ALZHEIMER DISEASE DETECTION USING ENSEMBLE MODELLING

Submitted in Partial fulfillment of Requirement for the award of the Degree of

BACHELOR OF ENGINEERING

in

Electronics and Communication

Name of the Student: Arun Karajagi USN:1MS19EC020

Name of the Student: Bhagyashree E USN:1MS19EC023

Name of the Student: Bhargavi V USN:1MS19EC025

Name of the Student: G Abhishek Reddy USN:1MS19EC038

Under Guidance of

Dr. Punya Prabha V

Assistant Professor, Department of E & C

Department of Electronics and Communication

RAMAIAH INSTITUTE OF TECHNOLOGY

(Autonomous Institute, Affiliated to VTU)
Accredited by National Board of Accreditation & NAAC with ‘A+’ Grade
MSR Nagar, MSRIT Post, Bangalore-560054
www.msrit.edu

2023

CERTIFICATE

This is to certify that the dessertation work entitled **"Alzheimer Disease Detection Using Ensemble Modelling"** is carried out by **Arun Karajagi(1MS19EC020), Bhagyashree E(1MS19EC023)**, **Bhargavi V(1MS19EC025) and G Abhishek Reddy (1MS19EC038)**, bonafide students of Ramaiah Institute of Technology, Bangalore, in partial fulfillment for the award of Bachelor of Engineering in **Electronics and Communication** of the Visvesvaraya Technological University, Belgaum, during the year 2022 -2023. It is certified that all corrections / suggestions indicated for Internal Assessment have been incorporated in the thesis. The thesis has been approved as it satisfies the academic requirements in respect to dessertation work prescribed for the said degree.

Guide	HOD	Principal
Dr. Punya Prabha V	**Dr. Maya V Karki**	**Dr. N V R Naidu**
Assistant Professor	Professor and HOD,	Principal,
Department of E & C	Department of E & C	RIT, Bangalore
RIT, Bangalore	RIT, Bangalore	

Name & Signature of Examiners with Date: -

1)

2)

RAMAIAH
Institute of Technology

DECLARATION

We hereby declare that the Project entitled "Alzheimer Disease Detection Using Ensemble Modelling" has been carried out independently at Ramaiah Institute of Technology under the guidance of **Dr. Punya Prabha V, Assistant Professor, Department of Electronics and Communication, RIT, Bangalore.**

We hereby declare that work submitted in this thesis is our own, except where acknowledged in the text and has not been previously submitted for the award of the degree of Visvesvaraya Technological University, Belgaum or any other institute or University

Signature of Students:

1. Arun Karajagi (USN:1MS19EC020)
2. Bhagyashree E (USN:1MS19EC023)
3. Bhargavi V (USN:1MS19EC025)
4. G Abhishek Reddy (USN:1MS19EC038)

Place:

Date:

ACKNOWLEDGEMENT

The immense satisfaction that accompanies the successful completion of the project would be incomplete without the mention of the people who made it possible. I consider it my honour to express my deepest gratitude and respect to the following people who always guided and inspired me during the course of the Project.

I am deeply indebted to **Dr. N. V. R. Naidu**, Principal, RIT, Bangalore for providing me with a rejuvenating master course under a very creative learning environment.

I am much obliged to **Dr. Maya V Karki**, Professor & HOD, Department of Electronics and Communication Engineering, RIT, Bangalore for her constant support and motivation.

I sincerely thank my guide **Dr. Punya Prabha V, Assistant Professor,** Department of Electronics and Communication Engineering, RIT, Bangalore and express my humble gratitude for her valuable guidance, inspiration, encouragement and immense help which made this project work a success.

I sincerely thank **name of reviewer 1** and **reviewer 2** for reviewing my project work and providing valuable suggestions. I also thank all the faculty members of Department of E&C, RIT for their kind support to carry out this project successfully.

Last but not the least I would like to express my heartfelt gratitude to my parents, relatives and friends for their constant support, motivation and encouragement.

ABSTRACT

A neurological disorder called Alzheimer's disease affects the brain and has an impact on memory, thought, and behaviour. It develops slowly and cannot be stopped. Dementia, a general term for a mental deterioration, severe enough to interfere with daily life, is most frequently caused by this. It causes cognitive function to deteriorate until it is ultimately unable to perform daily chores. Structural MRI can show the changes in the brain's structure.

