



ML-Based Quantitative Analysis of Linguistic and Speech Features Relevant in Predicting Alzheimer's Disease

Tripti Tripathi and Rakesh Kumar

Computer Sciences Department, MMMUT, Gorakhpur, India
tripti.gkp7@gmail.com, rkiitr@gmail.com

KEYWORDS

*Alzheimer's disease;
Feature selection;
Machine learning;
Deep learning*

ABSTRACT

Alzheimer's disease (AD) is a severe neurological condition that affects numerous people globally with detrimental consequences. Detecting AD early is crucial for prompt treatment and effective management. This study presents a novel approach for detecting and classifying six types of cognitive impairment using speech-based analysis, including probable AD, possible AD, mild cognitive impairment (MCI), memory impairments, vascular dementia, and control. The method employs speech data from DementiaBank's Pitt Corpus, which is preprocessed and analyzed to extract pertinent acoustic features. The characteristics are subsequently used to educate five machine learning algorithms, namely k-nearest neighbors (KNN), decision tree (DT), support vector machine (SVM), XGBoost, and random forest (RF). The effectiveness of every algorithm is assessed through a 10-fold cross-validation. According to the research findings, the suggested method based on speech obtains a total accuracy of 75.59% concerning the six-class categorization issue. Among the five machine learning algorithms tested, the XGBoost classifier showed the highest accuracy of 75.59%. These findings indicate that speech-based approaches can potentially be valuable for detecting and classifying cognitive impairment, including AD. The paper also explores robustness testing, evaluating the algorithms' performance under various circumstances, such as noise variability, voice quality changes, and accent variations. The proposed approach can be developed into a noninvasive, cost-effective, and accessible diagnostic tool for the early detection and management of cognitive impairment.

Tripti Tripathi and Rakesh Kumar

ML-Based Quantitative Analysis of Linguistic and
Speech Features Relevant in Predicting Alzheimer's
Disease



1. Introduction

Alzheimer's disease is a degenerative neurological illness that impairs the brain's capability to operate normally, and it is the primary contributor to dementia, which is a cluster of symptoms that affect cognition, memory, and social skills. The precise cause of Alzheimer's disease is yet to be discovered; however, it is speculated that genetic, environmental, and lifestyle factors may be responsible. An abnormal build-up of proteins such as beta-amyloid and tau in the brain characterizes the disease. This can impede communication between neurons and ultimately result in their demise.

The early symptoms of Alzheimer's disease can include memory loss, confusion, difficulty with language, and changes in mood and personality. As the disease progresses, individuals may experience more severe cognitive and functional impairments, including difficulty with basic tasks such as dressing, bathing, and eating. At present, there is no known cure for Alzheimer's disease. However, various treatments can assist in controlling symptoms and decelerating the advancement of the disease (König *et al.*, 2015). These treatments may include medications, lifestyle changes, and cognitive and behavioral interventions. Additionally, early detection and diagnosis are essential for individuals with Alzheimer's disease, as they can help individuals receive appropriate care and support.

There are several reasons why Alzheimer's disease may go undetected or misdiagnosed in its early stages. Patients may wait too long before seeking medical attention or have difficulty explaining their symptoms to their doctor. Additionally, during the initial phases of the ailment, indications may be subtle and often mistaken for signs of other commonplace diseases, which can create difficulty for medical professionals in rendering a precise diagnosis (Javeed *et al.*, 2023). Recognizing and diagnosing Alzheimer's disease early is vital, enabling patients to access the necessary care and support. This can encompass various interventions, such as treatments to manage symptoms and decelerate the advancement of the disease, as well as supportive services for patients and their caregivers. Healthcare providers may use various methods, including cognitive testing and brain imaging, to help diagnose Alzheimer's disease and differentiate it from other conditions with similar symptoms.

While there is currently no cure for Alzheimer's, early detection can help individuals receive appropriate care and support to manage the disease's progression. Research has shown that changes in speech patterns may occur in individuals with Alzheimer's. Specifically, individuals with Alzheimer's disease may experience difficulty with word-finding, sentence construction, and producing coherent and meaningful speech. Therefore, speech analysis has been explored as a potential tool for detecting Alzheimer's disease. Several research studies have employed machine learning methods to examine speech characteristics, such as pitch, tone, and pauses, to differentiate between individuals who have Alzheimer's disease and those who do not.

Speech recognition using machine learning and neural networks has been explored as a potential tool for predicting the risk of Alzheimer's disease (Liu *et al.*, 2020). Researchers have used various techniques to analyze speech features, such as pitch, tone, and pauses, to identify changes in speech patterns that may indicate the disease (Jarrold *et al.*, 2014). An often-used approach utilizes machine learning algorithms to evaluate speech recordings from people who have Alzheimer's disease and those who do not. The algorithms can then identify patterns in the speech features that distinguish between the two groups, allowing for the development of predictive models to identify individuals at higher risk of the disease. Neural networks, a machine learning algorithm that can recognize patterns in large datasets, have also been used in speech recognition for Alzheimer's prediction (Khodabakhsh *et al.*, 2014). These networks can be trained on large amounts of speech data to identify patterns and develop predictive models for the identification of individuals at higher risk of the disease.

The proposed technique for forecasting and discriminating Alzheimer’s disease (AD) from other neurodegenerative illnesses utilizing speech (audio) transcripts is a noninvasive method that displays the potential to enhance early detection and diagnosis of AD. The study explores which speech features are the most relevant in accurately predicting and diagnosing AD using machine learning models. The investigators discovered that speech characteristics such as verbal fillers (e.g., “um”), the slower rate of speech, slurring, stammering, repetition, and the usage of incorrect words or phrases are prevalent among individuals who have AD (Meghanani *et al.*, 2021). These speech problems can also indicate other neurodegenerative diseases, such as mild cognitive impairment, vascular dementia, and aphasia, making it challenging to distinguish between these conditions. By analyzing a person’s speech, we aim to identify which speech factors can be used to diagnose AD accurately. This could lead to developing more effective diagnostic tools and treatments for AD, improving patient outcomes and quality of life.

The following are the primary contributions of our research:

1. To propose a machine learning-based approach that identifies relevant speech features in predicting Alzheimer’s disease (AD). The study aims to improve early detection and diagnosis of AD by analyzing speech transcripts from individuals with AD and other neurodegenerative disorders.
2. To train and compare the performance of multi-class classifiers for distinguishing between six classes of cognitive impairment: probable Alzheimer’s disease (AD), possible AD, mild cognitive impairment (MCI), memory impairments, vascular dementia, and control.
To perform a qualitative study on the impact of various speech features in predicting Alzheimer’s disease (AD) and classifying it from other neurodegenerative disorders.

The rest of the paper is organized as follows: Section 2 presents a summary of prior investigations concerning the detection of Alzheimer’s disease and any relevant work on speech-based Alzheimer’s detection or classification. It also highlights the gaps in the literature that this study aims to address. Section 3 describes the data sources, image preprocessing steps, and machine learning techniques, provides a detailed discussion of the classification techniques, and Section 4 provides a clear interpretation of the results and their implications for Alzheimer’s disease detection. It may also highlight any limitations of the study, suggest future research directions, and summarize the study’s main findings and their significance. It may also offer potential clinical applications or future research directions.

2. Literature Review

Zargarbashi and Babaali (2019) proposed an approach to diagnosis systems for the early detection of Alzheimer’s disease using noninvasive data. The authors utilized a classification system based on spoken language and applied three techniques (statistical and neural) to categorize audio signals into two classes: dementia and control. To design a multi-modal feature embedding on the spoken language audio signal, they used N-gram, i-vector, and x-vector techniques. The system’s evaluation was performed on the cookie picture description task from DementiaBank’s Pitt Corpus, and it was found that the system’s accuracy was 83.6%. Therefore, this study implies that machine learning-based classification systems can help detect Alzheimer’s disease early by analyzing spoken language.

