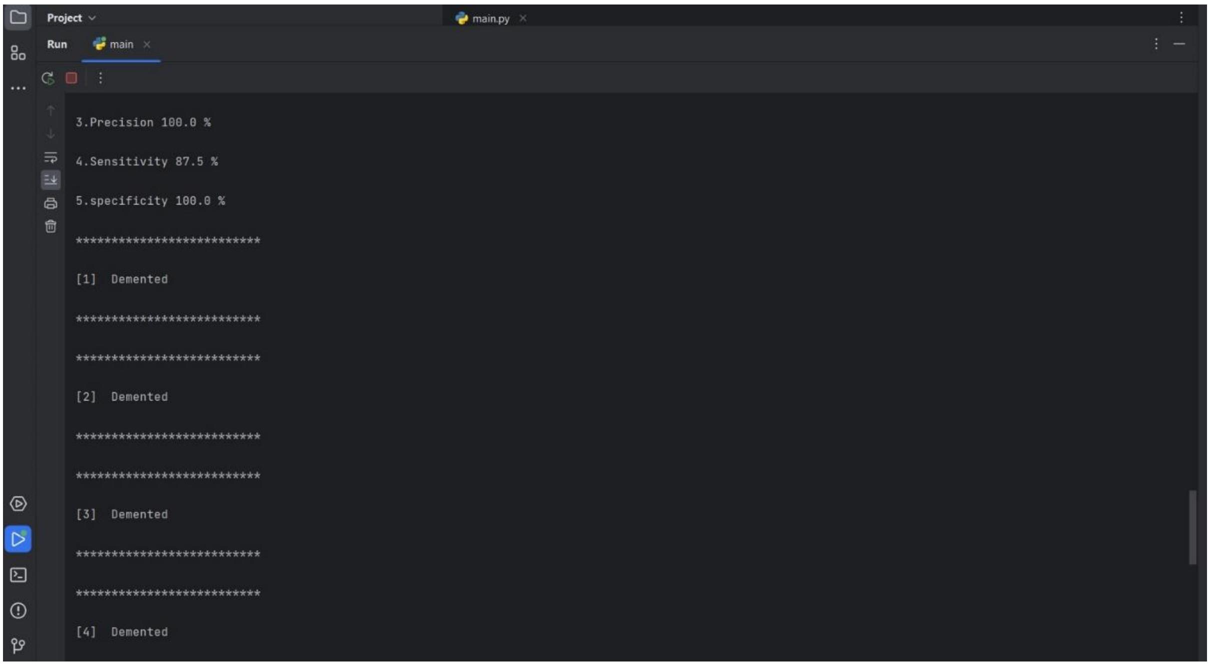
```
Hand          0
Age           0
EDUC          0
SES           0
MMSE          0
CDR           0
eTIV          0
nWBV          0
ASF           0
dtype: int64

*****************************************************

              Before Label Encoding

*****************************************************

0    Nondemented
1    Nondemented
2       Demented
3       Demented
4       Demented
5    Nondemented
6    Nondemented
7    Nondemented
8    Nondemented
9    Nondemented
Name: Group, dtype: object
*****************************************************
```