In this paper, we describe an ensemble model for the detection of Alzheimer's disease. Although Alzheimer's disease does not yet have a cure, there are therapies that can help control the symptoms and enhance the quality of life for people who are affected. This project's goal is to contribute to the creation of precise and trustworthy instruments for the early identification of AD, which can significantly affect patient outcomes and further our knowledge of the condition. In our research, we provide an ensemble model to precisely identify imaging and clinical data indicative of moderate cognitive impairment (MCI), Alzheimer's disease (AD) and Normal Conditioned (NC).

ACRONYMS

AD	Alzheimer Disease
MCI	Mild Cognitive Impairment
NC	Normal Control
MRI	Magnetic Resonance Imaging
CNN	Convolution Neural Network
DL	Deep Learning
PET	Positron Emission Tomography
ROC	Receiver Operating Characteristic
NIfTI	Neuroimaging Informatics Technology Initiative
ADNI	Alzheimer's Disease Neuroimaging Initiative
CSV	Comma separated Values
FC	Fully Connected
ML	Machine Learning
MLP	Multi-Layer Perceptron
KNN	K-Nearest Neighbour
AUC	Area Under ROC Curve

Contents

List of figures

List of tables

RAMAIAH
Institute of Technology

CHAPTER 1

INTRODUCTION

1.1. INTRODUCTION

Alzheimer's disease is a gradual and irreversible neurological condition that affects the brain and impairs memory, thinking, and behaviour. Dementia is a generic word used to describe a deterioration in mental ability that is severe enough to interfere with daily life. Its most common cause is Alzheimer's disease. It results in cognitive decline, which finally makes it unable to carry out daily duties. The alterations in brain structure can be seen on structural MRI.

The illness starts off with mild changes in memory and thinking, but as it worsens, it can make it difficult for a person to carry out everyday tasks, communicate, and operate independently. Beta-amyloid plaques and neurofibrillary tangles are two aberrant brain formations that are indicative of Alzheimer's disease. The symptoms of Alzheimer's disease can only be managed with present therapies because there is no known cure for the condition. Alzheimer's disease is a serious public health issue since it is predicted to become more prevalent as the population ages.

In this paper, we describe an ensemble model for the detection of Alzheimer's disease. Although the precise aetiology of Alzheimer's disease is still unknown, researchers think that a variety of genetic, environmental, and lifestyle variables may have a role in the disease's progression. Although Alzheimer's disease does not yet have a cure, there are therapies that can help control the symptoms and enhance the quality of life for people who are affected.

1.2. PROPLEM STATEMENT

- Screening for mild cognitive impairment (MCI) is a frequent practise, and early detection of Alzheimer's disease (AD) is essential for optimal care.
- Convolutional neural network (CNN) is one of several deep-learning techniques that have been used to assess structural changes in the brain on magnetic resonance imaging (MRI).

- This is because CNN is incredibly effective at automating feature learning using a variety of multilayer perceptrons.
- To accurately identify AD and moderate cognitive impairment (MCI) using imaging and clinical data, we provide an ensemble model.

1.3. MOTIVATION FOR THE WORK

Millions of people throughout the world are afflicted by the neurological ailment known as Alzheimer's disease. It is a gradual illness that results in memory loss, cognitive deterioration, and behavioural changes, ultimately making it unable to carry out even the most fundamental daily tasks. To effectively treat and manage Alzheimer's disease and to help patients and their families, early recognition of the condition is essential.

A machine learning technique called ensemble modelling mixes the output of various models to increase accuracy and robustness. By taking into account a variety of elements and variables, ensemble modelling can improve the accuracy of the diagnosis in the case of Alzheimer's disease detection.

The motivation of this study is to provide a precise and trustworthy approach for detecting Alzheimer's disease at its earliest stages, which will help in timely diagnosis and treatment. By combining the strengths of many machine learning models, such as neural networks, decision trees, and support vector machines, ensemble modelling offers a viable method to accomplish this goal. The project seeks to create a useful tool for medical professionals to diagnose Alzheimer's disease in its early stages, potentially enhancing the quality of life for patients and their families. This is accomplished by utilising the power of ensemble modelling.

1.4. OBJECTIVE

The symptoms of Alzheimer's disease (AD) include significant memory loss and cognitive decline. It is linked to major alterations in brain structure that can be seen by magnetic resonance imaging (MRI) scans. Using image classification technologies like convolutional neural network (CNN), the observed preclinical structural changes offer a chance for AD early identification. The sample size of the majority of AD-related studies, however, is currently a limitation. It is crucial to find a productive technique to train an image classifier with little data. In our effort, we show an ensemble model to successfully identify moderate cognitive impairment (MCI) and Alzheimer's disease (AD) from imaging and clinical data.