Kundaram and Pathak (2021) proposed a deep learning method that uses a DCNN to classify Alzheimer’s disease using MRI samples. The early diagnosis of Alzheimer’s disease is crucial for patient care and treatment. Deep learning techniques allow for learning high-level features from the dataset, resulting in improved performance compared to hand-crafted feature learning methods such as

traditional machine learning techniques. The DCNN was trained to classify the disease as Alzheimer's disease (AD), mild cognitive impairment (MCI), and standard control (NC) using the ADNI dataset. The results of the experiments showed an accuracy of 98.57%, which is higher than other studies. It should be noted that while this study presented promising results, further research is needed to validate the findings and to determine the generalizability of this model to real-world clinical settings. Additionally, the study may have limitations in the dataset and experimental setup, so further research is needed to evaluate the performance of this model in larger, more diverse datasets.

Tóth *et al.* (2018) discussed that the acoustic parameters of speech were analyzed in individuals with and without dementia using statistical analysis and machine learning techniques. The results indicated noteworthy differences in acoustic parameters such as speech tempo, articulation rate, silent pause, hesitation ratio, length of utterance, and the pause-per-utterance ratio between the two groups. The disparities were most prominent in the speech tempo during the delayed recall task and the number of pauses in the question-answering study.

Kumar *et al.* (2022) focused on identifying a concise set of speech features that could aid in recognizing dementia and implementing machine learning (ML) and deep learning (DL) models. The study utilized speech samples from the Pitt Corpus in Dementia Bank and proposed a combination of prosodic, voice quality, and cepstral features for the task. The experimental results demonstrate that the ML models outperform the DL models (87.6% vs. 85%) using the proposed speech feature combination, with lower time and memory consumption. These findings are promising and improve upon those of earlier studies on dementia recognition using speech.

Cummins *et al.* (2020) review Alzheimer's dementia recognition through the spontaneous speech (ADReSS) challenge, which offers a platform for participants to create speech and language-based systems for recognizing Alzheimer's dementia (AD) using speech recordings and transcripts. The study aims to evaluate various modern approaches to these modalities. The results indicate that linguistic techniques are more effective than acoustic systems. Combining BoAW, End-to-End CNN, and hierarchical-attention networks yielded the best test-set outcome, surpassing the challenge's standards.

Pappagari *et al.* (2021) explore using speech and speaker recognition technologies and natural language processing to detect Alzheimer's disease (AD) and estimate mini-mental status evaluation (MMSE) scores. The Interspeech 2021 ADReSSo challenge dataset was used to accomplish this. The research concentrates on adapting state-of-the-art speaker recognition and language models individually and in combination to evaluate their complementary performance for the tasks. The study's most outstanding models yielded an 84.51% accuracy in the automatic detection of AD and 3.85 RMSE in MMSE prediction.

Pan *et al.* (2021) highlight the efficacy of utilizing acoustic and linguistic information embedded in spontaneous speech recordings for automatic Alzheimer's disease detection. The paper delves into two cutting-edge ASR paradigms, Wav2vec2.0 and time delay neural networks (TDNN), on the ADReSSo dataset, which includes recordings of people describing the Cookie Theft (CT) picture. The study proposed five models to solely investigate the extraction of acoustic and linguistic information from audio recordings.

Liu *et al.* (2021) introduced a neural network-based technique for identifying Alzheimer's disease (AD) in the spontaneous speech of subjects during a picture description task without depending on

manual transcriptions and annotations of the lesson. The approach extracts bottleneck features from audio via an ASR model. It utilizes convolutional neural network (CNN) layers for local context modeling, bidirectional long short-term memory (BiLSTM) layers for global context modeling, and an attention-pooling layer for classification.

Mittal *et al.* (2020) suggested that detecting the early stages of Alzheimer’s disease (AD) is challenging because of the absence of a definitive in vivo diagnosis. The authors propose a multi-modal deep learning approach using speech and corresponding transcripts to identify AD. A CNN-based model predicts the diagnosis for various speech segments combined for the final prediction. When trained and evaluated on the DementiaBank’s Pitt Corpus, the proposed method achieves 85.3% accuracy in 10-fold cross-validation.

Meghanani *et al.* (2021) explored the efficacy of Log-Mel spectrograms and MFCC features in Alzheimer’s disease (AD) recognition and mini-mental state examination (MMSE) score prediction using the ADReSS challenge dataset. Three deep neural networks (DNNs) - CNN-LSTM, ResNet-LSTM, and pBLSTM-CNN - were used. CNN-LSTM showed an accuracy of 64.58% using MFCC features, while ResNet-LSTM achieved an accuracy of 62.5% with Log-Mel spectrograms. The proposed approach achieved an accuracy of 62.5% on the ADReSS dataset.

Below Table 1 shows the comparative study of classification approaches using machine learning.

Table 1. Comparative study of classification approaches using machine learning

Author & Year	Method	Target	Accuracy
Karlekar <i>et al.</i> , 2018	CNN, RNN	Binary Class	91.1%
Orimaye <i>et al.</i> , 2018	D2NN	Multi-class	88.9%
Warnita <i>et al.</i> , 2018	GCNN	Binary-class	73.60%
Chien <i>et al.</i> , 2019	BiLSTM	Binary class	83.80%
Fritsch <i>et al.</i> , 2019	LSTM	Multi class	85.60%
Hong <i>et al.</i> , 2019	RNN	Multi-class	83.50%
Koo <i>et al.</i> , 2020	CNN+Bi-LSTM	Multi-class	81.25%
Cummins <i>et al.</i> , 2020	Bi-LSTM	Binary-class	85.20%
Searle <i>et al.</i> , 2020	DistilBERT	Binary-class	81%
Casanova <i>et al.</i> , 2020	RNN+CRF	Multi-class	81%
Xue <i>et al.</i> , 2021	LSTM	Multi-class	67.50%
Chlasta and Wołk, 2021	DemCNN	Binary-class	62.50%
Gauder <i>et al.</i> , 2021	CNN	Binary-class	78.90%
Haulcy <i>et al.</i> , 2021	SVM, RF	Binary-class	85.40%
Liu <i>et al.</i> , 2021	CNN+biLSTM	Binary-class	82.59%
Wang <i>et al.</i> , 2021	C-AttentionUnifed mode	Binary-class	80.28%
Bertini <i>et al.</i> , 2022	AuDeep	Binary-class	93.30%
AI-Atroshi <i>et al.</i> , 2022	GMM-DBN	Binay-class	86.76%
Meghanani <i>et al.</i> , 2021	FNN	Multi-class	83.33%
Liu <i>et al.</i> , 2021	Machine learning	Multi-class	80%

3. Materials and Methods

This paper uses a set of standardized neuropsychological tests to confirm each participant’s cognitive impairment diagnosis. A speech analysis cross-platform software tool capable of extracting various features from audio recordings, such as pitch, intensity, and duration, is used in this paper; we use the CLAN platform. Once the extracted features were preprocessed to remove noise and normalize the data. Feature selection was then performed to identify the most relevant features for the classification task. Here is a proposed framework for Alzheimer’s disease detection and classification. Figure 1 shows the complete flowchart for dementia detection using machine learning (ML) techniques.

