Enhance communication for dementia patients: supervised and sequential learning to identify dementia behavior and overcoming speech incompletion.

By
Arif Faisal
211-15-14650

Md. Asif Akram
211-15-14640

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the **Degree of Bachelor of Science in Computer Science and Engineering**

Supervised by
Ms. Sharun Akter Khushbu
Lecturer (Senior Scale)
Department of Computer Science and EngineeringDaffodil International University

Co-Supervised by
Ms. Rabeya Khatun
Lecturer
Department of Computer Science and EngineeringDaffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

January 12, 2025

APPROVAL

This Project titled "Enhance communication for dementia patients: supervised and sequential learning to identify dementia behavior and overcoming speech incompletion", submitted by **Arif Faisal**, ID No: **211-15-14650** and Md Asif Akram, ID No: **211-15-14640** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 12 January, 2025.

BOARD OF EXAMINERS

12|01|2025

Dr. Md. Taimur Ahad
Associate Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Chairman

Mushfiqur Rahman
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Mr. Rahmatul Kabir Rasel Sarker
Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Sadat Hasan
Data Scientist (Senior Principal Officer)
Risk Management Division
BRAC Bank

External Examiner

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Ms. Sharun Khushbu, Lecturer (Senior Scale)**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

12.1.25

Ms. Sharun Akter Khushbu
Lecturer (Senior Scale)
Department of Computer Science and
Engineering, Daffodil International University

Co-Supervised by:

Rabeya 12.01.25

Ms. Rabeya Khatun
Lecturer
Department of Computer Science and
Engineering, Daffodil International University

Submitted by:

Faisal 12-01-25

Arif Faisal
Student ID: 211-15-14650
Department of Computer Science and
Engineering, Daffodil International University

Asif 12.1.25

Md. Asif Akram
Student ID: 211-15-14640
Department of Computer Science and
Engineering, Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project (FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Ms. Sharun Akter Khushbu, Lecturer (Senior Scale),** Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Machine Learning, Deep Learning, Natural Language Processing** carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting themat all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to Dr. Sheak Rashed Haider Noori Professor & Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

This study examines fundamental autonomous speech parameters and incorporates text analysis to detect dementia from voices, while also assisting dementia patients by providing sentence auto completion support to address their communication challenges. To accomplish these 740 voice recordings (370 dementia and 370 non-dementia) were collected. Features such as MFCCs and RMS were extracted from the audio and text data. Various machine learning models, including Random Forest (RF), Logistic Regression (LR) and Gradient Boosting (XGBoost), alongside a deep learning Long Short-Term Memory (LSTM) were trained. Among these, the LSTM model achieved the highest accuracy of 92.93%. The recorded voices were transcribed into text using Whisper model, and TF-IDF trigram features were extracted for detection. The models were implemented for text classification, with LR and LSTM achieving the best accuracies of 72.43% and 72.78% respectively. For sentence auto completion, a Bi-directional LSTM (Bi-LSTM) model with N-gram sequences was implemented and achieved 20.8% accuracy. This research highlights the integration of speech and text-based methods to analyze and detect dementia and assist dementia patients through sentence auto completion.

Keywords: Dementia, Cognitive Decline, Speech Analysis, Text Classification, Deep Learning, Machine Learning, MFCCs, TF-IDF, Bi-LSTM, Sentence Completion.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

'Dementia is not seen as a medical condition but a normal problem of ageing – No, it is not. Dementia is devastating not only for individuals who suffer from it but also for their family and friends.' – World Alzheimer's Report 2010 [1]

1.1 Introduction

Dementia is a neurological disorder that comes along with aging and results in an impairment in memory, thinking and reasoning capabilities. By 2050, approximately 15 million people will live in Bangladesh are age of 60 years or older. Based on a Bangladesh Bureau of Statistics (BBS) research, dementia is a serious public health [4]. Worldwide about 57.6 million individuals are now affected by dementia [5]. A large number of people with dementia cases worldwide [6]. The impact of dementia in the South Asia region becomes evident when we observe these findings. Almost 4.8 million people in South Asia are affected by this symptom [7]. These numbers highlight the urgent need for increased awareness, early diagnosis, and effective intervention to address this condition. Elderly individuals are the main sufferers of dementia, that is not a typical aspect of aging. Early identification of this chronic brain condition is essential, it may result in a major impact on everyday activities. Language, assessment and learning abilities are all affected. Rather than a specific disease, this health issue is a set of symptoms that may arise because of multiple neurological disorders. This occurs as cells in the brain lose their ability to interact with one another. As the disorder progresses, one's ability to function independently can suffer significantly. Communication difficulties are among the most serious challenges that people with dementia face, often resulting in frustration, social withdrawal, and isolation. These problems become more common as the disease progresses [8]. Patients become stuck in the middle of a sentence, making it difficult to communicate clearly. According to research, patients face significant communication challenges in the late stages of the disease, which affects their social interactions [9]. Furthermore, communication impairments are frequently linked to an increased risk of anxiety and depression. However, dementia patients often struggle to express their needs or participate in conversations, which negatively affects their mental and emotional well-being. Proper identification and successful treatment of dementia depends on a comprehensive understanding of the condition. In this study, we explore how machine learning and natural language processing (NLP) can be applied to help dementia patients or elderly people overcome communication barriers. By developing models that can predict individuals that are suffering from dementia.

When a person gets stuck mid-speech and cannot complete their sentences, we aim to improve their communication skills and reduce the frustration associated with cognitive decline.

1.2 Motivation

This study was motivated by the significant difficulties that elderly people in Bangladesh face, particularly those suffering from dementia. In many cultures, including Bangladesh, dementia symptoms are dismissed as a normal sign of aging, and growing older is frequently associated with forgetfulness. In Bangladeshi culture, dementia symptoms are often dismissed as a normal part of getting older. This misconception causes dementia to be underdiagnosed and undertreated, denying patients the care they require. Unfortunately, society frequently misjudges these individuals, isolating or neglecting them as a result. This can be a particularly difficult time for their families. Dementia patients frequently struggle to communicate, forget important information, and lose their ability to express themselves effectively. This leads to a cycle of exhaustion and frustration, as loved ones get confused and are unclear about whether to assist or appreciate them. In Bangladesh, where appropriate dementia care and treatment are limited. Families are left to shoulder the entire bear the full burden both emotionally and physically. The World Health Organization (WHO) recently published a report that has a striking impact on this research. Their data in 2020 shows an alarming reality: the number of Alzheimer's and dementia-related deaths in Bangladesh had reached 14,993 [10]. It is placing the country at 142nd globally [11]. Alzheimer's disease, a primary form of dementia, continues to be a major health issue worldwide. In 2021, Alzheimer's disease accounted for 3.4% of all deaths in the United States, ranking 7th among leading cause of death in US. With the elderly population expanding at an annual rate of 4%, the impact of dementia will continue to worsen. That is why the need for early detection, proper care, and moral support is more essential than ever. This study aims to figure out these issues by building methods to detect dementia from speech as well as text from the Bangla language. By improving early identification and communication, we can offer essential assistance to dementia patients and their families, assisting them in navigating the complexities of this condition. Ultimately the objective of this work is to raise awareness of marginalized issues, challenge the societal issues associated with dementia, and improve the quality of life for those who suffer in silence.

1.3 Objectives

The objective of this study is to differentiate between dementia and non-dementia and also improve speech perception via predicting speech with completion in real-time, with a particular focus on the development of technologies that assist dementia patients in communication. By analyzing speech patterns and the application of machine learning and deep learning techniques, this research will not only help in identifying and recognizing dementia but also improve the quality of life for individuals suffering from cognitive decline. The primary objective of this research includes:

i. Enhance speech patterns of dementia patients aged 60-70 for everyday communication.
ii. Simplified Analysis Framework for facilitating synthetic analysis of recorded speech data.