The project's primary goals are to:

1. The objective of this project is to contribute to the development of accurate and reliable tools for the early diagnosis of AD, which can have a significant impact on patient outcomes and improve our understanding of the disease.

1.5. SCOPE

The project's scope is to develop an accurate, reliable, and scalable system for the early detection of Alzheimer's disease using ensemble modeling techniques.

The project seeks to create a useful tool for medical professionals to diagnose Alzheimer's disease in its early stages, potentially enhancing the quality of life for patients and their families. This is accomplished by utilising the power of ensemble modelling.

1.6. ORGANIZATION OF THE REPORT

The contents of this thesis are coordinated in the following way.

1. Introduction: This section will provide a brief overview of the importance of Alzheimer disease detection in its early stages.
2. Literature Review: This section will review the existing research on plant Alzheimer disease detection using deep learning techniques. It will cover the various deep learning algorithms and architectures used for plant Alzheimer disease detection.
3. Methodology: This section will describe the methodology used for Alzheimer disease detection. It will cover training two different models with different types of data inputs and finally ensemble model is built by combining those two models.
4. Experimental Results: This section will present the experimental results of the proposed models for Alzheimer disease detection using ensemble modelling. It will include results of Inception V3 model trained with imaging data, different machine learning models trained with clinical data and the ensemble model.
5. Discussion: This section will discuss the results of the experiments and the implications for the use of ensemble modelling techniques for Alzheimer disease detection.
6. Conclusion: This section will summarise the main findings of the thesis and their

implications for the field of Alzheimer disease detection using ensemble modelling.

7. References: This section will include a list of references cited throughout the thesis.

In summary, a thesis on Alzheimer disease detection using ensemble modelling will cover the introduction and importance of the topic, a literature review of the existing research, the proposed methodology, experimental results and analysis, and future directions for research in this area.

CHAPTER 2

LITERATURE SURVEY

Here are some recent papers on Alzheimer disease detection using deep learning techniques:

"Alzheimer's Disease Diagnosis With Brain Structural MRI Using Multiview-Slice Attention and 3D Convolution Neural Network" by Lin Chen, Hezhe Qiao and Fan Zhu (2022) - The multiview-slice attention and 3D convolution neural network (3D-CNN) are the foundations for the unique AD diagnosis model presented in this paper. To be more precise, they start by employing several sub-networks to extract the local slice-level feature in different dimensions. Then, in order to eliminate the redundant characteristics, they designed a slice-level attention method to emphasise particular 2D-slices. A 3D-CNN was then used to record the global subject-level structural changes that had occurred after that. In order to create representations that are more discriminative, all of these 2D and 3D attributes were combined.

"Automated Detection of Alzheimer's Disease and Mild Cognitive Impairment Using Whole Brain MRI" by Fazal Ur Rehman and Goo-Rak Kwon (2022) - The proposed mechanism was utilised in the research to hierarchically transform the magnetic resonance imaging images into more condensed high-level features by combining features from several layers.The computation complexity is decreased by the proposed method's fewer parameters.The approach is then compared with current state-of-the-art works for AD classification, which demonstrate superior results for the frequently used assessment metrics, such as accuracy, an area under the ROC curve, etc., indicating that their suggested convolution operation is appropriate for the AD diagnosis.

"Automatic Diagnosis of Alzheimer's Disease and Mild Cognitive Impairment Based on CNN + SVM Networks with End-to-End Training" by Zhe Huang, Minglang Sun and Chengan Guo (2021)- They suggested a hybrid model in this research that integrates CNN and SVM networks to forecast NC, MCI, and AD. The suggested model's structure consists of two modules: a feature extraction module based on CNN with 3D kernels (3DCNN), and a classification module based on SVM.

"A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer Disease using MRI scans" by Shagun Sharma, Kalpana Gueria, Sunita Tiwari and Sushil Kumar (2022) - The research suggests utilising CNN and Deep Learning to diagnose Alzheimer's disease using MRI scans. The VGG16 model, a pre-trained CNN-based model, is used to extract features from the MRI scan pictures. Evaluation of the suggested model's performance using various datasets and determination of the optimal DL model for the prediction of AD stages in the MRI scan dataset is performed.

“CNNs Based Multi-Modality Classification for AD Diagnosis” by Danni Cheng, Manhua Liu (2017) – In order to classify AD using MRI and PET scans, this paper suggests building multi-level convolutional neural networks (CNNs) that gradually learn and incorporate the multi-modality features.

