A Comprehensive NLP-Based Voice Assistant System for Streamlined Information Retrieval in Metro Rail Services of Bangladesh

by

MD. Abir Rahman
19201003
Adnan Alamgir
19201126
Shaheedul Haque Chowdhury
19201058
Maliha Mushtari
19201090
Wasim Anzum
19201051

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
School of Data and Sciences
Brac University
January 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Abir Rahman

MD. Abir Rahman
19201003

Adnan Alamgir

Adnan Alamgir
19201126

Shaheedul Haque Chowdhury
19201058

Maliha Mushtari

Maliha Mushtari
19201090

Wasim

Wasim Anzum
19201051

Approval

The thesis/project titled "A Comprehensive NLP-Based Voice Assistant System for Streamlined Information Retrieval in Metro Rail Services of Bangladesh" submitted by

1. MD. Abir Rahman (19201003)
2. Adnan Alamgir (19201126)
3. Shaheedul Haque Chowdhury (19201058)
4. Maliha Mushtari (19201090)
5. Wasim Anzum (19201051)

Of Fall, 2023 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on 18 January 2024.

Examining Committee:

Supervisor:

Md. Khalilur Rhaman

Dr. Md. Khalilur Rahman
Professor
Department of Computer Science and Engineering
Brac University

Program Coordinator
(Member)

Dr. Md. Golam Rabiul Alam
Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

Our research paper, which we wrote from scratch, has been carefully constructed to assure its integrity and lack of plagiarism. To offer fresh perspectives and add to the body of knowledge, we have carried out in-depth research, gathered pertinent data, and critically analyzed the findings. Our research culminated in this publication, which demonstrates our dedication to rigorous scholarly work and academic honesty.

Abstract

Bangladesh's capital city Dhaka is served by the Dhaka Metro Rail. A metro-rail-based rapid transit system is considered one of the important technologies that may decrease the working-hour wasting issue in a developed nation owing to traffic congestion. It moves between communities within an urban region or the cities that constitute a metropolitan area. However, to reap the advantages for both passengers and metro-rail authorities, a voice assistance system is also necessary for a metro-rail-based transit system. Many passengers expressed their dissatisfaction and frustration at the appearance of such difficulties from the very beginning of the metro-rail service. Many people complained about experiencing trouble obtaining tickets from vending machines due to technological challenges when the mass transit system was opened to the public. Officials reported that vending machines stopped operating as people attempted to use them without understanding how to use them. This research proposes a noble approach for the general population of Bangladesh. General people will be able to interact with a voice assistant and get their job done, such as collecting information about the train and metro-rail station. We will be undertaking our research with the help of Natural Language Processing (NLP) based on the Artificial Intelligence Markup Language (AIML) structure for training the model. The primary dataset creation procedure is cautiously defined, comprising question generation, response formulation, and category assignment. To ensure the relevance and accuracy of our dataset, a thorough verification procedure was done in collaboration with the Managing Director of Dhaka Mass Transit Company Limited (DMTCL). Term Frequency-Inverse Document Frequency (TF-IDF), and a sequential neural network model are trained with the dataset. We designed a web application with the capability to receive voice input and provide spoken output. This application was developed by utilizing a voice recognition Application Programming Interface (API) for voice-to-text and text-to-voice conversion. A closed domain question answering (cdQA) NLP solution was utilized to acquire information about the given query. The paper intends to show how voice assistants can be used in daily life in metro rail stations with minimal effort and to analyze if there is potential for making them accessible to the general public.

Keywords: Natural Language Processing; Artificial Intelligent Markup Language; Dataset; Managing Director; Dhaka Mass Rapid Transit Company Limited; Term Frequency-Inverse Document Frequency; Long Short-Term Memory; Web Application; Application Programming Interface; Closed Domain Question Answering

Dedication

We dedicate this research paper to all the passengers and stakeholders of metro rail services. Our aim was to enhance the metro rail experience, making it more efficient, user-friendly, and accessible to all. This dedication is a testament to our unwavering commitment to improving urban transportation and ensuring a seamless journey for commuters.

Acknowledgement

Glory be to the Great Allah first and foremost, with whose assistance we were able to complete our thesis without too many difficulties. Secondly, we thank our renowned Supervisor Dr. Md. Khalilur Rahman sir for his advice and guidance. When we needed it, we could also count on the help of our parents, friends, and instructors. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