1.4 Methodology

This research focuses on detecting dementia through speech and text analysis and providing sentence auto-completion support for dementia patients. Initially, the research was initiated by consulting with a medical professional, who delivered guidance on the basic methodology. As per the recommendations, we focused on the fundamental requirements of individuals with dementia. Subsequently, sentences were composed to correspond with the phrases and expressions that these individuals apply daily. Then collect the audio recordings (370 dementia and 370 non-dementia) . The recordings were processed using "Audacity" for trimming to enhance quality. Key features such as MFCCs and RMS Energy were extracted using the Librosa library. These features were structured into a dataset and used to train machine learning models (Logistic Regression, Random Forest, Gradient Boosting) and a deep learning model (LSTM). The audio recordings were transcribed into text, followed by a systematic correction process to improve transcription accuracy. Preprocessing included tokenization, removal of unnecessary symbols, and TF-IDF trigram feature extraction. These features were used to train machine learning models (Logistic Regression, Random Forest, and SVM) and a deep learning model (LSTM) for text classification, enabling the distinction between dementia and non-dementia sentences. For sentence auto-completion, a Bi-LSTM model was implemented. Sentences were tokenized, converted into n-gram sequences, and padded for uniform input length. The Bi-LSTM model incorporates embedding layers and bidirectional LSTM layers. The model was trained to predict meaningful word completions and assist dementia patients in real-time communication

1.5 Project Outcome

The outcomes of this research contribute significantly to understanding and supporting mechanisms for dementia patients, focusing on their communication challenges and linguistic analysis. The study examined innovative approaches to address the cognitive and speech problems that are commonly associated with dementia, providing useful solutions for both patients and caregivers. To detect dementia, a combination of machine learning models and a deep learning-based LSTM, Bi-LSTM models were employed. These models effectively analyzed speech and textual data to identify patterns suggestive of dementia related cognitive decline. Linguistic feature analysis was conducted to provide crucial insights into the cognitive deficiencies of dementia patients. The primary focus of the research was on developing a predictive aid system for sentence completion, which would help dementia patients who sometimes face difficulties in completing sentences due to memory lapses. This technique directly contributes to the quality of life of dementia patients and decreases the stress of caregivers. In addition, a systematic methodology for assessing recorded speech data was designed, Enabling the identification of dementia-specific patterns. This framework provides a well-structured approach to future analysis and development in cognitive disorder analysis.

1.6 Organization of the Report

This report is organized into six chapters. It comprehensively documents the research journey from identifying the problem to proposing and evaluating the solutions.

Chapter 1 Introduction: Where discusses why dementia is a major problem in Bangladesh and how patients are treated. This was the main motivation behind this research. Here it also outlines the objectives and expected outcomes of this study.

Chapter 2 background: Here provides comprehensive analysis between existing works and identifying gaps particularly for the context of Bangla language.

Chapter 3 Methodology: In this chapter describes the steps taken during the research such as data preparation, transcription, preprocessing, and the implementation of machine learning and deep learning models for speech and text classification, as well as sentence auto-completion.

Chapter 4 Implementation and result: This chapter presents the performance metrics and result analysis for speech classification, text classification and sentence auto completion. Here it also provides insight into the effectiveness of the proposed approach.

Chapter 5 Engineering standards and design: Here investigates the compatibility of this research with engineering principles, mapping with knowledge profiles and discusses the challenges faced during the study.

Chapter 6 Conclusion: This last chapter summarizes the findings, highlighting the contributions of this study. Here outlines the potential future directions to improve and expand this research.

Chapter 2

Background

This chapter summarizes the essential background information and relevant required to comprehend the research. It analyses the technologies, methodologies, and existing research that are associated with the identification of dementia and sentence completion.

2.1 Introduction

Detecting dementia through text is a substantial body of work in other languages, particularly English, that explores speech and text-based dementia detection, similar studies for Bangla are lacking. However, existing studies on sentiment analysis, emotion detection, and sentence prediction in Bangla provide valuable groundwork and methodologies that can be adapted to this domain. Below is a review of relevant studies that, while not directly addressing dementia in Bangla, share similar methodologies and approaches.

2.2 Literature Review

Table 2.1: Summary of similar literature reviews.

Authors	Year	Title	Key Findings	Methodology
[13] Rifat, R., Khan, M. S. A., & Hasan, M. K.	2023	"A Comparative Study on Bengali Speech Sentiment Analysis Based on Audio Data"	7000 audio, & emotions. Features: MFCC, Zero Crossing Rate (ZCR), Chroma.	Random Forest, KNN, AdaBoost. RF achieved highest 90% accuracy
[14] Zhu, Y., Lin, N., Balivada, K. S., Haehn, D., & Liang, X.	2024	"Adversarial Text Generation using Large Language Models for Dementia Detection"	Task-specific feature contexts, attention to detail, language clarity	Adversarial Text Generation (ATG) accuracy 85.42%

[16] A. Rianti, S. Widodo, A. D. Ayuningtyas, and F. B. Hermawan	2022	"NEXT WORD PREDICTION USING LSTM"	180 Indonesian destinations about 4-7 words, used LSTM sequential one layer model	LSTM accuracy 75%
[17] Islam, M. R., Al Amin, & Zereen, A. N.	2024	Enhancing Bangla Language Next Word Prediction and Sentence Completion through Extended RNN with Bi-LSTM Model On N-gram Language	1.7 GB dataset from Bdnews24, Prothom Alo, BBC news Bangla. In methodology used Bi-LSTM.	Bi-LSTM accuracy in 4-grams (99%) and 5-grams (96.30%)

Despite the shortage of direct research on the identification of dementia from Bangla speech, numerous studies in other languages have highlighted the possibility of voice analysis for the early identification of dementia [18]. Using a dataset of 1264 voice recordings, the investigation" Detection on voice recordings using deep learning: a Framingham Heart Study" involved 330 dementia cases, 483 regular activities also 451 people with mild cognitive impairment. The procedure for extracting features was carried out using Mel-Frequency Cepstral Coefficients (MFCCs) alongside Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). The CNN model surpassed the LSTM model, which had an AUC of 0.740 and balanced accuracy of 64.7%, with an AUC of 0.805 and balanced accuracy of 74.3%. Thus, The CNN model revealed better results. This investigation demonstrates the possibility of deep learning for real-time dementia assessment, despite further validation with a variety of datasets, involving those in Bangla, is necessary [19. The Random Forest model achieved a maximum accuracy of 87.6% within the machine learning models, whereas with an accuracy of 85%, the PRCNN model surpassed the deep leaning models. The effectiveness of voice-based biomarkers and the stability of machine leaning models for identifying dementia are demonstrated in this research. The methodologies implemented in these investigations can be utilized for identical research in Bangla but there is some unavailability of resources [20]. Fortunately, identification of dementia in Bangla text has not been explored yet, there are works that focus on text classification such as sentiment analysis and emotion detection and these can be used for dementia research. The collection of data used in the research "Sentiment analysis of Bangla and Romanized Bangla Text using LSTM Networks" contains 9,337 instances that were classified into Positive, Negative and Ambiguous categories. The LSTM network has illustrated the usefulness of LSTM networks for analyzing text in Bangla by obtaining 78% success rate for binary classification and 55% for three-class classification. Despite knowing that this research works on sentiment analysis, its adoption of LSTM networks provides a relevant methodology for future investigation on the identification of dementia from Bangla language. The use of LSTM networks allows the detection of linguistic indicators related to cognitive impairment in text [21]. While there is research on sentence completion or predicting the next word in Bangla, none of this research focuses on dementia related text [22]. A research paper titled "Word Completion and Sequence Prediction in Bangla Language Using Trie and a Hybrid Approach of Sequential LSTM and N-gram" discovers the obstacles of sentence completion and sequence prediction in Bangla, an area with limited works in compared to other languages. The outcomes were positive, achieving an accuracy of 84% with the smaller dataset and the larger dataset achieved an accuracy of 81%. Eventually these models were not dementia focused but their methodology could be used to predictive text system in Bangla, for example sentence completion for people with dementia [23]. A dataset of 0.25 million words from a renowned newspaper "Prothom Alo" was used in an additional work on Bangla word prediction. To accurately predict each syllable in a sentence, with the use of five models: unigram, bigram, trigram, backoff and deleted interpolation. These results illustrate the possibility of stochastic language models in Bangla natural language processing, which might facilitate the advancement of word and sentence prediction systems in the field of identification of dementia.