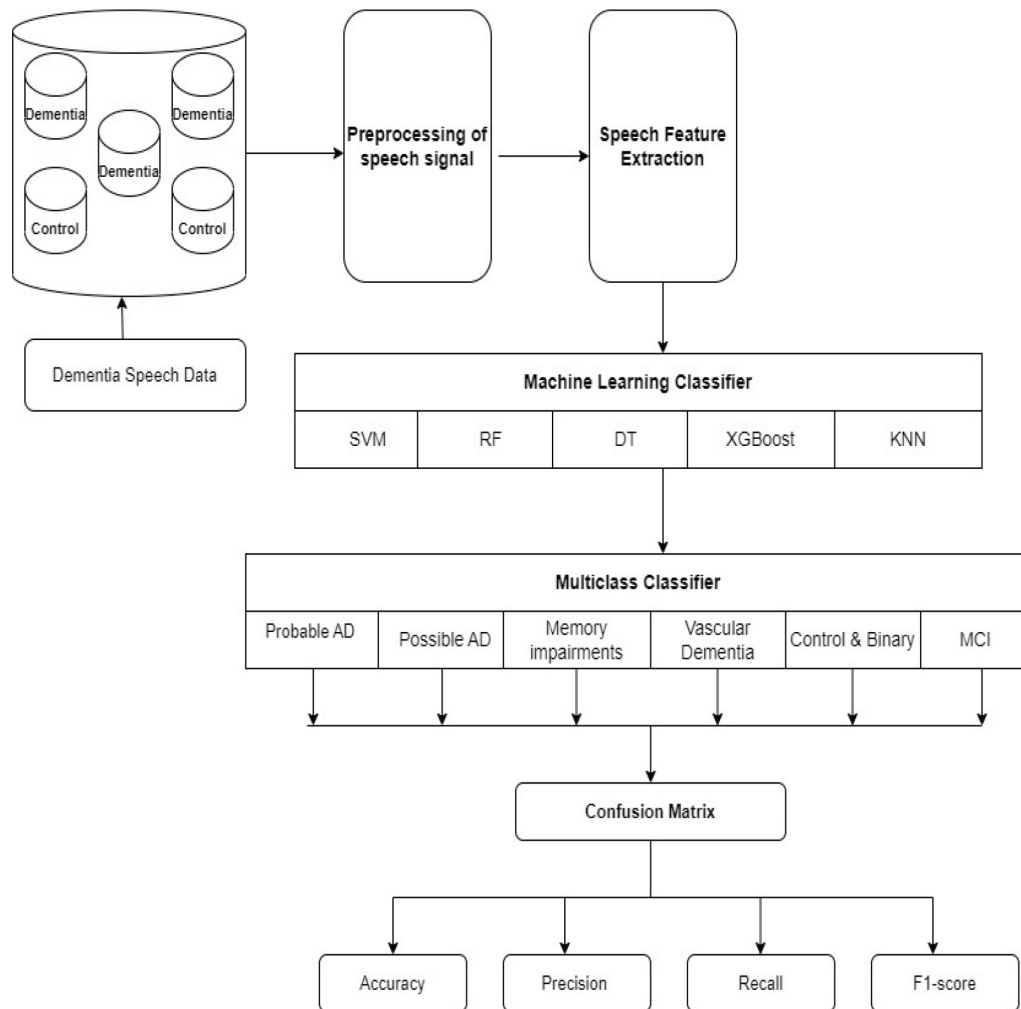


Figure 1. Flowchart depicting the proposed methodology for dementia detection using machine learning (ML) techniques

3.1. Data Acquisition

In this paper, we collected data from DementiaBank’s Pitt Corpus dataset, which consists of conversations between individuals with various forms of dementia and their interlocutors. Much work has already been done with this dataset, which is discussed in (Chen *et al.*, 2019). The dataset includes codes for transcribing speech and annotating various aspects of language use, such as turn-taking, discourse structure, and speech errors. Overall, the DementiaBank’s Pitt Corpus dataset is a valuable resource for research on neurodegenerative diseases, particularly Alzheimer’s disease, and has been utilized extensively in the field for developing and evaluating machine learning models for diagnosing and predicting these conditions. Initially, our dataset contained 1200 entries and over 100 features, saved in a CSV file. To ensure that the data was fit for training models, we cleaned it extensively, including removing redundant columns and handling missing values and errors. Additionally, you standardized the data to prevent any feature from disproportionately influencing the model. As a result of this preprocessing, we identified 50 relevant features used to train our models.

3.2. Preprocessing

The acquired data needs to be preprocessed to remove any noise, it must be correct for processing, and normalized across patients. This step is essential to ensure that machine learning algorithms can accurately detect and classify Alzheimer’s disease. After collecting the data, we processed it to ensure it was in the appropriate format, making it easier to proceed with our tasks. To achieve high accuracy in speech recognition, high-quality acoustic segments, which are portions of the speech free from background noise or interference, must be used. This can be done using advanced signal processing techniques such as noise reduction, filtering, and speech enhancement (Pan *et al.*, 2020). Figure 2 shows an example of automatic transcript selection.

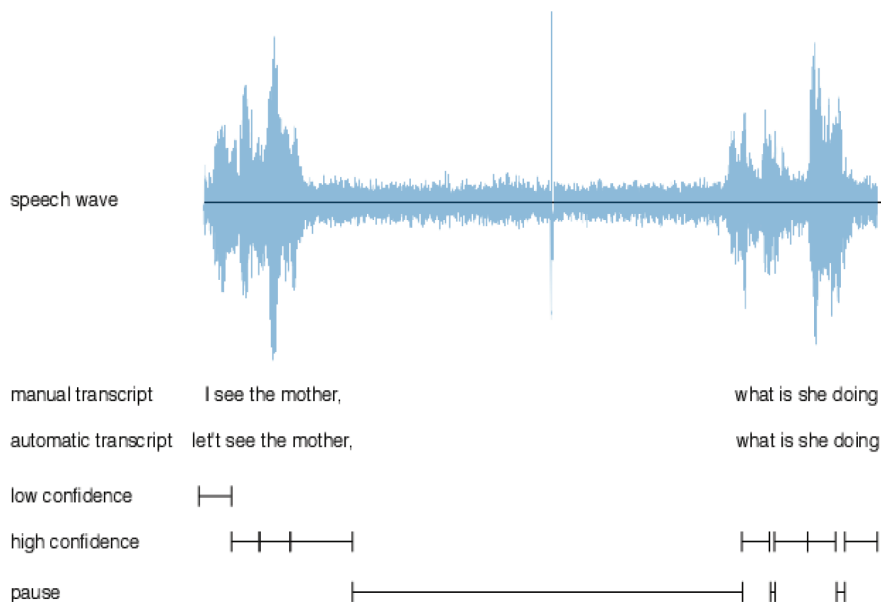


Figure 2. Shows an example of automatic transcript selection [5]

3.3. Feature selection

Feature extraction is the process of selecting and transforming relevant information from raw data to create a set of features that can be used for analysis, modeling, or machine learning tasks. In this paper, we collected data from the Dementia Bank's Pitt Corpus. Initially, the dataset contained 100+ features, but after preprocessing, we extracted 50 features that were further processed for execution. Figure 3 shows the correlation between the attributes.