CHAPTER 3

METHODOLOGY

3.1. IMPLEMENTATION

1. Data inputs - clinical data (demographics, memory tests, balance score, etc.) and imaging (MRI scans).
2. We train two different models using these two different types of data as input for each.
3. Clinical data is trained using Random Forest and Decision Tree machine learning algorithm.
4. MRI data is obtained in nii format from ADNI which is converted into png format later into jpeg.
5. Jpeg MRI images are trained using CNN model.
6. After training the models, ensemble modeling is done to combine the results of two models.
7. Lastly, we can give the test dataset as an input to the ensemble model for the classification of Alzheimer's disease stages (CN, MCI, and AD)

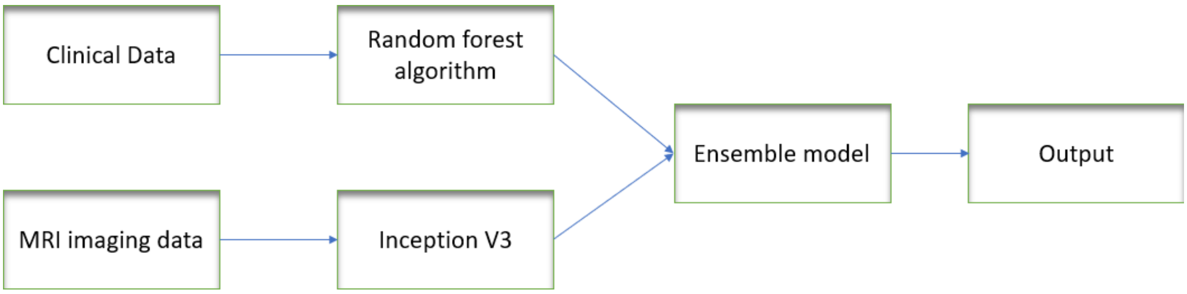

Figure 1: Block diagram representation of methodology

3.2. CLINICAL DATA

A sizable clinical dataset for Alzheimer's disease has been produced by the Alzheimer's Disease Neuroimaging Initiative (ADNI), a research initiative. This dataset contains a variety of information from people with normal cognition, mild cognitive impairment (MCI), and Alzheimer's disease, including clinical evaluations, brain imaging, and biomarkers.

The ADNI clinical dataset includes clinical details for each person together with data on recruiting, demographics, physical examinations, and cognitive tests. CSV files (comma separated values) are downloaded in bulk, containing the whole collection of clinical data. The ADNI dataset has been used to research the course of Alzheimer's disease, find potential biomarkers for the disease's early diagnosis, and test new treatments.

The ADNI dataset contains information from more than 1,500 participants, including data gathered over several years at various points in time. Researchers can access the data through the ADNI website, and access is given after the ADNI Data and Publication Committee has reviewed the proposed research topic.

The collection contains a variety of clinical measurements, including demographic data, functional assessments, medical histories, and cognitive tests. Additionally, there is data on brain imaging, such as positron emission tomography (PET) and magnetic resonance imaging (MRI) scans.

Overall, the ADNI dataset is a useful tool for researchers looking into Alzheimer's disease and other illnesses, and it has greatly advanced our knowledge of the condition and the quest for novel treatments.

We employ 2384 patients' clinical data from neurological testing (such as balance tests), cognitive evaluations (such as memory tests), and patient demographics (such as age). There are 29 features in total in the clinical data, which is quantitative, categorical, or binary. When a patient takes medication to treat existing conditions, for example, we deleted any feature that could indicate a direct indicator of AD.

Clincal data set is trained using three different machine learning algorithms i.e., Random Forest, Decision Tree and Support Vector Machine.

3.2.1. Decision Tree

A supervised machine learning technique known as a decision tree is employed for both classification and regression problems. Each internal node represents a test on a property or feature, each branch indicates the test's result, and each leaf node represents a class label or a numerical value. The structure is similar to a flowchart. The decision tree algorithm divides the data into subsets recursively according to the values of the input features, starting with the complete dataset as the root node. The algorithm's objective is to produce a tree that can accurately categorise the incoming data by reducing impurity or entropy at each node of the tree.

3.2.2. SVM

A popular supervised machine learning method called Support Vector Machine, or SVM for short, is used for classification, regression, and detection of outliers applications. In order to partition the data into two or more groups, a discriminative method known as SVM identifies the optimum hyperplane in a high-dimensional feature space.

The hyperplane's margin, or the separation between it and the nearest data points in each class, is chosen to be maximized. The support vectors, which are the points closest to the hyperplane, determine the hyperplane's location. The objective of SVM is to identify the hyperplane with the biggest margin and the best generalization performance on untested data.