ListofFigures

ListofTables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

$AIML$ Artificial Intelligent Markup Language

ANN Artificial Neural Network

API Application Programming Interface

$cdQA$ Closed Domain Question Answering

$DMTCL$ Dhaka Mass Transit Company Limited

$LSTM$ Long Short-Term Memory

NLP Natural Language Processing

NLU Natural Language Understanding

RNN Recurrent Neural Network

$TF-IDF$ Term Frequency-Inverse Document Frequency

Chapter 1

Introduction

1.1 Why AI in Metro-rail

Bangladesh has successfully launched its metro rail service, making it the third South Asian nation to carry out such a revolutionary project. The impact of metro rail has already been seen in neighboring countries like India, which opened its first metro train service in 1984 in Kolkata, and Pakistan, which finished building its first metro system in Lahore in 2020 [37]. This is a huge step towards modernizing intercity travel. Bangladesh has made significant progress towards constructing a critical component of its transportation infrastructure, even though it entered the world of metro rail projects later than some of its regional peers. This feat is especially important for Dhaka, the biggest city in the nation, where over 20 million people are growing and struggle with traffic and clogged roadways. The metro system in Dhaka provides residents with a quicker and more effective way to go about, relieving the frustrations brought on by traffic jams. However, smooth communication of timely and accurate information to the general public is just as important to the smooth functioning of a metro rail system as the physical infrastructure.

The problem of guaranteeing timely and accurate information delivery increases with the number of trains and passengers on the metro rail network. The dense traffic and complex metro line system make it difficult for commuters to get important information like train timetables, ticketing alternatives, and other routes. Seeing the growing need for a novel approach, this study investigates artificial intelligence's (AI) unrealized potential for improving metro train operations.

Information dissemination in metro rail systems is a difficulty that can only be successfully addressed by using AI technology. Metro rail authorities can provide passengers with accurate and timely access to vital information by using artificial intelligence. Key areas where AI can transform metro rail operations and provide a streamlined and efficient commuter experience include dynamic communication, intelligent ticketing systems, and real-time train scheduling.

Specifically, this research supports the incorporation of an artificial intelligence voice assistant (AI) as a viable means of enabling smooth interactions between passengers and the metro rail system. It is planned to replace traditional vending machines with voice assistant AI, which would provide commuters with an easy-to-use interface for finding information, buying tickets, and using the metro rail system.

The research paper highlights the practical uses of voice assistance technology for increasing productivity, lowering operational bottlenecks, and improving user experience. The results offered here provide significant context to the discussion on how this technology may be strategically used in metro rail systems to modernize urban transportation.

Figure 1.1: Dhaka Metro Rail

1.2 Significance

Regarding the future configuration of metro rail networks, the integration of voice assistants represents a revolutionary and disruptive development. A new age of efficiency, control, and user-centricity might be brought in by this novel approach, which has the power to completely transform the passenger experience.
Voice assistant-facilitated real-time information cascades are at the foundation of this revolutionary change. Important information like train timetables, route modifications, and station upgrades are no longer unknown to commuters. Commuters can now get current and accurate data instantly via the voice assistant, which gives them a feeling of control and removes the uncertainty that often accompanies using public transportation. This significant adjustment helps to completely shift the commuting experience in its entirety.
The voice assistant makes things easier, such as checking fares and getting relevant information, in addition to the obvious advantages of time savings and less hassle at stations. The voice assistant essentially becomes a digital concierge, simplifying travel and promoting hassle-free commutes beyond its function as a simple tool.
We explore the many ways that this revolutionary technology might change everyday commuter routines. The focus is on encouraging passenger engagement in a more responsive, connected, and user-centered metro rail ecosystem—a notion that surpasses the traditional notion of system improvement.

We anticipate a transition towards a more informed, efficient, and seamlessly linked metro rail ecosystem by building the foundation for a future in which voice assistance technology becomes an essential component of every trip. The current study adds to the ongoing discussion about the future of urban transport by offering a prospective viewpoint on the incorporation of voice assistants. The research highlights the significant influence of metro rail on commuter experiences and the possibility of establishing a more accommodating and intuitive metro rail environment.

1.3 Aims and Objectives

The aim of integrating a voice assistant in the metro rail system is to streamline fare-related processes and general queries about the metro system and security. The following objectives are set to achieve this aim:
Creating a voice assistant that can answer ticketing-related questions and provide details about prices and routes. Allow for smooth interaction with the smart ticketing systems used by the metro train, enabling voice commands to be used by customers to buy tickets and for general queries.
Real-time information delivery of accurate and current train schedules, delays, and platform adjustments. Integrate with the data management systems of the metro rail to offer real-time updates on service advisories, disruptions, and detours to enhance passenger convenience.
Ascertain that the voice assistant can be used using voice commands, text-to-speech conversion options, and support for many languages by passengers with disabilities. Utilize tools like audio announcements and graphic displays.
This voice assistant aims to guarantee a seamless and generally accessible experience for each and every traveler. Among the specific goals is the creation of a voice assistant designed to provide smooth ticketing support and promote effective and user-friendly interactions. In addition, the integration seeks to work in parallel, enabling simple voice commands for ticket purchases and data retrieval, improving accessibility and ease of use.
The provision of accurate train information in real-time, the incorporation of multiple accessibility features, and the integration of service advisories with data management systems all serve to highlight the dedication to providing passenger-centric services. The overall goal is to develop a cutting-edge, inclusive, and user-friendly metro rail system that prioritizes passenger needs over operational issues.