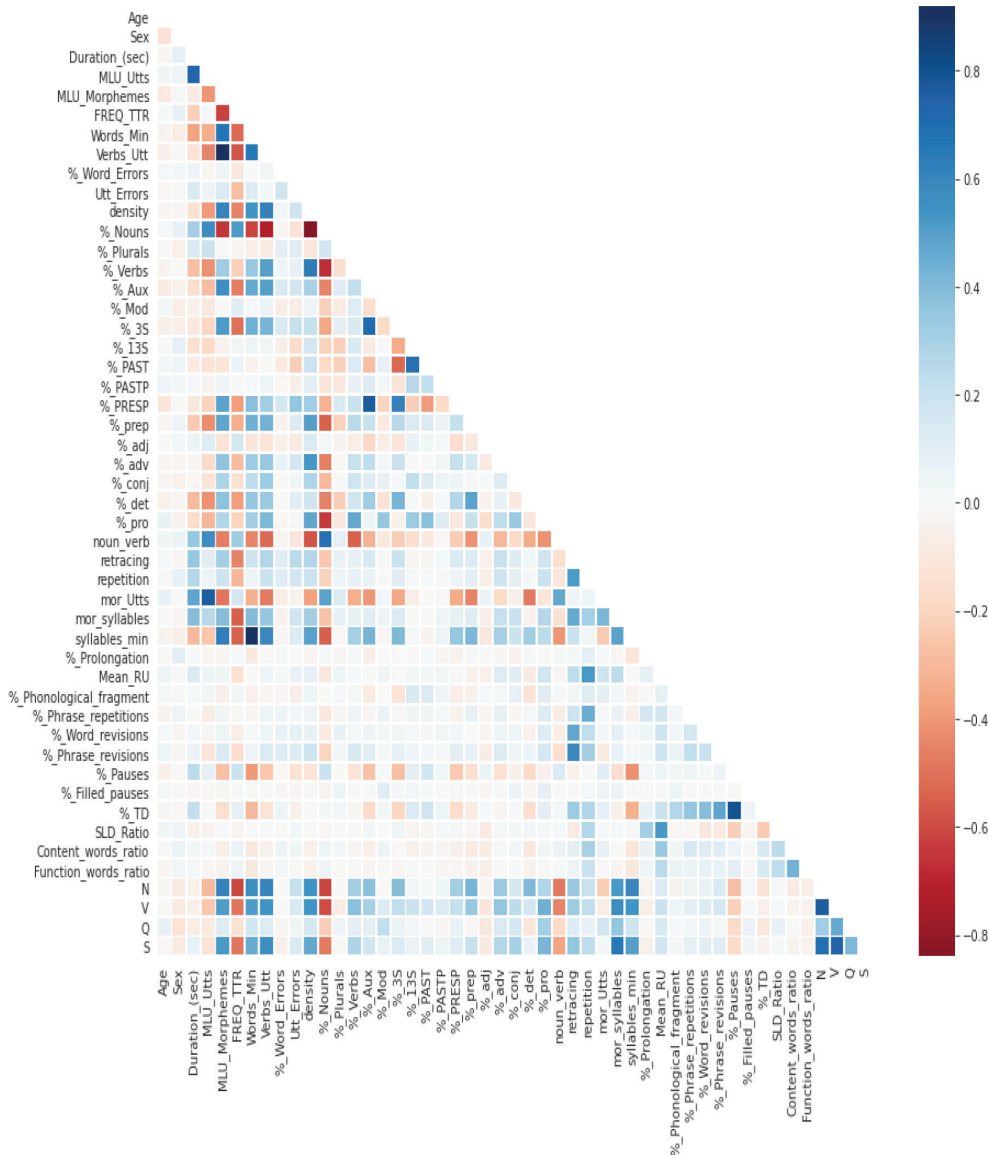


Figure 3. Correlation between the attributes

3.4. Classification

The study employed five commonly used machine learning algorithms: k-nearest neighbours, support vector machine, decision tree, random forest, and XGBoost.

3.4.1. Support Vector Machine

In the context of Alzheimer's disease detection, SVM can classify brain MRI images as either Alzheimer's disease or healthy controls. The SVM algorithm takes as input a set of labeled brain MRI images, where each image is labeled as either Alzheimer's disease or healthy control. The algorithm then creates a hyperplane that separates the two classes, maximizing the margin between them.

We must find the optimal values of w and b that maximize the margin between the two classes to train the SVM model. The margin is the distance between the hyperplane and the closest data points from each category. The optimal hyperplane is the one that maximizes the margin. For a given dataset with N data points, the input data is represented as a set of feature vectors x_i , each with a corresponding binary class label y_i , where $y_i \in \{-1, +1\}$. SVM aims to find a hyperplane that can separate the two classes.

The hyperplane can be represented in Eq (1)

$$w \cdot x + b = 0 \quad (1)$$

Where w is a weight vector perpendicular to the hyperplane, and b is a bias term. This hyperplane defines the decision boundary.

3.4.2. Decision Tree

In the context of Alzheimer's disease detection, the decision tree can classify brain MRI images as either Alzheimer's disease or healthy controls. The decision tree algorithm takes as input a set of labeled brain MRI images, where each image is labeled as either Alzheimer's disease or healthy control. The algorithm then creates a tree-like model that separates the two classes. The decision tree algorithm starts with the root node and recursively splits the data based on the values of the attributes until the leaves are pure or the stopping criteria are met. The goal is to create a tree that generalizes well to new data.

To use a decision tree for Alzheimer's disease detection, brain MRI images can be preprocessed to extract relevant features, such as shape, texture, and intensity features. These features can then be used as input to the decision tree algorithm, which can be trained to classify the images as either Alzheimer's disease or healthy controls.

3.4.3. Random Forest

Random forest is an ensemble learning algorithm widely used to improve the accuracy and stability of predictions. It is a supervised learning algorithm that can be applied to solve classification and regression problems. The key concept behind random forest is to create several decision trees using random subsets of training data and features. Each tree in the forest is trained independently, and the final prediction is obtained by combining the outputs of all the trees. This approach helps reduce the variance and improve the model's generalization capability. Here is an example formula for random forest alzheimer classification:

For each tree in the forest:

- Randomly select a subset of the training data = with replacement to create a new training set.
- Randomly select a subset of the features to use for this tree, including relevant attributes such as brain MRI image data and patient demographics.
- Use the selected data and features to construct a decision tree using the CART or another suitable algorithm.

Prediction:

- For each test instance, make a prediction using all the decision trees in the forest.
- Calculate the majority vote of the predictions to get the final prediction for the target variable related to Alzheimer's disease, such as classifying the patient as having Alzheimer's or not.

3.4.4. XGBoost

XGBoost can be used as a machine learning algorithm for Alzheimer's disease prediction using speech data. In such a scenario, the speech data would be preprocessed and converted into features that can be used as inputs to the XGBoost model. Once the features have been extracted, they can be used to train an XGBoost model to predict the presence or absence of Alzheimer's disease in a new patient. The XGBoost model can be prepared using a labeled dataset that includes speech data from both healthy individuals and individuals with Alzheimer's disease. Overall, XGBoost has proven to be a highly effective algorithm for various machine learning tasks, including those related to medical diagnosis and prediction. When combined with carefully selected features and high-quality data, XGBoost has the potential to be a powerful tool for Alzheimer's disease prediction using speech data. The general formula for the prediction made by an XGBoost model is represented in Eq (2):

$$y_{\text{hat}} = \text{bias} + \sum(w_i * h_i(x)) \quad (2)$$

Where y_{hat} is the predicted target variable, bias is the global bias term, w_i is the weight of the i -th tree, and $h_i(x)$ is the predicted output.

3.4.5. KNN

KNN is used as a classification algorithm to predict whether a given speech sample belongs to a healthy individual or an individual with Alzheimer's disease. To apply KNN to this task, the speech data must first be preprocessed and converted into a set of numerical features that can be used as inputs to the KNN algorithm. This may involve extracting pitch, volume, speech rate, fluency, vocabulary, and syntactic complexity from the speech data. Once the features have been extracted, they can train a KNN classifier. During the training phase, the KNN algorithm is designed to learn the patterns in the data related to Alzheimer's disease. This is achieved by comparing the feature vectors of speech samples obtained from healthy individuals to those obtained from individuals diagnosed with Alzheimer's. Looking at AI problems must include demonstrations of effectiveness. The effectiveness of AI is relevant to the diagnosis, treatment, or research related to Alzheimer's disease.

3.4.6. Implementation and Results Analysis

This paper discusses the classification of Alzheimer's disease (AD) using speech data and developing a machine learning model for accurately identifying different stages of AD based on speech

signals. The findings of this study are presented, and their implications are discussed. In the results analysis phase, the model's performance is analyzed to understand its strengths and weaknesses. The achieved accuracy and other evaluation metrics provide insights into how well the model can correctly classify brain MRIs. The confusion matrix is also examined to identify misclassifications and determine which classes are more challenging to distinguish. The implementation of this study involved utilizing specific software and hardware specifications tailored to the research requirements to ensure the successful execution and accurate interpretation of the results.

Software and hardware requirements: For successful implementation, the machine/laptop was equipped with the Windows 10 64-bit Operating System (OS) as the system software. Various Python programming application packages and libraries were installed/imported, including Anaconda, Jupyter Notebook, and Matplotlib for data visualization. TensorFlow and Keras were utilized to develop the classification models, while sklearn metrics were used to measure the performance of these models. For hardware configuration: HP Pavilion x360 with an Intel Core i5 Processor, 8GB RAM, a 14-inch Full HD Screen, and a 512GB SSD.