2.3 Gap Analysis

Most of the current study tend to focus on primarily on widely spoken languages such as English. Even though some studies carried out by Das and Bandyopadhyay have examined the classification of Bangla text [20]. In similar ways, research carried out by Xue et al. and Kumar et al. investigates the identification of dementia through voice, but their findings are only applicable to datasets that are not part of Bangla [18][19]. There is a critical void in multilingual dementia research due to the absence of work concentrating on the identification of dementia using Bangla voice data. In addition, there have been no prior studies that have attempted to create sentence auto-completion systems to address their communicational challenges.

2.4 Summary

These reviews show significant progress in dementia detection using voice and text data, but there is a notable absence of research focusing on Bangla. Existing studies on dementia detection from voice primarily utilize datasets in English or other languages, leaving Bangla voice data unexplored. While methodologies such as MFCCs and deep learning architecture demonstrate potentiality and can be applied to Bangla specific dataset. Similarly, research in text analysis has focused on tasks such as sentiment classification and emotion detection in Bangla, employing models like LSTM and Bangla-BERT. These studies were impactful but not specific to dementia-related challenges. Works on sentence completion and word prediction in Bangla have explored hybrid approaches, but lack of potentiality on dementia-specific text or suggest the communication difficulties faced by dementia patients. The gap highlights the need for research into Bangla voice and text for dementia detection, as well as the development of sentence auto-completion.

Chapter 3

Research Methodology

The research strategy, data collection procedures and experimental approach implemented in the research have been outlined in this chapter. It highlights the approaches and methods that are used to achieve a precise identification of dementia.

3.1 Methodology

The audio recordings of individuals with dementia and non-dementia were collected and stored in .wav format. Then process the voices with noise reduction to improve quality. Important audio features such as MFCCs and RMS and pitch were extracted from the audio data. Various classification models, including LSTM, Random Forest, Gradient Boosting and Logistic Regression were employed to differentiate between dementia and non-dementia voices based on extracted features. Then transcribed the voices into text for classify with the text also. Likewise, the transcribed text undergoes feature extraction using Term Frequency-Inverse Document Frequency (TF-IDF) and trigrams to capture contextual relationships between words. These features are then used for classification to distinguish dementia and non-dementia related sentences. For sentence auto completion, a Bi-LSTM models generation helping dementia patients complete their speech. By combining these methods, this research aims to improve both dementia detection and speech completion capabilities for dementia patients.

3.1.1 Overview

This research focuses on detecting dementia from speech and text and providing auto sentence completion support for dementia patients. For speech-based classification, audio recordings in WAV format were processed using Audacity for trimming and noise reduction. Features like MFCCs and RMS were extracted and used to train models, including LSTM, Random Forest, Logistic Regression and Gradient Boosting. For text-based classification, audio transcriptions were employed by using the Whisper model and then match the transcription for correction. Text preprocessing involved tokenization and TF-IDF extraction of trigram features. Models like LT, RF, SVM, and LSTM were trained to classify sentences as dementia or non-dementia. Finally, the sentence auto-completion phase utilized a tokenizer to generate n-gram sequences and padded inputs. A BiLSTM model was applied to predict meaningful words to complete the sentence. Model evaluations ensured consistent results across all tasks.

3.1.2 Proposed Methodology

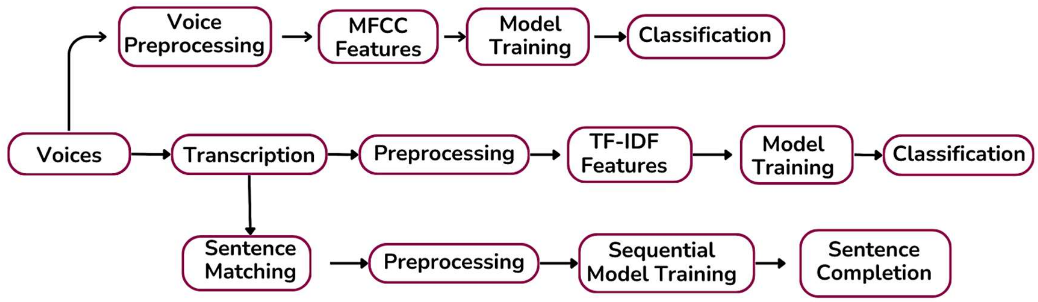

Fig 3.1.2.1: Proposed methodology.

The speech data for this study was prepared systematically. Initially, the audio recordings of the individuals with dementia and non-dementia were stored in .wav format. Subsequently, Audacity, a free software was used to trim recorded audio precisely. Noise reduction preprocesses were applied to the audio data to enhance its quality and suitability for analysis. Mel Frequency Cepstral Coefficients (MFCCs), and Root Mean Square (RMS) Energy, and Pitch (for jitter computation) were among the important audio features that were extracted from the audio [20]. Custom functions were used to manipulate and merge the feature vectors into a comprehensive feature vector. Various classification models, including LSTM, Random Forest, Gradient Boosting, and Logistic Regression, were used to distinguish between voices of dementia and non-dementia based on the extracted features. For the text-based classification, the dataset preparation involved transcription of dementia and non-dementia audio samples into text using the Whisper model, followed by manual correction of transcription errors to ensure accuracy. Then the text data went for preprocessing, including tokenization of Bangla text and removing resulting in a clean data frame. For feature extraction, TF-IDF (Term Frequency-Inverse Document Frequency) was employed to generate numerical representations of text from the Bangla tokenizer [21]. The classification phase utilized these features to train Logistic Regression, Random Forest, Support Vector Machine (SVM) in machine learning, and Long Short-Term Memory (LSTM) models in deep learning to classify sentences as dementia or non-dementia. In the final phase, the focus shifted to sentence completion, assisting dementia patients in generating meaningful continuations of incomplete thoughts. Sentences were transformed into n-gram sequences to prepare input-output pairs and then padded to ensure consistent lengths. The model architecture includes an embedding layer to transform words into dense vector representations, followed by Bidirectional LSTM (BiLSTM) layers to capture sequential patterns and contextual dependencies. Model evaluations involved confusion matrices, precision, recall, F1-scores, and cross-validation reports for classification tasks, while sentence completion models were focused on the contextual relevance of generated outputs.

3.2 Detailed Methodology and Design

3.2.1 Dataset

The primary motive of this research is to analyze and detect dementia from both speech and text while providing sentence autocompletion support. It consists of a combination of voice recordings and text data. The voice dataset comprises 740 recordings, divided equally between dementia (370) and non-dementia (370) categories. The voices were recorded based on the sentences that were specifically written for this study. 370 dementia voice samples were recorded from 10 participants. Focusing on this age group because dementia symptoms are more prominent here. Similarly, 370 non-dementia recordings were gathered who represented healthy individuals. All voices were recorded in a format to ensure high-quality audio suitable for advanced analysis.

3.2.2 Overall Execution Strategy

A third-party mobile application, Voice Pro, was used to collect audio data for this study. This application enabled the recording of audio in the wav format. Once the data capture process was finished, Audacity, a free and widely used audio editing software, was employed to remove any extraneous sections of the audio recordings.

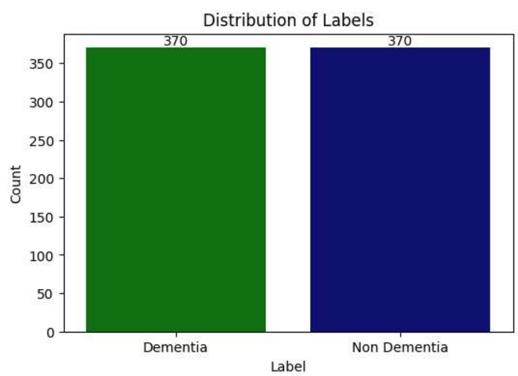

Fig 3.2.2.1: Distribution of classes.

Figure 3.2.2.1 demonstrates the proportional representation of both classes on the dataset. Insights into class balance are additionally offered, which is necessary for model performance evaluation. Data cleaning was then implemented through noise reduction and soundfile to balance signals and noise. The preprocessing procedures described above were highly beneficial in reducing extraneous noise, enhancing the quality of the audio data, and rendering the data suitable for subsequent analysis.