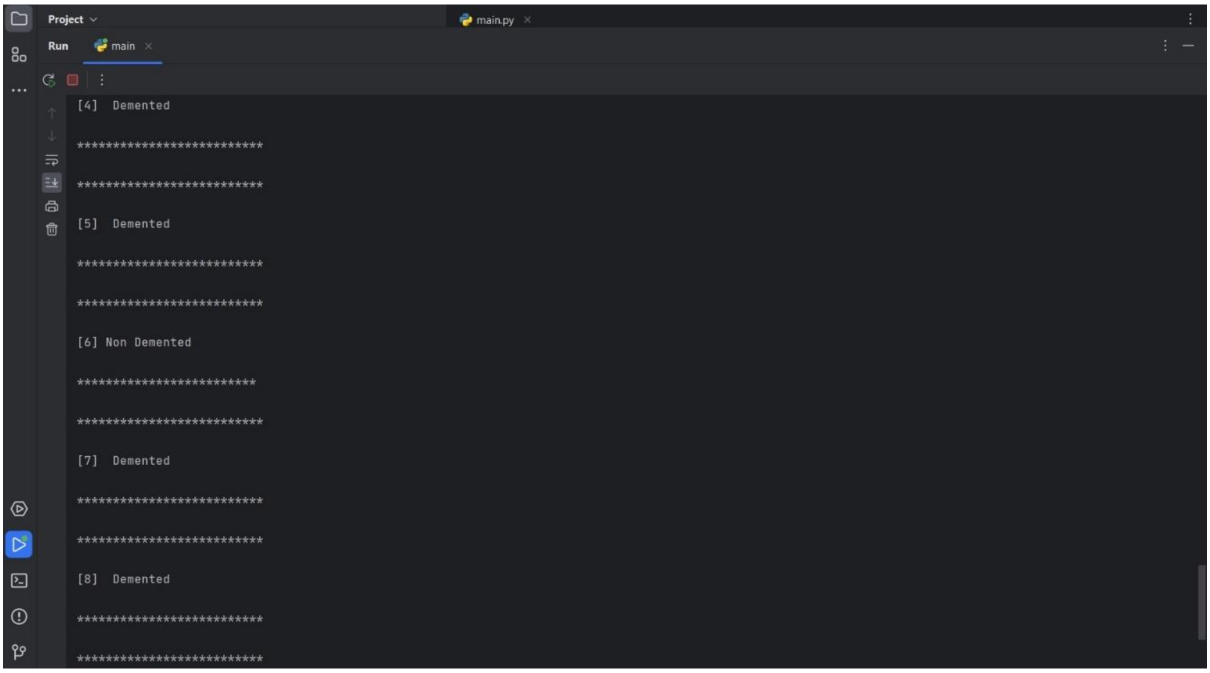
```
4       Demented
5    Nondemented
6    Nondemented
7    Nondemented
8    Nondemented
9    Nondemented
Name: Group, dtype: object
*****************************************************

              After Label Encoding

*****************************************************

0    2
1    2
2    1
3    1
4    1
5    2
6    2
7    2
8    2
9    2
Name: Group, dtype: int32
*******************************************************

          Performance Metrics for SVM

*******************************************************
```



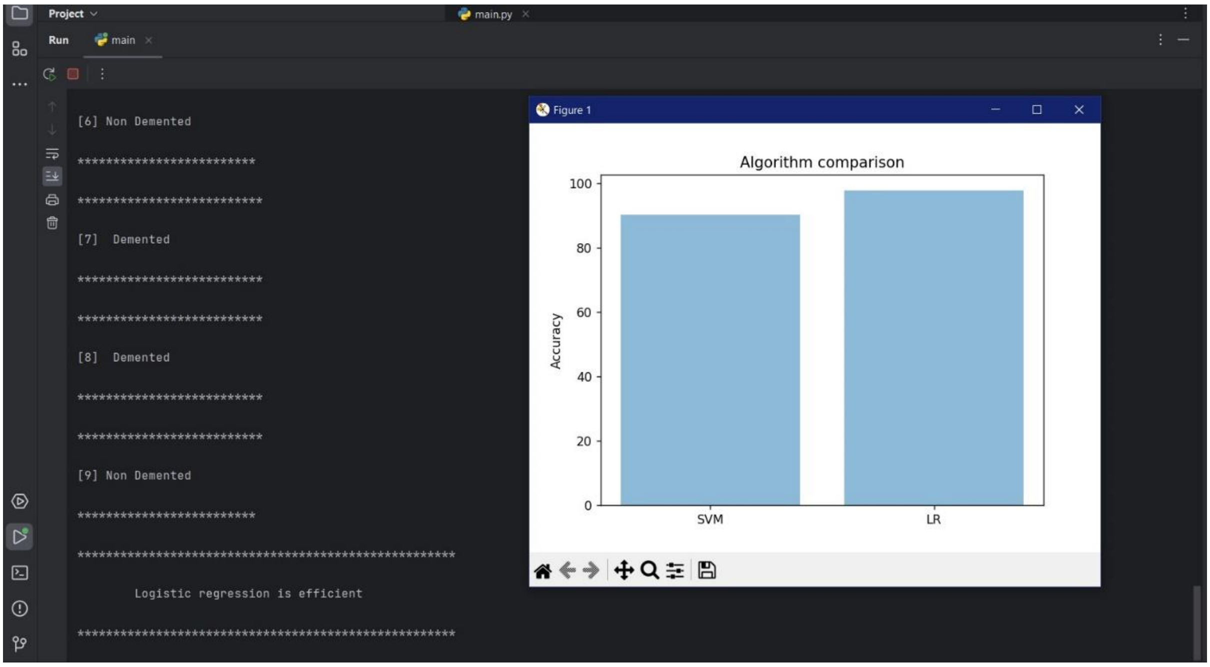
```
******************************************************

1.Confusion Matrix [[ 1  3  3]
 [ 2 43 15]
 [ 0  6 77]]

2.Accuracy 90.19607843137256 %

3.Precision 50.0 %

4.Sensitivity 60.0 %

5.specificity 93.47826086956522 %

******************************************************

          Performance Metrics for LR

******************************************************


1.Confusion Matrix [[ 0  0  7]
 [ 1 37 22]
 [ 4 10 69]]

2.Accuracy 97.77777777777777 %

3.Precision 100.0 %
```



```
3.Precision 100.0 %

4.Sensitivity 87.5 %

5.specificity 100.0 %

**************************

[1]  Demented

**************************

**************************

[2]  Demented

**************************

**************************

[3]  Demented

**************************

**************************

[4]  Demented
```



```
[4]  Demented

**************************

**************************

[5]  Demented

**************************

**************************

[6] Non Demented

*************************

**************************

[7]  Demented

**************************

**************************

[8]  Demented

**************************

**************************
```



```
[6] Non Demented

*************************

**************************

[7]  Demented

**************************

**************************

[8]  Demented

**************************

**************************

[9] Non Demented

*************************

*****************************************************

        Logistic regression is efficient

*****************************************************
```