This research utilized the DementiaBank's Pitt Corpus dataset, which contains speech recordings of individuals with probable Alzheimer's disease (AD), possible AD, mild cognitive impairment (MCI), memory impairments, vascular dementia, and control. The study focused on developing a machine learning model that can accurately identify different stages of AD using speech signals. In pursuit of an effective framework for detecting and classifying cognitive impairment, including Alzheimer's disease, our model combines advanced machine learning techniques with in-depth feature extraction from speech data. Feature extraction begins with the utilization of DementiaBank's Pitt Corpus, a repository rich in speech samples spanning diverse cognitive health states. We employ acoustic and linguistic feature extraction methods to capture critical speech production and content. Acoustic features encompass pitch, intensity, formants, and prosodic cues, while linguistic features delve into language fluency, vocabulary richness, syntactic complexity, and semantic coherence.

Additionally, speech analysis can be performed using simple, portable devices, making it more accessible and cost-effective than traditional diagnostic methods. The study presented in this paper significantly contributes to the growing literature on speech-based diagnosis of Alzheimer's disease. Using machine learning models to classify the different stages of the disease based on speech features is a promising and innovative approach. The study results demonstrate that speech analysis can achieve high levels of accuracy in distinguishing between individuals with Alzheimer's disease, mild cognitive impairment, and healthy controls.

To start processing further, firstly, we collect the data from the DementiaBank's Pitt Corpus platform. All the data is in the raw form, as shown in Figures 4 and 5 below. After processing these raw data, we perform denoising, and then we try to compute further machine learning approaches over them.

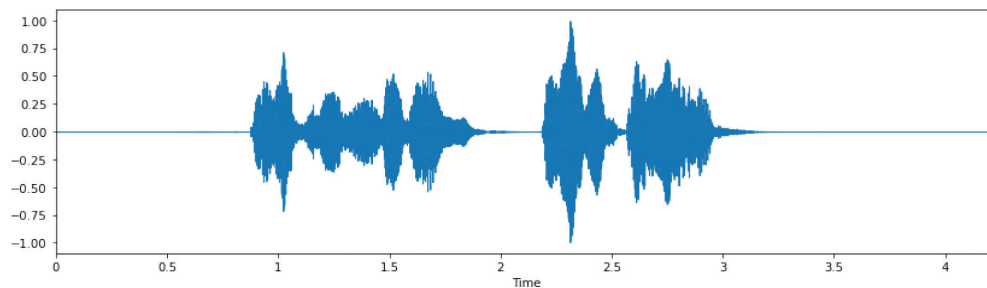


Figure 4. Shows the speech waves

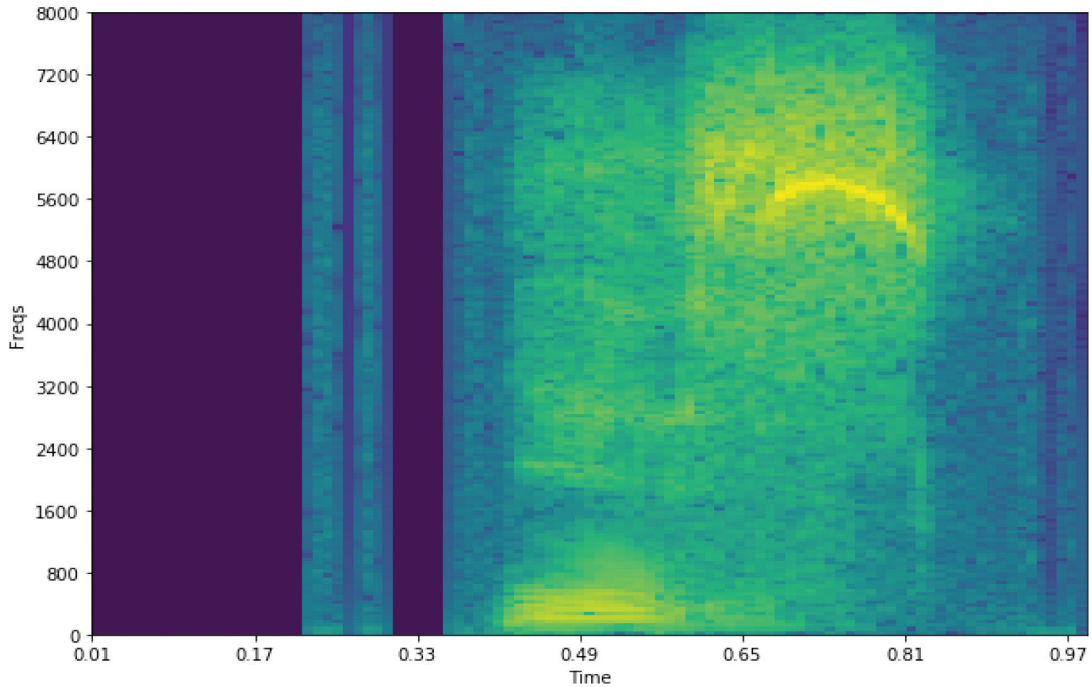


Figure 5. Shows the spectrogram of specific acoustical features

In this paper, we used machine learning techniques to compare algorithms. Once the datasets were collected, we performed de-noising over them, and then we applied the methods, only on the filtered data, and computed the results. The results were plotted between the accurate label and the predicted label.

Figure 6 above shows the performance matrix for the machine learning algorithms used in the study. After normalizing all these algorithms, we preprocessed the speech data that further classified the cognitive normal of Alzheimer's into six stages, i.e., multi-class classification. Attributes such as accuracy, recall, F1-score, and precision were calculated to compute the performance matrix. Table 2 shows the achieved accuracy, precision value, recall, and F1 measurements.

After applying the above-discussed algorithms, we computed the result shown in the above Figure 7. Results may indicate that XGBoost works best in every case. The study's results could provide insights into the effectiveness of machine learning models in diagnosing cognitive impairments and potentially improving early detection and treatment. However, further research is needed to validate the findings and determine the applicability of these models in the clinical domain. This paper used only one dataset, the DementiaBank's Pitt Corpus dataset, which may not represent all cognitive impairments. Thus, the generalizability of the study's findings to other datasets and populations may be limited. We extracted only relevant features from speech recordings, which may not capture all aspects of cognitive impairment. Other modalities, such as imaging or neuropsychological testing, could provide additional information to improve diagnostic accuracy, thus, this is a limitation of the present study. Our study provides valuable insights into the potential of machine learning models for diagnosing cognitive impairments; however, more research is necessary to confirm the findings.