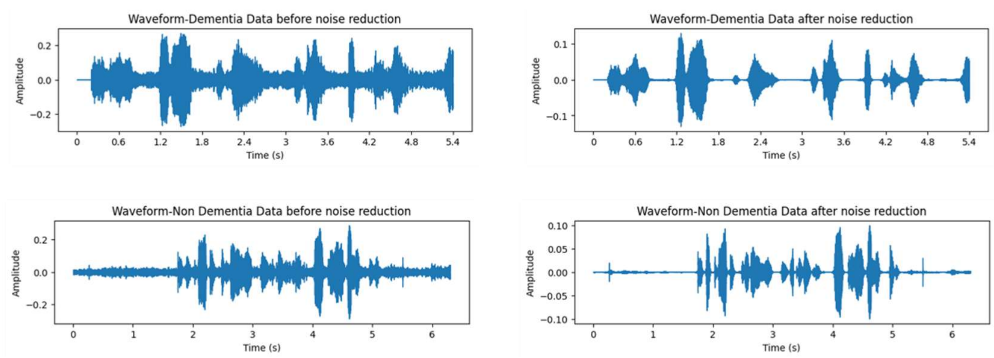

Fig 3.2.2.2: Audio frequency before and after preprocessing.

Fig 3.2.2.2 demonstrates that noise reduction extensively cleans the signals for dementia and non-dementia data, reducing unwanted noise and enhancing different speech patterns. This improves the clearness of signals, leading to precise analysis. Audio and feature extraction and data preparation Key features (MFCCs and Root Mean Square (RMS) energy, and Pitch for jitter calculation) were extracted. MFCC1, MFCC2, and MFCC5 critical coefficients, which are strongly impacted by vocal energy and articulation clarity, are consistently greater in non-dementia speech. The majority of MFCCs show lower mean values for dementia speech, suggesting a reduction in vocal control and energy distribution, even if it sometimes surpasses non-dementia speech in certain coefficients. To facilitate the application of machine learning models, label encoding was implemented after all features were extracted to convert categorical labels into a numerical format. Dementia data were encoded as 0 and non-dementia data as 1. Finally, a train-test division was established with a fixed train-test ratio and a random state to ensure the results were reproducible. We employed a variety of machine learning models, including Logistic Regression, Random Forest, and Gradient Boosting, to complete the classification task. Furthermore, to enhance the accuracy of data classification, a deep learning model based on Long-Short-Term Memory (LSTM) was implemented. LSTM was implemented as a single-layer model with 128 concealed units. A sigmoid activation function was implemented in the output layer to facilitate multi-class classification, while tanh was implemented in the LSTM layer for cell states and sigmoid was implemented for gate operations. The final training procedure of the LSTM model consisted of no more than 50 pushes, and it was trained using 10 epochs during cross validation. Early halting was implemented to prevent overfitting. In an effort to identify dementia in the audio signals, the models were trained and validated. By this, the methodology denotes the appropriate utilization of the encoded labels and extracted features, as it was implemented within a suitable data analysis and classification framework for the audio data. After completing the voice classification process, the next phase of this research focuses on text-based analysis. This stage involves two key components: classifying dementia and non-dementia individuals based on textual data and developing a sentence autocompletion system to assist in real-time communication. All audio recordings were derived from the predefined sentences generated earlier for this study, ensuring a structured and consistent dataset for transcription and analysis. The recorded voices were transcribed using the Whisper model from the hugging face ecosystem. The transcribed outputs were then compared against the predefined sentences to evaluate the transcription accuracy. This initial comparison yielded an average similarity score of 86.74%.

Table 3.2.2.1: Transcription and matching accuracy.

Predefines sentences	Transcription output	class
আমি কি খেয়েছি মনে করতে পারছি না।	তিনি কী খেয়াসি মনে করতে পাচ্ছি না।	Dementia
তোমার মা এখন মাঝেমধ্যে আমাকে খাইয়ে দেয়।	মার মা এখন মাঝে মধ্যে আমাকে খাইয়ে দায়।	Dementia
কাজ শেষে আরামের জন্য বই পড়া আমার পছন্দের।	কাজ শেষে আড়ামের জন্য বৈপড়া আমার পছন্দের।	Non-Dementia

The next step involved preprocessing the text dataset to prepare it for feature extraction and model training. The preprocessing was performed systematically to retain only the essential parts of the sentences, ensuring the extraction of meaningful features. Initially, Bangla punctuation marks, including full stops (।), commas (,), and question marks(?), were removed from the dataset. Following this, predefined Bangla stop words were identified and removed from all sentences. This step eliminated common but non-informative words, leaving behind the main components which can be shown in (Fig 3.2.2.6).

Fig 3.2.2.6: Word cloud for both classes.

After preprocessing the sentences, tokenization was performed. Following tokenization, TF-IDF features were extracted to quantify the importance of words and phrases within the dataset which is shown in (Table 3.2.2.2). For this research, trigrams were used, and the process resulted in a total of 2,000 features representing the linguistic characteristics of the sentences in a detailed and structured manner. For the text classification task, three machine learning models Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM) were trained using extracted TF-IDF feature with a tri-gram representation. Logistic Regression was implemented with regularization and class balancing to effectively manage class imbalance, Random Forest utilized an ensemble of 10 decision trees and Support Vector Machine employed a linear kernel to capture the discriminative patterns between dementia and non-dementia text classes effectively. The processed text data was split into 80% training and 20% testing sets for all the models. Then cross-validation was conducted to assess the consistency and reliability of the models. For developing an LSTM model, the text data was first preprocessed by tokenizing sentences into sequences of integers, followed by padding to ensure uniform sequence lengths. The model architecture included an embedding layer to convert integer-encoded words into dense vector representation with a dimension of 60. Batch normalization was applied to stabilize and accelerate training, while a fully connected layer and ReLU activation function for processing the data. The output layer consisted of a single neuron with sigmoid activation for binary classification. The next phase of this research focuses on sentence auto-completion, a larger dataset was required beyond the 370 transcribed sentences. So far, the classification task was done by the transcribed audio data. This tiny dataset is not enough for tasks like autocompletion. To address this, we utilized the generated sentences for this study. These sentences were preprocessed by removing punctuation marks to standardize the text. After preprocessing, a total vocabulary of 1,968 unique words was identified, with the maximum sentence length being 14. The sentences were then tokenized into integer sequences, and N-gram sequences were generated for model training. Padding was used to ensure uniform sequence length. For each N-gram sequence, the last token of each sentence was used as the target output (test), while the preceding tokens served as the input (train). This approach allowed the Bi-LSTM model to learn the contextual flow of sentences efficiently, easing accurate sentence prediction and completion. The Bi-LSTM model for sentence completion integrates an embedding layer to map words into dense vectors, followed by a Bidirectional LSTM (Bi-LSTM) layer with 150 units to learn contextual patterns in both forward and backward directions. An attention mechanism enhances the model's focus on relevant parts of the input, with a second Bi-LSTM layer fine-tuning the output. A dropout layer (0.5) prevents overfitting, and a dense output layer with SoftMax activation predicts the next word from a vocabulary of 1,969 words. The model, compiled with categorical cross-entropy and the Adam optimizer, was trained for 70 epochs. Its architecture effectively captures sequential and contextual relationships, making it suitable for assisting dementia patients by completing incomplete sentences.

3.3 Project Plan

The project plan outlines the oriented timetable, resources and objectives that are crucial for the effective execution of the study on the identification of dementia. It provides an outline to ensure that the study is carried out in an organized manner while following the established objectives and timelines.

Table 3.3.1: Detailed project plan.