By using different kernel functions that increase the dimensionality, SVM can handle problems with both linear and nonlinear classification.

3.2.3. Random Forest

A popular machine learning method called Random Forest is utilised for both classification and regression problems. A large number of decision trees are built during the training phase of this ensemble learning method, which then produces the class or mean prediction of each tree.

An example of a bagging algorithm is Random Forest, which mixes the predictions of many models trained on various subsets of the training data to increase accuracy and minimise overfitting. The individual models used in Random Forest are decision trees that were trained using arbitrary subsets of the characteristics and data samples. This randomization in feature and data selection lowers the correlation between the trees and enhances the model's generalization capabilities.

Each decision tree in the Random Forest predicts the output class or regression value during the prediction phase, and the final output is derived by combining all of the trees' predictions. Depending on the job, the aggregation process can be carried out in a variety of methods, such as taking the majority vote for classification or the mean value for regression.

When compared to other machine learning methods, Random Forest provides a number of benefits. It is simple to use, needs minimum preprocessing of the data, effectively handles missing values and noisy data, and performs well even with high-dimensional data. It can also manage categorical and numerical characteristics and rate the importance of the features, which is helpful for feature selection and comprehending the underlying data.

3.3. IMAGING DATA

ADNI 1 basline data is taken in nii format.

The ADNI 1 Baseline dataset in NIfTI (.nii) format is a collection of neuroimaging data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study. The dataset includes T1-weighted magnetic resonance imaging (MRI) scans of the brain, which were acquired at the baseline visit of the ADNI 1 study.

The ADNI 1 study was conducted to identify biomarkers of Alzheimer's disease (AD) progression and to develop better methods for clinical trials in AD. The study collected a wide range of data from participants, including neuroimaging, cognitive testing, and biomarker measurements.

The NIfTI format is a widely used file format for neuroimaging data. It stores the MRI data in a 3D grid of voxels, with each voxel representing a small unit of the brain volume. The ADNI 1 Baseline dataset in NIfTI format allows researchers to analyze and compare the brain structure and function of participants at the baseline visit of the ADNI 1 study.

The dataset is trained using Inception V3 model.

3.3.1. Inception V3 Architecture

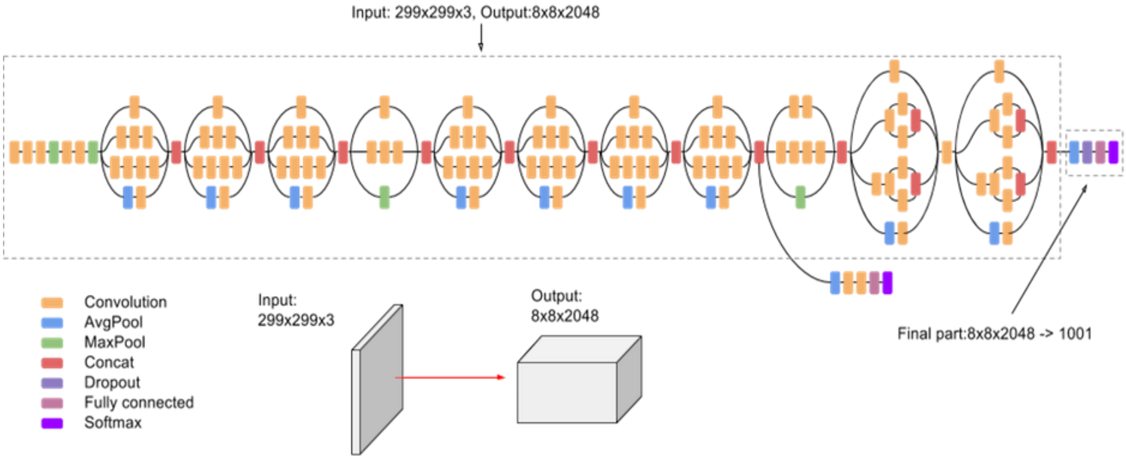

Figure 2: Inception V3 Architecture

- The Inception V3 is an image categorization deep learning model based on convolutional neural networks. The Inception V3, which was released in 2014 as GoogLeNet, is an improved version of the initial model, Inception V1. It was created by a team at Google, as the name would imply.
- In various biomedical applications, Inception V3 has demonstrated strong classification performance. As a result, we experimented with using Inception V3 architecture in our model.
- Inception V3 is a convolution neural network with 48 layers, which is slightly more than the previous iterations of Inception. However, this model's efficiency is quite outstanding.
- Convolution filters, average pools, and sometimes max pools are included in each layer, and each layer is followed by a concatenate to merge the outputs of all the filters. Fully connected layers and a softmax layer are placed after these.