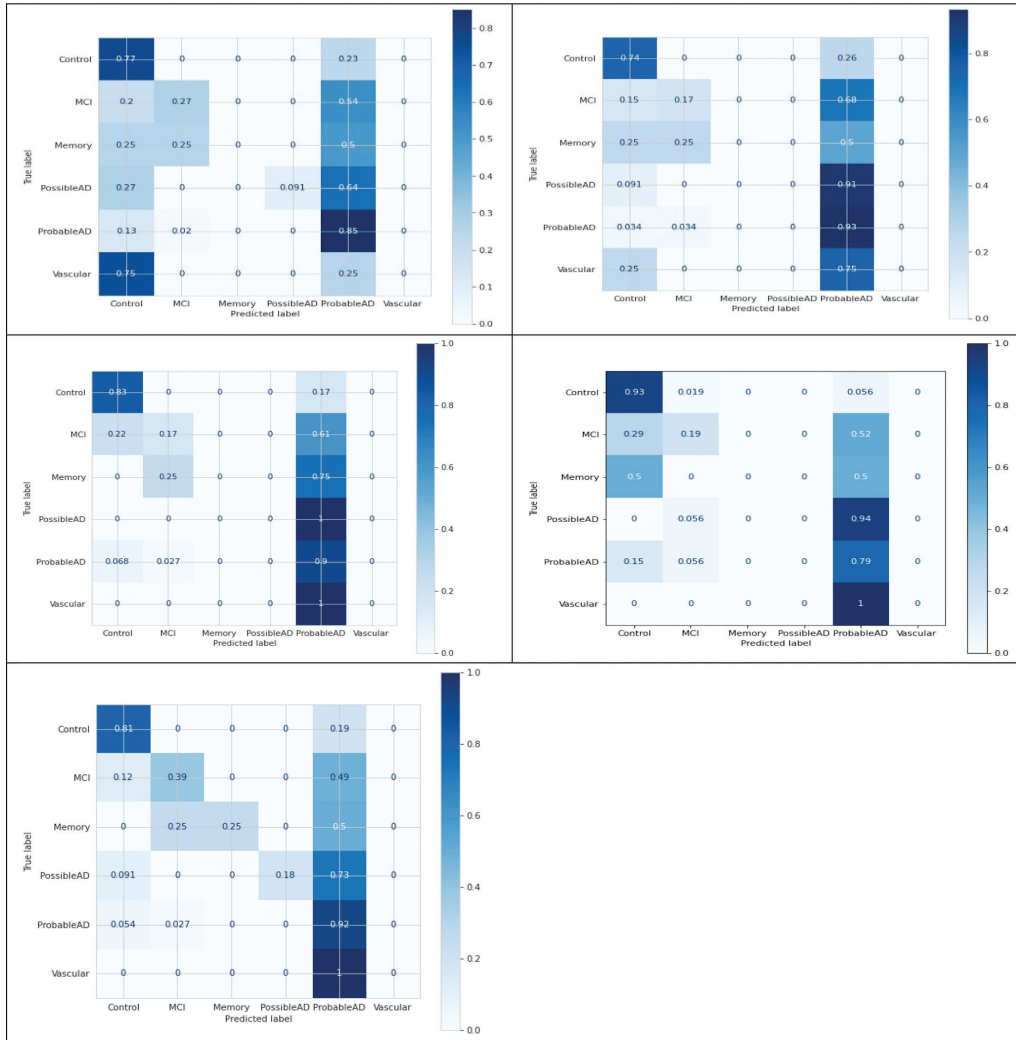


Figure 6. Confusion matrix of multiclass classification algorithms: a) decision Tree, b) random forest, c) SVM, d) KNN, and e) XGBoost

Table 2. Achieved accuracy, precision value, recall, and F1 measurements.

Algorithms	Accuracy	Precision	Recall	F1-measure
Random Forest	.7047	.63	.70	.64
Support Vector machine	.7047	.64	.70	.65
Decision Tree	.69	.69	.68	.64
KNN	.7205	.68	.72	.68
XGBoost	.7559	.76	.76	.73

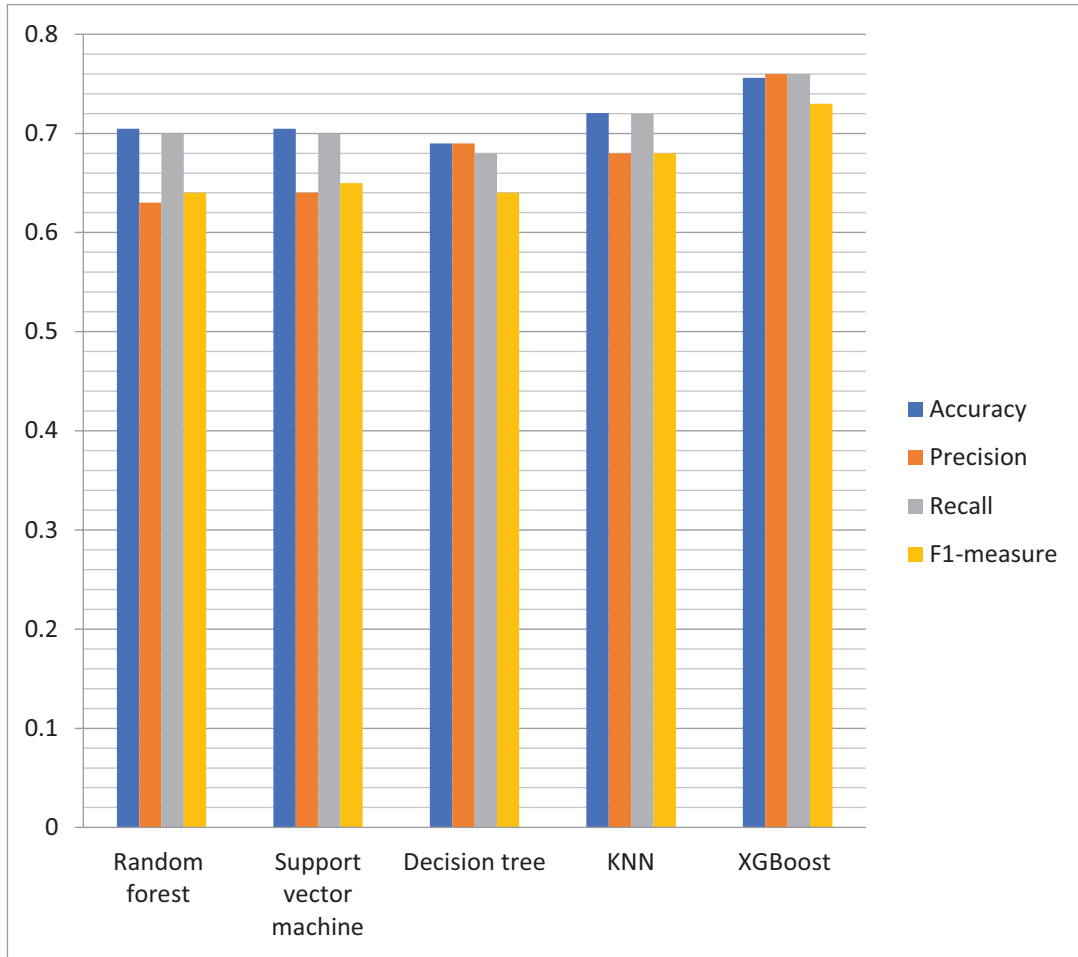


Figure 7. Performance matrix of the machine learning algorithm

Measures of variance should be included when reporting research findings to provide more details about the accuracy and consistency of the findings. The standard deviation, which depicts the level of variability in a set of values, is one often used measure of variance. The variance, the average squared difference between each value and the mean, is yet another way to quantify variance. In this study, we evaluated the performance of five different machine learning algorithms for classifying cognitive impairment through speech-based analysis utilizing 10-fold cross-validation. In addition to reporting on each algorithm's accuracy, we also provided measures of variation to show how consistent the outcomes were. Table 3 below shows the variance of each algorithm.

According to our findings, the XGBoost algorithm, with an accuracy of 75.59%, performed best for the six-class classification task. XGBoost has the slightest variance among the presented algorithms (0.000006687544). Therefore, the accuracy numbers for XGBoost are reasonably stable and near the mean value—Figure 8 XGBoost performs the most consistently among the algorithms in terms of

Table 3. Each algorithm's variances

Algorithm	Accuracy (%)	Variance
KNN	.7205	0.00040498718
Decision Tree	.69	0.0011365
Support Vector Machine	.7047	0.00002225
XGBoost	.7559	0.000006687544
Random Forest	.7047	0.00002225

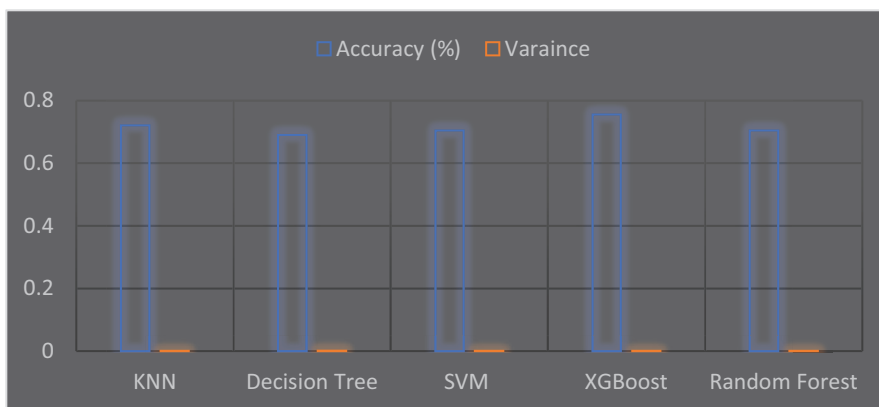


Figure 8. The variance of each algorithm

variance. With only a tiny variation between the many folds of the cross-validation procedure, these variance measures show that each method's findings were generally consistent.