Phase Name	Activities	Outcomes	Duration
Phase1: Sentence Generation	- Collect dementia and non-dementia sentences. - Design linguistically relevant sentences. - Ensure diversity in sentence structures and cognitive considerations.	Dataset of 1520 sentences divided into dementia (1150) and non-dementia (370) categories.	3 weeks
Phase 2: Speech Collection	Record dementia and non-dementia voice samples.	Audio dataset with dementia and non-dementia speech.	4 weeks
Phase 3: Speech and Sentence Preprocessing	- Remove noise and irrelevant data from speech files. - Transcribe recorded audio files. - Validate transcription accuracy. - Remove punctuation. - Tokenize sentences. - Match speech and text datasets.	Clean and preprocessed datasets (speech and text) ready for model input.	3 weeks
Phase 4: Model Development and Analysis	- Train ML/DL models (e.g., CNN, LSTM) for dementia detection. - Fine-tune hyperparameters. - Evaluate models using metrics (accuracy). - Perform comparative analysis.	- Developed and evaluated models for dementia detection. - Performance comparison report.	4 weeks
Phase 5: Visualization and Documentation	- Create visualizations (e.g., confusion matrix). - Analyze outcomes and highlight key findings. - Document the methodology, results, and conclusions.	Finalized documentation and visuals for the thesis/report.	2 weeks

Phase 6: Submission	- Compile the thesis/report. - Submit the project. - Present results if required.	Successfully submitted project and presentation.	2 weeks

3.4 Task Allocation

The thesis was done jointly by two individuals, with allocated responsibilities according to their respective areas of expertise. The data gathering process and literature review were engaged by both of us. One member focused on the objective of identifying dementia through speech analysis associated with that aspect. Another member effectively managed sentence completion and text classification. Although we collaborated, we were able to independently manage our own commitments. We collaborated to oversee the evaluation of the model, the analysis of outcomes, the visualization of the data and the resulting documentation.

3.5 Summary

After writing and collecting the sentence and speech data, the audio data preprocessed by trimming and reducing noise, followed by the extraction of key audio features, including Mel Frequency Cepstral Coefficients (MFCCs), RMS Energy, and Pitch, using the LIBROSA library. For the text data, audio recordings were transcribed using the Whisper model, with subsequent manual corrections for accuracy. The text was then tokenized (for Bangla text) and cleaned by removing unnecessary symbols. To represent the cleaned text numerically, TF-IDF (Term Frequency-Inverse Document Frequency) was applied, with a focus on trigram features. For classification, various machine learning models were employed. Logistic Regression, Random Forest, Gradient Boosting, and LSTM models were used to classify the speech data based on the extracted audio features. For the transcribed text, Logistic Regression, Random Forest, Support Vector Machine (SVM), and LSTM models were applied to differentiate between dementia and non-dementia sentences. Additionally, a Bidirectional LSTM (BiLSTM) model was designed for sentence autocompletion, helping patients with dementia generate meaningful continuations of incomplete sentences. This model architecture included an embedding layer to represent words as dense vectors, followed by BiLSTM layers to capture the sequential and contextual dependencies of the sentences. The models were evaluated using standard metrics such as confusion matrices, precision, recall, F1-scores, and cross-validation reports for the classification tasks. For sentence autocompletion, the focus was on ensuring the contextual relevance and coherence of the generated outputs, providing meaningful sentence continuations for patients.

Chapter 4

Implementation and Results

This chapter describes the execution of the recommended models, the results of the studies and the evaluation of the model's efficacy. This involves the analysis and insights that emerged from the results.

4.1 Environment Setup

Dementia and non-dementia data are analyzed using machine learning and deep learning techniques for the purpose of identification of dementia in this study. Python was the primary language for programming, with scikit-learn being used for constructing machine learning models and TensorFlow and Keras being implemented to develop deep learning models. Visual Studio offered interactive programming environments, and the models were trained and tested in VS code with the aid of dedicated GPU support for faster model training. Key libraries included scikit-learn for ML models (Logistic Regression, Random Forest, SVM, Gradient Booosting), Pandas and Numpy for data processing and TensorFlow/Keras for the building of deep learning model, including LSTM and Bi-LSTM. The dataset consists of 740 speeches, divided into dementia and non-dementia categories and was split into training and testing sets.

4.2 Comparative Analysis

To evaluate the effectiveness of the proposed methodology in this research, a comparative analysis is conducted in relation to existing studies in this field. The comparison shows diverse datasets, techniques, and evaluation metrics across both voice and text-based approaches as well as sentence autocompletion. The table below highlights the distinct contributions of this research.

Table 4.2.2.1: Compare with other studies.

Das, A., & Bandyopadhyay, S.	9,337 text entries (6,698 Bangla, 2,639 Romanized)	Hesitation counts, TTR, Word2Vec, GloVe. The model used was LSTM networks.	LSTM networks 78% (2-class), 55% (3-class classification)
Hasan, M. M., Haque, M. R., Al Amin	108,950 comments	Multinomial Naive Bayes (MNB), SVM, KNN, LSTM, BiLSTM, Bangla-BERT.	82.64% (MNB), 83.23% (Bangla-BERT-Base), 79.14% (BiLSTM)

This study	740 sentences (370 Dementia and 370 Non-dementia)	LR, RF, SVM, LSTM with Tf-IDF tri-gram features.	LR (82.43%), RF (74.32%), SVM (81.08%) and LSTM (83.78%)

4.3 Results and Discussion

The following figure shows the test accuracy of the four models: Logistic Regression, Random Forest, Gradient Boosting and Long Short-Term Memory (LSTM). The accuracy of Logistic Regression is 85.86%, which is the lowest one. While Random Forest and Gradient Boosting achieved similar accuracies of 90.62%. The LSTM model surpasses the other models and shows its superiority in this context with an accuracy of 92.84%.

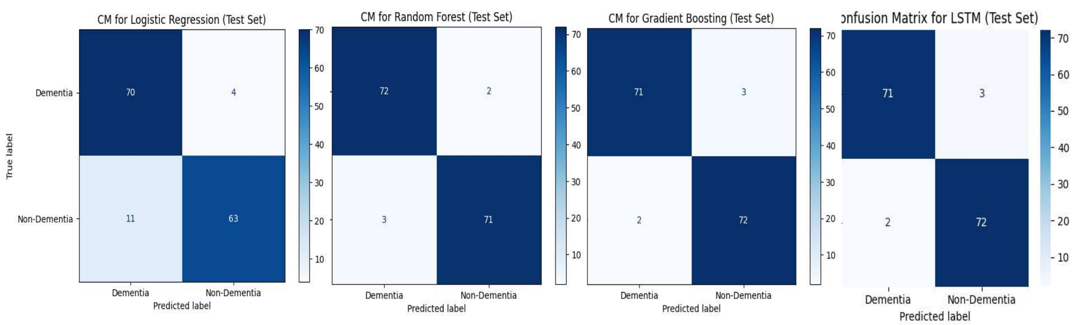

Fig 4.3.2: Confusion matrices of speech data.

The confusion matrices plainly illustrate the performance of four models in the classification of Dementia and Non-Dementia cases. The rate of misclassification is higher in Logistic Regression, with 11 false positives and 4 false negatives. Gradient Boosting and Random Forest exhibit superior outcomes with negligible misclassification (3 and 2 false positives, respectively). The LSTM model demonstrates an exceptional success rate, with only three false negatives and two false positives. Random Forest, Gradient Boosting, and LSTM exhibit exceptional performance; however, LSTM is the most effective. In the text classification task, the evaluated four models were Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), and Long Short-Term Memory (LSTM). The performance of each model is shown on accuracy graph (Fig 4.3.3). The Random Forest model achieved an accuracy of 70.32%, while SVM and LR performed better with the accuracy of 72.08% and 77.43%. Despite the higher accuracy SVM and LR models made some incorrect predictions. However, these misclassifications were within an acceptable range. On the other hand, RF was more biased to predict dementia, that's why accuracy was lower than SVM and LR.

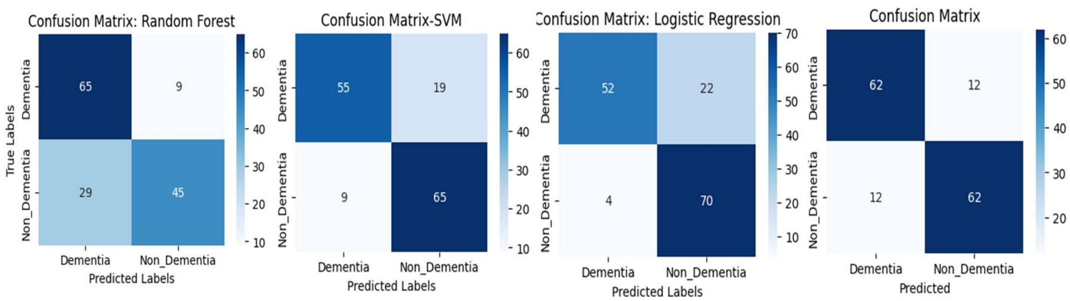

Fig 4.3.4: Confusion matrices of text classification.