3.3.2. Inception V3 model summary

```
Model: "inception_cnn_model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 inception_v3 (Functional)   (None, 4, 4, 2048)        21802784

 dropout (Dropout)           (None, 4, 4, 2048)        0

 global_average_pooling2d (G  (None, 2048)             0
 lobalAveragePooling2D)

 flatten (Flatten)           (None, 2048)              0

 batch_normalization_94 (Bat  (None, 2048)             8192
 chNormalization)

 dense (Dense)               (None, 512)               1049088

 batch_normalization_95 (Bat  (None, 512)              2048
 chNormalization)

 dropout_1 (Dropout)         (None, 512)               0

 dense_1 (Dense)             (None, 256)               131328

 batch_normalization_96 (Bat  (None, 256)              1024
 chNormalization)

 dropout_2 (Dropout)         (None, 256)               0

 dense_2 (Dense)             (None, 128)               32896

 batch_normalization_97 (Bat  (None, 128)              512
 chNormalization)

 dropout_3 (Dropout)         (None, 128)               0

 dense_3 (Dense)             (None, 64)                8256

 dropout_4 (Dropout)         (None, 64)                0

 batch_normalization_98 (Bat  (None, 64)               256
 chNormalization)

 dense_4 (Dense)             (None, 3)                 195

=================================================================
Total params: 23,036,579
Trainable params: 1,227,779
Non-trainable params: 21,808,800
_________________________________________________________________
```

Figure 3: Inception V3 model summary

3.4. ENSEMBLE MODEL

We utilized two different models, Random Forest and Inception V3, to classify patients with Alzheimer's disease using both clinical and imaging data. The results of both models were combined using an ensemble method to improve overall performance.

Ensemble model is trained using Random forest ML algorithm.

CHAPTER 4

RESULTS & DISCUSSION

4.1. CLINICAL DATA

Clinical data is trained using SVM algorithm, we got an accuracy of 76.98%.

```
# Fitting the classifier into the Training set
DT = DecisionTreeClassifier(max_depth=6, random_state=1)
DT.fit(X_train, y_train)
```

```
DecisionTreeClassifier(max_depth=6, random_state=1)
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

```
# Predicting the test set results
y_Pred = DT.predict(X_test)
```

```
accuracy = accuracy_score(y_test, y_Pred)
print("accuracy:",accuracy)
```

```
accuracy: 0.769857433808554
```

Figure 4: Image showing accuracy of Decision Tree model

Clinical data is trained using SVM algorithm, we got an accuracy of 79%.

```
# Fitting SVM with the training set
SVM = SVC(kernel='linear', random_state=0)
SVM.fit(X_train, y_train)
```

```
]: SVC(kernel='linear', random_state=0)
```

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

```
# Testing the model by classifying the test set
y_pred = SVM.predict(X_test)
```

```
accuracy = accuracy_score(y_test, y_pred)
print("accuracy:",accuracy)
```

```
accuracy: 0.790224032586558
```

Figure 5: Image showing accuracy of SVM model

Clinical data is trained using Random forest algorithm, we got an accuracy of 85.98%.

```
# Predicting the test set results
y_Pred = RF.predict(X_test)
```

```
print(classification_report(y_test, y_Pred))
```

```
              precision    recall  f1-score   support

           0       0.89      0.92      0.91       181
           1       0.82      0.75      0.79       161
           2       0.85      0.91      0.88       129

    accuracy                           0.86       471
   macro avg       0.86      0.86      0.86       471
weighted avg       0.86      0.86      0.86       471
```

```
accuracy = accuracy_score(y_test, y_Pred)
print("accuracy:",accuracy)
```

```
accuracy: 0.8598726114649682
```

Figure 6: Image showing accuracy of Randomforest model

ML Algorithm	Accuracy
Decision Tree	76.98%
SVM	79%
Random forest	85.98%

Table 1: Comaparision of ML Algorithms

From the above table, we got best accuracy from Random forest algorithm.

4.2. MRI IMAGING DATA

Using Inception V3 model, we got 88.06% training accuracy and 87.24% validation accuracy.