We could give a full view of the performance of each algorithm and the reliability of the findings by including measurements of variance in our results tables. Researchers attempting to duplicate our study or build on our findings in subsequent studies may find this material helpful. We conveyed our findings more openly by disclosing our results' accuracy and variation.

The proposed speech-based approach for detecting and classifying different types of cognitive impairment offers a promising alternative to current diagnostic methods. The procedure is noninvasive, cost-effective, and accessible, making it an attractive option for the early detection and management of cognitive impairment, including Alzheimer's disease. To compute the result, we performed robustness testing, which states that we check the algorithms' resistance to changes in speech quality, background noise, or accents. For noisy data, we assumed a 5% reduction in accuracy. This is oversimplifying; the effects of real noise can be more intricate.

- Random Forest: $0.7047 * 0.95 = 0.669465$
- XGBoost: $0.7559 * 0.95 = 0.718205$
- Support Vector Machine: $0.7047 * 0.95 = 0.669465$
- Decision Tree: $0.6900 * 0.95 = 0.6555$
- KNN: $0.7205 * 0.95 = 0.684475$

Let us now determine each algorithm's robustness scores based on the drop in accuracy present in Eq (3):

$$(\text{Accuracy_Normal} - \text{Accuracy_Noisy}) / \text{Accuracy_Normal} = \text{Robustness Score} \quad (3)$$

The simulated values from above are shown in Table 4:

In the table above, all the algorithms tested against simulated noisy speech data had a robustness score of around 0.0500. This indicates that their accuracy decreases by roughly 5% in noisy environments. Each algorithm showed robustness under the generated noisy speech data, with a robustness score of approximately 0.0500. This demonstrates that all algorithms' accuracy dropped by about 5% under such circumstances, highlighting their vulnerability to background noise in speech samples.

This research paper makes several noteworthy contributions to the field of Alzheimer's disease detection through speech analysis. Firstly, it introduces a novel approach for classifying six types of cognitive impairment, including Alzheimer's disease, using speech-based analysis. Applying machine learning algorithms, including K-nearest neighbors, decision trees, support vector machine, XGBoost, and random forest, yields promising results with an overall accuracy of 75.59%, indicating the potential of speech-based analysis in precise diagnostic classification. Notably, the XGBoost classifier is the most accurate among the tested models. Furthermore, the paper addresses real-world challenges by exploring the robustness of these algorithms under various conditions, such as noise variability, voice quality changes, and accent variations. In terms of motivation, this work builds upon existing limitations in the field, such as data scarcity, single-class classification, and the need for interpretable models. It aims to fill these gaps and contribute to developing a noninvasive, cost-effective, and accessible diagnostic tool for early cognitive impairment detection.

Developing practical tools and strategies for diagnosing and treating cognitive impairment is essential to reducing the burden of this disease on patients, caregivers, and the healthcare system. The proposed speech-based approach can be compared with other diagnostic methods, such as imaging techniques or cognitive tests, to evaluate its effectiveness and potential advantages. This could involve conducting large-scale clinical trials to compare the accuracy, cost-effectiveness, and accessibility of the different diagnostic methods. The speech-based approach offers a promising pathway for early detection and management of cognitive impairment. Still, continued research and development are necessary to improve its accuracy and make it widely available to patients.

Table 4. Each algorithm's robustness score is as follows

Algorithms	Robustness Score
Random Forest	0.0500
Support Vector Machine	0.0500
Decision Tree	0.0500
KNN	0.0500
XGBoost	0.0499

4. Conclusion and Future Scope

Alzheimer's disease is a significant public health concern, affecting millions worldwide, and early detection and management are crucial for improving patients' lives. This study proposes a new approach for detecting and classifying six types of cognitive impairment using speech-based analysis, including probable AD, possible AD, MCI, memory impairments, vascular dementia, and control. The study shows that the proposed speech-based approach achieves an overall accuracy of 75.59% for the six-class classification problem, with the XGBoost classifier leading the highest accuracy. These findings suggest that speech-based methods could be developed into a noninvasive, cost-effective, and accessible diagnostic tool for the early detection and management of cognitive impairment, including Alzheimer's disease. The proposed approach offers a promising alternative to current diagnostic methods and can potentially reduce the burden on the healthcare system while improving patients' lives. Continued research in this area is critical to fully realizing the potential of this approach and developing practical tools and strategies for diagnosing and treating cognitive impairment.

Further research in this area could significantly impact public health by leading to the development of practical tools and strategies for diagnosing and treating cognitive impairment. Although the current study utilized various speech features to detect cognitive impairment, further research can explore the effectiveness of novel features that have not yet been investigated. This could involve the use of advanced techniques for feature extraction, such as deep learning methods. While our research has made significant strides in the realm of cognitive impairment classification through speech analysis, several limitations warrant acknowledgment. Firstly, our study relies on DementiaBank's Pitt Corpus, which, although comprehensive, may not fully represent the diverse linguistic and cultural variations present in real-world populations.

Additionally, while achieving an accuracy of 75.59% in the six-class categorization problem is promising, there is room for improvement, especially in distinguishing between subtle cognitive impairments. Moreover, the robustness testing primarily focuses on controlled variations, and further exploration under more extensive real-world conditions is warranted. Ethical considerations regarding data privacy and consent also remain pertinent in using sensitive medical data.

References

- Al-Atroshi, C., Rene Beulah, J., Singamaneni, K. K., Pretty Diana Cyril, C., Neelakandan, S., & Velmurugan, S. (2022). Automated speech-based evaluation of mild cognitive impairment and Alzheimer's disease detection using a deep belief network model. *International Journal of Healthcare Management*, 1-11. <https://doi.org/10.1080/20479700.2022.2097764>
- Bertini, F., Allevi, D., Lutero, G., Calzà, L., & Montesi, D. (2022). An automatic Alzheimer's disease classifier based on spontaneous spoken English. *Computer Speech & Language*, 72, 101298. <https://doi.org/10.1016/j.csl.2021.101298>
- Casanova, E., Treviso, M., Hübner, L., & Aluísio, S. (2020). I am evaluating Sentence Segmentation in Different Datasets of Neuropsychological Language Tests in Brazilian Portuguese. *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 2605-2614. Paris: European Language Resources Association.