The LSTM model outperformed all other models with an accuracy of 77.78%. The confusion matrix for the LSTM model shows that its predictions were well-balanced for both dementia and non-dementia classes shown in (Fig 4.3.4). This balanced performance makes LSTM an ideal model for this task. Overall, after evaluating the results, it can be said that SVM, LR, and LSTM models all performed well, but the LSTM model achieved the highest accuracy.

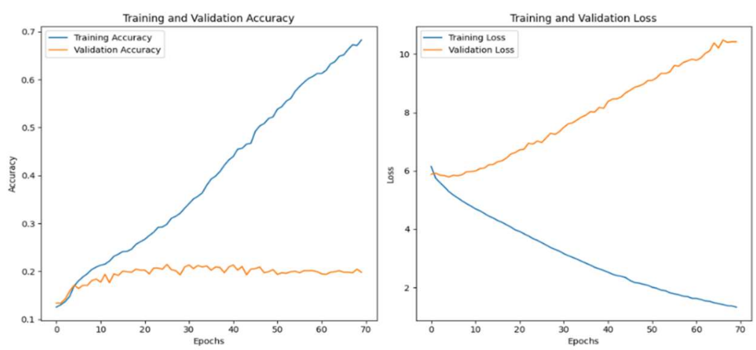

Fig 4.3.5: Bi-LSTM accuracy and loss graph

In the sentence auto-completion task, a Bidirectional LSTM (BiLSTM) model was employed to assist dementia patients with Bangla text data. The model had an accuracy of 20.8%. The model demonstrated a training accuracy of 70.4% and a validation loss of 20% (Fig 4.3.5). While this result is considered a little bit low. It is noted that, given the limited amount of data available in the Bangla language and then process that a significant challenge. Despite these constraints, the 23.8% accuracy is respectable, especially when compared to other studies. This performance indicated in the (Table) shows that the BiLSTM model can generate useful sentence completions for Bangla-speaking dementia patients, marking a valuable contribution to dementia care in our society.

Table 4.2.2.2: Sentence completion evaluation.

Input	Output
এবার করবানির ছাগলের ...	এবার করবানির ছাগলের দাম অনেক বেশি দাও
সকালে আমি নাস্তা করেছিলাম ...	সকালে আমি নাস্তা করেছিলাম কিনা মনে পড়ছে না
আমার রুটি দিয়ে হাঁসের ...	আমার রুটি দিয়ে হাঁসের মাংস খেতে ইচ্ছা করছে

4.4 Summary

The accuracy results for the classification models showed that the LSTM model outperformed all others, achieving the highest accuracy of 92.84% in speech classification, surpassing Random Forest and Gradient Boosting, which both reached 90.62%. In text classification, the LSTM model again demonstrated superior performance with an accuracy of 77.78%, outperforming Support Vector Machine (72.08%) and Logistic Regression (77.43%), while Random Forest lagged at 70.32%. For the sentence autocompletion task, the BiLSTM model achieved an accuracy of 20.8%, which although not so good. But it gives a promising outcome considering the challenges associated with the limited availability of Bangla language data. Overall, LSTM models consistently showed the highest accuracy across tasks, highlighting their effectiveness in both speech and text classification.

Chapter 5

Engineering Standards and Design Challenges

This chapter dives into the engineering standards which were followed and the challenges that were faced throughout the procedure. The techniques applied to overcome these challenges are the subject of this chapter.

5.1 Compliance with the Standards

This research was carried out in accordance with different engineering and design standards to ensure a methodology that was both structured and dependable. It explores the industry's standard-set criteria for the efficacy of machine learning frameworks and applications, dataset preprocessing stages, and matrix evaluation. This guarantees the system's scalability, correctness, and trustworthiness. Additionally, it’s adherence to the ethical and professional standards of artificial intelligence and machine learning.

5.1.1 Communication Standards

We used transparent and effective communication protocols throughout our research to ensure the successful completion of the project and smooth teamwork. The responsibilities were split between the two of us according to our respective areas of expertise. One member was responsible for audio analysis, while the other focused on text classification and sentence completion. Progress was assessed, and further actions were determined during the weekly discussions. These discussions were conducted both in person and remotely through Google Meet, which enabled us to improve the quality of our research and maintain our alignment. Face-to-face discussions were particularly beneficial for the development of new ideas and the solution of complex issues, as they facilitated the improvement of methods and stronger collaboration. The research process was carefully documented, that includes data collection, preprocessing, model training, and evaluation of the models. This ensured the availability of comprehensive recordings for future analysis and replication. The real-time sharing of research materials, code, and files was facilitated by collaboration tools, such as Google Collab, Google Drive, Google Sheets, and Google Docs, while version control was simultaneously maintained. In order to ensure that the findings were understood and agreed upon by all parties, the results of each phase were transparently communicated, evaluated, and validated in a collaborative manner. The resolution of any disagreements or obstacles through open discussions encouraged a collaborative and supportive environment. By integrating both virtual and face-to-face interactions, our communication standards ensured the research's overall success, continuous development, and accountability.

5.2 Impact on Society, Environment and Sustainability

5.2.1 Impact on Life

The lives of those affected are heavily impacted by dementia, resulting in a gradual decline of cognitive and functional capacities. The daily responsibilities of individuals who are suffering from dementia, like recognizing familiar surroundings, managing personal matters, or remembering names, become difficult. Communication challenges further restrict their capacity to articulate their thoughts, emotions, and needs in the later stages of the disease, resulting in social withdrawal, isolation, and frustration. The emotional burden felt by dementia patients gets worse by feelings of insignificance and the loss of independence. These cognitive impairments have an impact on not only the patients, but also their families, who are required to adapt to the emotional and financial responsibilities of caregiving. The quality of life for both patients and caregivers might decrease as dementia advances, as patients frequently depend on them for additional support. These challenges are intended to be mitigated through the implementation of machine learning and natural language processing (NLP) models in this research. This will be achieved by enhancing communication capabilities, helping with sentence completion, and facilitating early diagnosis. By addressing these aspects, the study has the potential to substantially enhance the daily lives of individuals with dementia, providing them with a means to communicate more effectively and maintain a level of independence for an extended period as the disease progresses. Caregivers may experience diminished frustration, enhanced emotional well-being, and ultimately, superior care because of enhanced communication techniques.

5.2.2 Impact on Society & Environment

Society and the environment are substantially affected by our research. By emphasizing early detection of dementia, it provides a non-invasive, accessible approach that improves the quality of life for patients and alleviates the burden on families and guardians. Fostering awareness and comprehension of dementia, it addresses societal stigma. By enabling the development of personalized treatment plans and the surveillance of cognitive decline, the developed instruments enhance healthcare systems. The study implements energy-efficient technologies and digital data to guarantee optimal resource utilization with minimal environmental impact. This fosters a more informed and healthier society while simultaneously upholding environmental responsibility.

5.2.3 Ethical Aspects

The ethical considerations of this research focus on several key areas:

i. Ensuring dementia patients or their legal guardians fully understand and consent to the research, especially since patients with advanced dementia may not be capable of consenting themselves.
ii. Protecting sensitive health data by anonymizing it and maintaining strict confidentiality to safeguard participants' personal information.
iii. Ensuring that the machine learning models used for dementia prediction are unbiased, and fair.
iv. Considering the role of caregivers and ensuring the technology supports rather than replaces them, alleviating their burden without compromising there.

5.2.4 Sustainability Plan

To guarantee the long-term effectiveness and impact of this research, it is essential to establish a complete sustainability plan. The strategy highlights the importance of both technical and social factors to ensure that the technology that has been developed is accessible, adaptable, and advantageous to dementia patients in the long term. The sustainability plan includes several critical elements, including:

Continuous Improvement: The system is refined by regularly updating the models with updated data and collecting feedback from users, caregivers, and healthcare professionals.