```
Epoch 16/20
195/195 [==============================] - 200s 1s/step - loss: 0.4033 - acc: 0.8416 - auc: 0.9558 - f1_score: 0.8417 - val_
loss: 0.3810 - val_acc: 0.8468 - val_auc: 0.9612 - val_f1_score: 0.8460 - lr: 0.0010
Epoch 17/20
195/195 [==============================] - 197s 1s/step - loss: 0.3914 - acc: 0.8482 - auc: 0.9583 - f1_score: 0.8483 - val_
loss: 0.3416 - val_acc: 0.8649 - val_auc: 0.9687 - val_f1_score: 0.8646 - lr: 0.0010
Epoch 18/20
195/195 [==============================] - 196s 1s/step - loss: 0.3474 - acc: 0.8693 - auc: 0.9670 - f1_score: 0.8694 - val_
loss: 0.3214 - val_acc: 0.8752 - val_auc: 0.9723 - val_f1_score: 0.8752 - lr: 0.0010
Epoch 19/20
195/195 [==============================] - 196s 1s/step - loss: 0.3335 - acc: 0.8774 - auc: 0.9690 - f1_score: 0.8774 - val_
loss: 0.3329 - val_acc: 0.8732 - val_auc: 0.9701 - val_f1_score: 0.8731 - lr: 0.0010
Epoch 20/20
195/195 [==============================] - 194s 996ms/step - loss: 0.3283 - acc: 0.8806 - auc: 0.9699 - f1_score: 0.8806 - v
al_loss: 0.3143 - val_acc: 0.8829 - val_auc: 0.9734 - val_f1_score: 0.8828 - lr: 0.0010
```

```
#Evaluating the model on the data

#train_scores = model.evaluate(train_data, train_labels)
#val_scores = model.evaluate(val_data, val_labels)
test_scores = custom_inception_model.evaluate(test_data, test_labels)

#print("Training Accuracy: %.2f%%"%(train_scores[1] * 100))
#print("Validation Accuracy: %.2f%%"%(val_scores[1] * 100))
print("Testing Accuracy: %.2f%%"%(test_scores[1] * 100))
```

```
61/61 [==============================] - 51s 842ms/step - loss: 0.3346 - acc: 0.8724 - auc: 0.9692 - f1_score: 0.8723
Testing Accuracy: 87.24%
```

Figure 7: Results of Inception V3 model

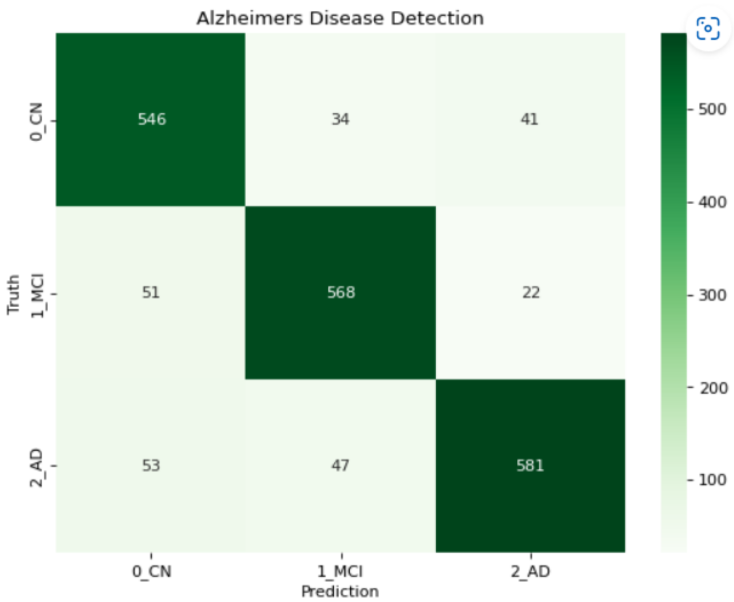

Figure 8: confusion matrix of Inception V3 model

The confusion matrix for the test result obtained for identification of alzheimer disease with Inception V3 model as shown in figure. The x-axis depicts predicted labels and y-axis depicts actual labels. The diagonal values of the matrix portrayed the number of images that were correctly classified by the model whereas; the non-diagonal elements depict the misclassifications.

4.3. ENSEMBLE MODEL

The Random Forest model was trained on clinical data where we got an accuracy of 85.98%, while the Inception V3 model was trained on imaging data where we got an accuracy of 87.24%. Both models showed good performance on their respective datasets. From ensemble model we got 95.91% accuracy. The ensemble of these models showed improved accuracy compared to individual models.