CHAPTER 10

REFERENCES

1. Association, "2018 Alzheimer's disease facts and figures," Alzheimer's Dementia, vol. 14, no. 3, pp. 367–429, 2018.

2. M. Prince, A. Comas-Herrera, M. Knapp, M. Guerchet, and M. Karagiannidou, World Alzheimer Report 2016: Improving Healthcare for People Living with Dementia: Coverage, Quality and Costs now and in the Future, London, U.K.: Alzheimer's disease International, 2016.

3. B. Dubois et al., "Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria," Alzheimer's Dementia, vol. 12, no. 3, pp. 292–323, 2016.

4. M. S. Albert et al., "The diagnosis of mild cognitive impairment 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 Dementia, vol. 7, no. 3, pp. 270–279, 2011.

5. 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 Dementia, vol. 7, no. 3, pp. 263–269, 2011.

6. R. A. Sperling et al., "Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," Alzheimer's Dementia, vol. 7, no. 3, pp. 280–292, 2011.

7. G. P. Morris, I. A. Clark, and B. Vissel, "Questions concerning the role of amyloid-β in the definition, aetiology and diagnosis of Alzheimer's disease," Acta Neuropathologica, vol. 136, no. 5, pp. 663–689, 2018.

8. K. H. Tse and K. Herrup, “Re-imagining Alzheimer’s disease–the diminishing importance of amyloid and a glimpse of what lies ahead,” J. Neurochemistry, vol. 143, no. 4, pp. 432–444, 2017.

9. F. Zhang, J. Wei, X. Li, C. Ma, and Y. Gao, “Early candidate urine biomarkers for detecting Alzheimer’s disease before amyloid-β plaque deposition in an APP (swe)/PSEN1 dE9 transgenic mouse model,” J. Alzheimer’s Disease, vol. 66, pp. 613–637, 2018.

10. F. Kametani and M. Hasegawa, “Reconsideration of amyloid hypothesis and tau hypothesis in Alzheimer’s disease,” Frontiers Neuroscience, vol. 12, 2018, Paper 25.

11.M. Gold, “Phase II clinical trials of anti–amyloid β antibodies: when is enough, enough?” Alzheimer’s Dementia: Translational Res. Clin. Interventions, vol. 3, no. 3, pp. 402–409, 2017.