Training and Education: Developing training programs for healthcare providers, caregivers, and families to ensure effective use of the application and enhance dementia care.

Collaboration: Establishing partnerships with local healthcare organizations, governments, and NGOs to promote technology adoption and integrate it into public health systems.

Cost-Effective Solutions: Ensuring the technology remains affordable by seeking funding or grants to support deployment and maintenance in low- and middle-income countries.

5.3 Project Management and Financial Analysis

Provide a cost analysis in terms of budget required and revenue model. In the case of budget,you must show an alternate budget and rationales.

Table 5.3.1: Cost analysis.

Name	Estimated Cost
GPU (GeForce RTX 3050)	30,500 BDT
Processor (intel i5 12th gen)	16,000 BDT
RAM (32 GB)	15,800 BDT
Storage (1 TB)	16,000 BDT
Total Hardware cost	**78,300 BDT**
Data collection cost	12,000 BDT
Development and research cost	40,000 BDT
Web-based application	55,000 BDT
Total cost	**188,000 BDT**

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

In this section, provide a mapping with problem solving categories. For each mapping add subsections to put rationale (Use Table 5.1). For P1, you need to put another mapping with Knowledge profile and rational thereof.

Table 5.4.1.1: Mapping with complex problem solving.

EP1 Dept of Knowledge	EP2 Range Of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent Of Stake-holder Involvement	EP7 Interdependence
✓		✓				✓

Mapping with Knowledge Profile for EP1

Table 5.4.1.2: Mapping with knowledge Profile (EP1).

K1 Natura l Science s	K2 Mathematic s	K3 Engineeri Ng Fundame n tals	K4 Special i st Knowl e dge	K5 Enginee r ing Design	K6 Enginee r ing Practice	K7 Comprehe n sion	K8 Resear c h Literat ure
		✓	✓	✓	✓		✓

K3(Engineering Fundamentals): Comprehending the fundamental principles of audio signal processing, Bangla text analysis, natural language processing, machine learning and deep learning.

K4(Specialist Knowledge): Exceptional proficiency in feature extraction techniques (MFCCs, Chroma Features, TF-IDF, N-Grams) and deep learning models such as LSTM, Bi-LSTM.

K5(Engineering Design): The development of models and systems that are specifically designed for the detection of dementia from Bangla speech and text. Creating workflows for the preprocessing, feature extraction, model training, and evaluation of data.

K6(Engineering Practice): The practical application of ML and DL pipelines using tools such as scikit-learn, keras, and TensorFlow. The deployment of solutions that are both efficient and scalable on platforms such as Visual Studio.

K8(Research Literature): Examination and analysis of previous research on the detection of dementia, Bangla text classification and language modeling.

Mapping with Knowledge Profile for EP3

Table 5.4.1.3: Mapping with knowledge Profile (EP3).

K1 Natura l Science s	K2 Mathematic s	K3 Engineeri Ng Fundame n tals	K4 Special i st Knowl e dge	K5 Enginee r ing Design	K6 Enginee r ing Practice	K7 Comprehe n sion	K8 Resear c h Literat ure
	✓	✓	✓	✓			✓

K2 (Mathematics): Applying mathematical techniques for feature extraction and performance evaluation. Models like Bi-LSTM and LSTM are implemented with the help of activation functions such as ReLU and SoftMax. Utilize mathematical and statistical principles to implement models like LR, SVM and Gradient Boosting.

K3(Engineering Fundamentals): Utilizing the principles of machine learning, deep learning and natural language processing (NLP) to extract features and analyze data.

K4(Specialist Knowledge): Employing dementia-specific data and advanced deep learning and machine learning models. Analyze and understand dementia patient behavior and collect the voices from dementia and non-dementia individuals.

K5(Engineering Design): The formation of workflows for optimization, hyperparameter customization, and model comparison.

K6(Engineering Practice): Development of data analysis pipelines with the help of scikit-learn, keras, and TensorFlow. Visualize the findings and save necessary files for future use.

K8(Research Literature): The steps for reviewing previous studies and implementing results to improve models and evaluation metrics. It helps to analyze the gaps, also compare and evaluate our outcomes.

Mapping with Knowledge Profile for EP7

Table 5.4.1.4: Mapping with knowledge Profile (EP7).

K1 Natural Sciences	K2 Mathematics	K3 EngineeriNg Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K7 Comprehension	K8 Research Literature
✓	✓	✓	✓	✓	✓	✓	✓

K1 (Natural Sciences): The breakdown of speech signals through the perception of acoustics and phonetics.

K2 (Mathematics): Applying statistical techniques to extract features such as MFCCs, AUC-ROC, TF-IDF and Implement models like Logistic Regression, SVM and LSTM. Activation functions such as RELU and Softmax are also used for evaluating metrics.

K3 (Engineering Fundamentals): Implementation of natural language processing (NLP) and machine learning techniques for interconnected system components.

K4 (Specialist Knowledge): Advanced tasks related to classification, sentence auto completion, feature extraction and model implementations.

K5 (Engineering Design): The procedures that combine preliminary processing, training and deployment with ensuring scalability.

K6 (Engineering Practice): Processing data effectively by pipeline management and integrating techniques like Numpy, TensorFlow, Keras, and scikit-learn.

K7 (Comprehension): Evaluating the ethical and sociological consequences to ensure that the healthcare system meets all the requirements for the greater population.

K8 (Research Literature): Review more than twenty different research to manage complex interdependencies in an effective manner.

5.4.2 Engineering Activities

Table 5.4.2.1 Mapping with complex engineering activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
✓	✓	✓	✓	✓

EA1 (Range of resources): Applies a wide range of resources to achieve the goals, consisting of human expertise, machine learning and deep learning frameworks.

EA2 (Level of interaction): Resolving obstacles in the integration of speech and text processing with modeling while addressing conflicting requirements. This study combines audio signal analysis with textual data while balancing high clinical accuracy and efficiency for real-time performance.

EA3 (Innovation): The research applies deep learning creatively to classify dementia with Bangla language to handle speech and text data. This approach fills the gap in research for dementia in Bangla language and builds a sentence auto completion model specifically for dementia patients.

EA4 (Consequences for society and the environment): This study contributes to society by enabling early dementia detection, improving patient care, supporting dementia patients and families.

EA5 (Familiarity): By implementing engineering principles to an innovative approach to solve new challenges in dementia detection from Bangla speech and text.

5.5 Summary

The objective of this research is to improve the quality of dementia people and reduce the financial and emotional stress on the caregivers and family members by early identification of dementia and sentence auto completion. It contributes to the formation of a better society by increasing awareness and reducing the discrimination related to dementia. The research focuses on the enhancement of healthcare accessibility in an environmentally sustainable way with the use of digital data and computational resources. The technology is meant to be user friendly, adaptable as well as inclusive, with an emphasis on moral principles to meet the requirements of various communities. This commitment to sustainable and efficient problem solving is demonstrated by the strategic allocation of resources for data collection, model training and further analysis. Using innovative approaches to analyze speech and text data, this study addresses the real-world challenges by combining expertise from natural language processing, deep learning and medical sector. This cross-disciplinary teamwork is nurturing the advancement of the solutions in the field of healthcare.

Chapter 6

Conclusion

As the research moves forward, continuous collaboration with the healthcare professionals, regular model enhancement and real-world testing will play a vital role for ensuring its long-term success. By making innovation in dementia care, we aim to provide more accessible and effective solutions for individuals impacted by this condition.

6.1 Summary

This research focuses on applying machine learning and natural language processing (NLP) techniques to tackle the challenges faced by dementia patients. This study concern particularly for the peoples who speaks Bengali. This investigation aims to early diagnosis, enhance communication and assist with sentence auto completion for dementia patients through sequential learning. It utilizes both speech and text data. Extracting features such as Mel Frequency Cepstral Coefficients (MFCC), Chroma features, TF-IDF and N-Gram for classification. Transcription and text processing were also carried out for sentence analysis. Various machine learning models including Long Short-Term Memory (LSTM), were used to classify dementia and non-dementia cases. Results showed that LSTM performed best with the highest accuracy in both speech and text classification tasks. Additionally, the study also investigates the potential of these technologies in supporting sentence completion for dementia patients particularly in Bangla. The bidirectional LSTM (Bi-LSTM) was implemented to sequentially generate and try to complete the sentences. The overall goal is to enhance communication and quality of life for individuals suffering from dementia. Also contributing to dementia care solutions in Bangladesh.