```
# Train a new classifier on top of the concatenated predictions
from sklearn.ensemble import RandomForestClassifier
ensemble_model = RandomForestClassifier(n_estimators=100, max_depth=10, random_state=0)
ensemble_model.fit(ensemble_X_test, y_test)
```

```
RandomForestClassifier
RandomForestClassifier(max_depth=10, random_state=0)
```

```
# Evaluate the ensemble model
from sklearn.metrics import accuracy_score
ensemble_preds = ensemble_model.predict(ensemble_X_test)
ensemble_score = accuracy_score(y_test, ensemble_preds)
print("Ensemble accuracy:", ensemble_score)
```

```
Ensemble accuracy: 0.9591836734693877
```

Figure 9: Results of ensemble model

CHAPTER 5

CONCLUSION & FUTURE WORK

5.1. CONCLUSION

The project titled "Alzheimer's Disease Detection Using Ensemble Learning" utilized two different models, Random Forest and Inception V3, to classify patients with Alzheimer's disease using both clinical and imaging data. The results of both models were combined using an ensemble method to improve overall performance.

The Random Forest model was trained on clinical data where we got an accuracy of 85.98%, while the Inception V3 model was trained on imaging data where we got an accuracy of 87.24%. Both models showed good performance on their respective datasets. The ensemble of these models showed improved accuracy compared to individual models.

The results of this project suggest that combining clinical and imaging data can lead to more accurate classification of patients with Alzheimer's disease. Furthermore, the use of ensemble learning can further improve the accuracy of the classification.

5.2. FUTURE WORK

There are several potential directions for future work on the project titled "Alzheimer's disease Detection using ensemble learning," where a random forest model is trained with clinical data and an Inception V3 model is trained with imaging data, and the outputs of the two models are ensembled. Here are a few future scopes of the project:

Expand the dataset: The performance of machine learning models is often correlated with the size and quality of the training data. Collecting and incorporating more data into the model training process could lead to better accuracy and generalization. For example, we could try to acquire more clinical data or imaging data, or combine data from multiple sources to increase the diversity of the dataset.

Incorporate more advanced machine learning techniques: Ensemble learning is a powerful method for combining the outputs of multiple models, but there are many other advanced machine learning techniques that could be incorporated into the pipeline to improve performance. For example, you could try using deep learning models such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs) to analyze the imaging data, or using more advanced ensemble learning methods such as stacking or boosting.

Investigate feature importance: Random forest models are able to rank the importance of different features in the dataset, which can provide insight into which clinical factors are most strongly associated with Alzheimer's disease. Similarly, we could investigate which imaging features are most informative for predicting Alzheimer's disease using techniques such as feature selection or principal component analysis (PCA).

Evaluate the model on external datasets: To test the generalizability of the model, it would be valuable to evaluate its performance on external datasets that were not used in training. This could help identify potential issues with overfitting or bias in the model, and provide insights into how well the model is likely to perform in real-world clinical settings.

Incorporate additional modalities: In addition to clinical and imaging data, there may be other modalities that could provide valuable information for predicting Alzheimer's disease. For example, genetic data or data from wearable devices could be incorporated into the model training process to improve accuracy and predictive power.

RAMAIAH
Institute of Technology

REFERENCES

[1] J. Venugopalan, L. Tong, H. R. Hassanzadeh, and M. D. Wang, "Multimodal deep learning models for early detection of Alzheimer's disease stage," Scientific reports, vol. 11, no. 1, pp. 1–13, 2021.

[2] C. Fang, C. Li, P. Forouzannezhad, M. Cabrerizo, R. E. Curiel, D. Loewenstein, R. Duara, and M. Adjouadi, "Gaussian discriminative component analysis for early detection of alzheimer's disease: A supervised dimensionality reduction algorithm," Journal of Neuroscience Methods, vol. 344, p. 108 856, 2020

[3] P. Forouzannezhad, A. Abbaspour, C. Li, C. Fang, U. Williams, M. Cabrerizo, A. Barreto, J. Andrian, N. Rishe, R. E. Curiel, et al., "A gaussian-based model for early detection of mild cognitive impairment using multimodal neuroimaging," Journal of neuroscience methods, vol. 333, p. 108 544, 2020.

[4] S. El-Sappagh, T. Abuhmed, S. Riazul Islam, and K. S. Kwak, "Multimodal multitask deep learning model for alzheimer's disease progression detection based on time series data," Neurocomputing, vol. 412, pp. 197–215, 2020

[5] T. Abuhmed, S. El-Sappagh, and J. M. Alonso, "Robust hybrid deep learning models for alzheimer's progression detection," Knowledge-Based Systems, vol. 213, p. 106 688, 2021