12.S. Makin, “The amyloid hypothesis on trial,” Nature, vol. 559, no. 7715, pp. S4–S4, 2018.

13.H. D. Soares, Y. Chen, M. Sabbagh, A. Rohrer, E. Schrijvers, and M. Breteler, “Identifying early markers of Alzheimer’s disease using quantitative multiplex proteomic immunoassay panels,” Annals. New York Academy Sci., vol. 1180, no. 1, pp. 56–67, 2009.

14.H. D. Soares et al., “Plasma biomarkers associated with the apolipoprotein E genotype and Alzheimer disease,” Archives Neurology, vol. 69, no. 10, pp. 1310–1317, 2012.

15. M. Thambisetty et al., “Proteome-based identification of plasma proteins associated with hippocampal metabolism in early Alzheimer’s disease,” J. Neurology, vol. 255, no. 11, pp. 1712–1720, 2008.

16. A. Hye et al., “Proteome-based plasma biomarkers for Alzheimer’s disease,” Brain, vol. 129, no. 11, pp. 3042–3050, 2006.

17. E. Bullmore and O. Sporns, ‘‘Complex brain networks: Graph theoretical analysis of structural and functional systems,’’ Nature Rev. Neurosci., vol. 10, no. 3, pp. 186–198, Mar. 2009.

18.S. Eickhoff, T. E. Nichols, J. D. Van Horn, and J. A. Turner, ‘‘Sharing the wealth: Neuroimaging data repositories,’’ NeuroImage, vol. 124, pp. 1065–1068, Jan. 2016.

19. N. Bhagwat, ‘‘Prognostic applications for Alzheimer’s disease using magnetic resonance imaging and machine-learning,’’ Ph.D. dissertation, Graduate Dept. Inst. Biomater. Biomed. Eng., Univ. Toronto, Toronto, ON, Canada, 2018.

20.J. Yao, ‘‘Development of a multimodal framework for cardiac computed tomography gating,’’ Ph.D. dissertation, School Elect. Comput. Eng., Georgia Inst. Technol., Atlanta, GA, USA, 2018.

21. S. Ahmed, K. Y. Choi, J. J. Lee, B. C. Kim, G.-R. Kwon, K. H. Lee, and H. Y. Jung, ‘‘Ensembles of patch-based classifiers for diagnosis of Alzheimer diseases,’’ IEEE Access, vol. 7, pp. 73373–73383, 2019.

22. Chaddad, C. Desrosiers, and T. Niazi, ‘‘Deep radiomic analysis of MRI related to Alzheimer’s disease,’’ IEEE Access, vol. 6, pp. 58213–58221, 2018.

23. K. F. Hunter, M. Northwood, V. Haggar, and F. Bates, Management of Fecal Incontinence for the Advanced Practice Nurse. Springer, 2018, pp. 127–148.

24.J.-Y. Han, L. M. Besser, C. Xiong, W. A. Kukull, and J. C. Morris, ‘‘Cholinesterase inhibitors may not benefit mild cognitive impairment and mild Alzheimer disease dementia,’’ Alzheimer Disease Associated Disorders, vol. 33, no. 2, pp. 87–94, 2019.

25. I. P. Vatanabe, P. R. Manzine, and M. R. Cominetti, ‘‘Historic concepts of dementia and Alzheimer’s disease: From ancient times to the present,’’ Revue Neurologique, vol. 176, no. 3, pp. 140–147, Mar. 2020.

26. E. R. Adlard and R. J. Sundberg, ‘‘the chemical century: Molecular manipulation and its impact on the 20th century,’’ Chromatographia, vol. 80, no. 10, p. 1599, Oct. 2017.

27. L. Keuck, ‘‘History as a biomedical matter: Recent reassessments of the first cases of Alzheimer’s disease,’’ Hist. Philosophy Life Sci., vol. 40, no. 1, p. 10, Mar. 2018.

28.C. Saraiva, C. Praça, R. Ferreira, T. Santos, L. Ferreira, and L. Bernardino, ‘‘Nanoparticle-mediated brain drug delivery: Overcoming blood–brain barrier to treat neurodegenerative diseases,’’ J. Controlled Release, vol. 235, pp. 34–47, Aug. 2016.