6.2 Limitation

The following limitations were identified and considered for future improvements:

i. Data Limitations: The study used a relatively small dataset for both speech and text, which may not fully represent the diverse range of dementia cases. The limited availability of Bangla sentence data may have impacted on the model's ability to generalize across different demographics and dialects.

ii. Technology Access: Machine learning models, and web-based applications required internet access and modern devices, which could be a barrier to adoption in remote or rural areas.

iii. Ethical Concerns: This research involves vulnerable populations, and they raise important ethical issues including informed consent, data privacy and fairness. These issues must be addressed continuously.

6.3 Future Work

Future work in this area should attend several key holes that would allow this research to have a greater impact and greater effectiveness. First resolving the data sparsity issues which is essential to make the model robust and generalizable. The main future work will be collecting a larger and more diverse dataset (e.g. different dialects and specific types of dementia sentences) to ensure prediction accuracy and make sure the models can generalize across different populations. Furthermore, refining existing models through techniques such as transfer learning can improve their performance. The system could achieve more accuracy if using more advanced neural network architectures, such as Transformer-based models like Bangla-BERT. This could capture deeper contextual patterns in both speech and text. The important factor is real-world testing and validation are also important for assessing the utility of the models. Testing them in clinical settings or home care environments will provide valuable insights into the effectiveness of the models. It helps to refine the models for diverse real-world scenarios. Another important area of work is collaborating with healthcare providers. By working closely with healthcare professionals and caregivers, this technology can be better integrated into the clinical workflows. In conclusion, by considering these future directions, we seek to provide a better solution that improves the quality of life for dementia patients and for their caregivers which can be used in scalable and effective manner.

References

[1] Alzheimer's Disease International, “World Alzheimer Report 2010: The Global Economic Impact of Dementia”, p. 6.

[2] TBS News (2024), “1 in 12 Bangladeshi elderly have dementia: Survey”, TBS News, https://www.tbsnews.net/bangladesh/health/1-12-bangladeshi-elderly-have-dementia-survey-268675 [Accessed: Dec. 13, 2024].

[3] United Nations Bangladesh, "Fulfilling social protection for older persons in Bangladesh," United Nations Bangladesh, Aug. 29, 2023. [Online]. Available: https://bangladesh.un.org/en/248154-fulfilling-social-protection-older-persons-bangladesh [Accessed: Dec. 13, 2024].

[4] The Business Standard, "Number of elderly people grows 3.41 times faster than population growth," The Business Standard, Oct. 2, 2022. [Online]. Available: https://www.tbsnews.net/bangladesh/number-elderly-people-grows-341-times-faster-population-growth-506314 [Accessed: Dec. 13, 2024].

[5] E. Nichols, C.E. Szoeke, S.E. Vollset, N. Abbasi, F. Abd-Allah, J. Abdela, M.T.E. Aichour, et al., "Global, regional, and national burden of Alzheimer's disease and other dementias, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016," Lancet Neurol., vol. 18, no. 1, pp. 88–106, 2019

[6] World Health Organization, “Global action plan on the public health response to dementia 2017-2025”, Geneva: World Health Organization, 2017, pp. 2.

[7] E. Nichols, J. D. Steinmetz, S. E. Vollset, K. Fukutaki, J. Chalek, F. Abd-Allah, A. Abdoli, et al., "Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019," The Lancet Public Health, vol. 7, no. 2, pp. e105–e125, 2022.

[8] World Health Organization, "Dementia," World Health Organization, Mar. 15, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia [Accessed: Dec. 13, 2024].

[9] E. Vasse, et al., "Improving person-centered communication in dementia care," International Psychogeriatrics, vol. 22, no. 8, pp. 1235–1242, 2010.

[10] Centers for Disease Control and Prevention (CDC), "Leading Causes of Death," CDC Wonder, 2021. [Online]. Available: https://wonder.cdc.gov/controller/datarequest/D158;jsessionid=807535333530DEE5A4771D84DB64?stage=results&action=toggle&p=O_show_rank&v=true. [Accessed: Dec. 9, 2024].

[11] World Life Expectancy, "Bangladesh: Alzheimer’s & Dementia," 2020. [Online]. Available: https://www.worldlifeexpectancy.com/bangladesh-alzheimers-dementia. [Accessed: Dec. 9, 2024].

[12] Campbell, E. L., Docío-Fernández, L., Jiménez Raboso, J., & García-Mateo, C. (2020). Alzheimer’s Dementia Detection from Audio and Text Modalities. INTERSPEECH 2020. https://arxiv.org/abs/2004.06833

[13] Rifat, R., Khan, M. S. A., & Hasan, M. K. (2023). A Comparative Study on Bengali Speech Sentiment Analysis Based on Audio Data. ResearchGate. https://www.researchgate.net/publication/369395001

[14] Zhu, Y., Lin, N., Balivada, K. S., Haehn, D., & Liang, X. (2024). Adversarial Text Generation using Large Language Models for Dementia Detection. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 21918–21933.

[15] Ilias, L., & Askounis, D. (2022). Explainable Identification of Dementia from Transcripts using Transformer Networks. IEEE Journal of Biomedical and Health Informatics. https://arxiv.org/abs/2109.06980

[16] A. Rianti, S. Widodo, A. D. Ayuningtyas, and F. B. Hermawan, "NEXT WORD PREDICTION USING LSTM" J. Inf. Technol. Its Utiliz., vol. 5, no. 1, pp. 10, June 2022. [Online]. Available: https://doi.org/10.2654-802X. [Accessed: Dec. 12, 2024].

[17] Islam, M. R., Al Amin, & Zereen, A. N. (2024). Enhancing Bangla Language Next Word Prediction and Sentence Completion through Extended RNN with Bi-LSTM Model on N-gram Language. ArXiv preprint. Available: https://arxiv.org/html/2405.01873v1.

[18] 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(146). https://doi.org/10.1186/s13195-021-00888-3

[19] Kumar, M. R., Vekkot, S., Lalitha, S., Gupta, D., Govindraj, V. J., Shaukat, K., Alotaibi, Y. A., & Zakariah, M. (2022). Dementia Detection from Speech Using Machine Learning and Deep Learning Architectures. Sensors, 22(9311). https://doi.org/10.3390/s22239311

[20] Das, A., & Bandyopadhyay, S. (2016). Sentiment Analysis of Bangla and Romanized Bangla Text using LSTM Networks. arXiv preprint arXiv:1610.00369v2. https://arxiv.org/abs/1610.00369v2

[21] Hasan, M. M., Haque, M. R., Al Amin, M. M., & Ali, M. M. (2023). 'MONOVAB': An Annotated Corpus for Bangla Multi-label Emotion Detection. arXiv preprint arXiv:2309.15670v1. https://arxiv.org/abs/2309.15670v1

[22] M. R. Islam, A. Amin, and A. N. Zereen, "Word Completion and Sequence Prediction in Bangla Language Using Trie and a Hybrid Approach of Sequential LSTM and N-gram," arXiv, vol. 2405.01873v1, 2024. [Online]. Available: https://arxiv.org/html/2405.01873v1.

[23] Haque, M. M., Habib, M. T., & Rahman, M. M. (2015). Automated Word Prediction in Bangla Language Using Stochastic Language Models. International Journal in Foundations of Computer Science & Technology, 5(6), 67-75.

Dementia AI & Plag Check Report

ORIGINALITY REPORT

12%	10%	5%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	1%
2	Submitted to United International University Student Paper	1%
3	doctorpenguin.com Internet Source	1%
4	www.frontiersin.org Internet Source	1%
5	Submitted to Queen Mary and Westfield College Student Paper	<1%
6	H.L. Gururaj, Francesco Flammini, S. Srividhya, M.L. Chayadevi, Sheba Selvam. "Computer Science Engineering", CRC Press, 2024 Publication	<1%
7	link.springer.com Internet Source	<1%