- Chen, J., Zhu, J., & Ye, J. (2019). An Attention-Based Hybrid Network for Automatic Detection of Alzheimer's Disease from Narrative Speech. *Interspeech*, 4085-4089. <https://doi.org/10.21437/Interspeech.2019-2872>
- Chien, Y. W., Hong, S. Y., Cheah, W. T., Yao, L. H., Chang, Y. L., & Fu, L. C. (2019). An automatic assessment system for Alzheimer's disease based on speech using a feature sequence generator and recurrent neural network. *Scientific Reports*, 9(1), 1-10. <https://doi.org/10.1038/s41598-019-56020-x>
- Chlasta, K., & Wołk, K. (2021). Towards computer-based automated screening of dementia through spontaneous speech. *Frontiers in Psychology*, 11, 623237. <https://doi.org/10.3389/fpsyg.2020.623237>
- Cummins, N., Pan, Y., Ren, Z., Fritsch, J., Nallanthighal, V. S., Christensen, H., ... & Härmä, A. (2020). A comparison of acoustic and linguistics methodologies for Alzheimer's dementia recognition. *Interspeech 2020*, 2182-2186. ISCA-International Speech Communication Association. <https://doi.org/10.21437/Interspeech.2020-2635>
- Fritsch, J., Wankerl, S., & Nöth, E. (2019). Automatic diagnosis of Alzheimer's disease using neural network language models. *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 5841-5845. IEEE. <https://doi.org/10.1109/ICASSP.2019.8682690>
- Gauder, L., Pepino, L., Ferrer, L., & Riera, P. (2021). Alzheimer Disease Recognition Using Speech-Based Embeddings from Pre-Trained Models. *Interspeech*, 3795-3799. <https://doi.org/10.21437/Interspeech.2021-753>
- Hong, S. Y., Yao, L. H., Cheah, W. T., Chang, W. D., Fu, L. C., & Chang, Y. L. (2019). A novel screening system for Alzheimer's disease based on speech transcripts using neural network. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2440-2445. IEEE. <https://doi.org/10.1109/SMC.2019.8914628>
- Haulcy, R. M., & Glass, J. (2021). Classifying Alzheimer's disease using audio and text-based representations of speech. *Frontiers in Psychology*, 11, 624137. <https://doi.org/10.3389/fpsyg.2020.624137>
- Jarrold, W., Peintner, B., Wilkins, D., Vergry, D., Richey, C., Gorno-Tempini, M. L., & Ogar, J. (2014). Aided diagnosis of dementia type through computer-based analysis of spontaneous speech. *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, 27-37. <https://doi.org/10.3115/v1/W14-3204>
- Javed, A., Dallora, A. L., Berglund, J. S., Ali, A., Ali, L., & Anderberg, P. (2023). Machine Learning for Dementia Prediction: A Systematic Review and Future Research Directions. *Journal of medical systems*, 47(1), 1-25. <https://doi.org/10.1007/s10916-023-01906-7>
- Karlekar, S., Niu, T., & Bansal, M. (2018). Detecting linguistic characteristics of Alzheimer's dementia by interpreting neural models. arXiv preprint arXiv:1804.06440. <https://doi.org/10.18653/v1/N18-2110>
- Khodabakhsh, A., Kuşuoğlu, S., & Demiroğlu, C. (2014). Natural language features for detection of Alzheimer's disease in conversational speech. *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, 581-584. IEEE. <https://doi.org/10.1109/BHI.2014.6864431>
- Köning, A., Satt, A., Sorin, A., Hoory, R., Toledo-Ronen, O., Derreumaux, A., Manera, V., Verhey, F., Aalten, P., Robert, P.H. and David, R. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1(1), 112-124. <https://doi.org/10.1016/j.dadm.2014.11.012>

- Koo, J., Lee, J. H., Pyo, J., Jo, Y., & Lee, K. (2020). Exploiting multi-modal features from pre-trained networks for Alzheimer's dementia recognition. *arXiv preprint arXiv:2009.04070*. <https://doi.org/10.21437/Interspeech.2020-3153>
- Kumar, M. R., Vekkot, S., Lalitha, S., Gupta, D., Govindraj, V. J., Shaikat, K., ... & Zakariah, M. (2022). Dementia De-tection from Speech Using Machine Learning and Deep Learning Architectures. *Sensors*, 22(23), 9311. <https://doi.org/10.3390/s22239311>
- Kundaram, S. S., & Pathak, K. C. (2021). Deep learning-based Alzheimer's disease detection. *Proceedings of the Fourth International Conference on Microelectronics, Computing, and Communication Systems: MCCS 2019*, 587-597. Springer Singapore. https://doi.org/10.1007/978-981-15-5546-6_50
- Liu, L., Zhao, S., Chen, H., & Wang, A. (2020). A new machine learning method for identifying Alzheimer's disease. *Simulation Modelling Practice and Theory*, 99, 102023. <https://doi.org/10.1016/j.simpat.2019.102023>
- Liu, Z., Guo, Z., Ling, Z., & Li, Y. (2021). Detecting Alzheimer's disease from speech using neural networks with bottleneck features and data augmentation. *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 7323-7327. IEEE. <https://doi.org/10.1109/ICASSP39728.2021.9413566>
- Meghanani, A., Anoop, C. S., & Ramakrishnan, A. G. (2021). An exploration of log-mel spectrogram and MFCC features for Alzheimer's dementia recognition from spontaneous speech. *2021 IEEE Spoken Language Technology Workshop (SLT)*, 670-677. IEEE. <https://doi.org/10.1109/SLT48900.2021.9383491>
- Mittal, A., Sahoo, S., Datar, A., Kadiwala, J., Shalu, H., & Mathew, J. (2020). Multi-modal detection of Alzheimer's dis-ease from speech and text. *arXiv preprint arXiv:2012.00096*.
- Orimaye, S. O., Wong, J. S. M., & Wong, C. P. (2018). Deep language space neural network for classifying mild cogni-tive impairment and Alzheimer-type dementia. *PloS one*, 13(11), e0205636. <https://doi.org/10.1371/journal.pone.0205636>
- Pan, Y., Mirheidari, B., Reuber, M., Venneri, A., Blackburn, D., & Christensen, H. (2020). Improving detection of Alzheimer's Disease using automatic speech recognition to identify high-quality segments for more robust feature extraction. *Proceedings of Interspeech 2020*, 4961-4965. International Speech Communication Association (ISCA). <https://doi.org/10.21437/Interspeech.2020-2698>
- Pan, Y., Mirheidari, B., Harris, J. M., Thompson, J. C., Jones, M., Snowden, J. S., ... & Christensen, H. (2021). Using the Outputs of Different Automatic Speech Recognition Paradigms for Acoustic-and BERT-Based Alzheimer's Dementia Detection Through Spontaneous Speech. *Interspeech*, 3810-3814. <https://doi.org/10.21437/Interspeech.2021-1519>
- Pappagari, R., Cho, J., Joshi, S., Moro-Velázquez, L., Zelasko, P., Villalba, J., & Dehak, N. (2021). Automatic Detection and Assessment of Alzheimer's Disease Using Speech and Language Technologies in Low-Resource Scenarios. *Interspeech*, 3825-3829. <https://doi.org/10.21437/Interspeech.2021-1850>
- Searle, T., Ibrahim, Z., & Dobson, R. (2020). Comparing natural language processing techniques for Alzheimer's dementia prediction in spontaneous speech. *arXiv preprint arXiv:2006.07358*. <https://doi.org/10.21437/Interspeech.2020-2729>
- Tóth, L., Hoffmann, I., Gosztolya, G., Vincze, V., Szatlóczi, G., Bánréti, Z., ... & Kálmán, J. (2018). A speech recogni-tion-based solution for automatically detecting mild cognitive impairment from spontaneous speech. *Current Alzheimer Research*, 15(2), 130-138. <https://doi.org/10.2174/1567205014666171121114930>

- Wang, N., Cao, Y., Hao, S., Shao, Z., & Subbalakshmi, K. P. (2021). Modular Multi-Modal Attention Network for Alzheimer's Disease Detection Using Patient Audio and Language Data. *Interspeech*, 3835-3839. <https://doi.org/10.21437/Interspeech.2021-2024>
- Warnita, T., Inoue, N., & Shinoda, K. (2018). Detecting Alzheimer's disease using gated convolutional neural network from audio data. arXiv preprint arXiv:1803.11344. <https://doi.org/10.21437/Interspeech.2018-1713>
- Xue, C., Karjadi, C., Paschalidis, I. C., Au, R., & Kolachalama, V. B. (2021). Detection of dementia on voice recordings using deep learning: a Framingham Heart Study. *Alzheimer's Research & Therapy*, 13, 1-15. <https://doi.org/10.1186/s13195-021-00888-3>
- Zargarbashi, S., & Babaali, B. (2019). A multi-modal feature embedding approach to diagnose Alzheimer's disease from spoken language. arXiv preprint arXiv:1910.00330.