1.4 Working Method

1. Speech Recognition API: We can integrate voice recognition into any HTML web page using the Web Speech API, a JavaScript function. Chrome versions 25 and later are compatible with this API. Speech-to-text conversion in web applications is supported via the Web Speech API. It's becoming more and more common and will dominate speech recognition in the future [13]. To precisely identify and transcribe the spoken words, this procedure entails evaluating and processing the audio input. This is used to deal with different accents, background noise, and variances in speech.
2. Natural Language Understanding (NLU): The AI voice assistant's NLU feature kicks in after speech-to-text conversion. The goal of NLU is to decipher the user's

meaning and intent from their spoken input. In order to extract pertinent data, such as user commands, queries, or requests, it analyzes the transcribed text. Voice assistants may now swiftly produce intelligible replies because of recent developments in computational linguistics, often known as natural language processing. These recent advancements in natural language processing, according to Hirschberg and Manning, are the result of four factors: (i) a significant increase in computing power; (ii) the availability of extremely large amounts of linguistic data; (iii) the development of highly successful machine learning (ML) methods; and (iv) a much deeper understanding of the structure of human language and how it is used in social contexts [14].
3. Dialog Management: Dialog management is in charge of managing convoluted user dialogues. It controls the flow of information and responses while maintaining the conversation's context. Dialog management ensures that the voice assistant may respond in a way that is logical and pertinent while taking into consideration the user's prior remarks or inquiries.

1.5 Problem Statement

Due to the growing number of passengers and trains, the metro rail service industry encounters difficulties in providing accurate and timely information to clients. Customers may find it challenging to immediately access the information they require. As consumer requirements increase, conventional approaches like customer care agents or static displays are becoming less effective and efficient. A more sophisticated and user-friendly system is required so that customers may have rapid access to accurate and current information.
Many consumers have reported having problems utilizing ticket vending machines due to technological issues. According to officials, there have been operational problems as a result of people trying to use the devices without fully comprehending how they operate. The noisy settings at metro stations make it difficult for voice assistants to understand spoken language. Speech recognition systems can be affected by background noise, which includes the sounds of trains, announcements, and crowds. Voice assistants may find it challenging to correctly perceive and comprehend consumer requests or inquiries as a result of these background disturbances.
Further complicating the voice assistant's job is the existence of diverse voices from people of different ages. Variations in pitch, tone, and pronunciation can affect the capacity of voice recognition systems to effectively perceive and react to user input since they are normally trained on a certain range of voices. This is especially crucial in the metro train service industry, since clients from all age groups and language backgrounds may use a voice assistant.
Communication issues between the voice assistant and the general public are also possible. Customers could struggle to communicate their needs in a way that the voice assistant can comprehend them, or they might not know how to successfully connect with one. Customers may become confused and irritated as a result of this. Technical challenges might arise when trying to distinguish regional languages spoken in Bangladesh by speakers with various accents. Voice assistants must be taught to comprehend and accommodate these regional variances, as accents can differ greatly within a single nation, to offer correct and pertinent information. Incorrect replies or misconceptions may result from failing to consider accent differences.

Finally, utilizing a voice assistant in the workplace might cause communication delays that can be problematic.

1.6 Research Objective

To help consumers with a variety of activities in metro train stations, this research focuses on the creation and implementation of an NLP-based model and voice assistance system that utilizes natural language processing (NLP) and incorporates automated question-answering technology. The following are further explanations of the research's goals:

1. The main objective is to develop a system that can assist users with various metro train service-related tasks. This involves giving out data on train schedules, assisting with ticketing services, figuring out prices, and offering route planning assistance. Customers will be able to communicate organically while getting accurate and timely information.
2. Investigating the possible benefits of using an NLP-based voice assistance system in the metro train service is a crucial component of the project. These benefits might include greater customer service efficiency, cost savings by automating some operations, and increased customer satisfaction.
3. The research attempts to assess the proposed system's usability from a customer standpoint and the veracity of the information given by it. Building trust and ensuring customer satisfaction depend on evaluating the system's capacity to deliver accurate and trustworthy information. By validating user experience and the accuracy of information this research aims to contribute insights into enhancing the trustworthiness and effectiveness of the proposed system.
4. Examining possible efficiency gains that may be made by putting the NLP model and voice assistant system into practice is another goal. Automating processes like ticketing or rate computations may enhance operations and cut down on the need for human involvement, which lowers costs and boosts the effectiveness of metro train services.
5. Strong data security measures, ongoing performance monitoring and optimization of the model, and extensive training programs for the voice assistant system to handle a variety of language and cultural situations are a few examples of these recommendations.
6. The project intends to contribute to ongoing efforts to improve the efficacy and user-friendliness of metro rail services by performing a thorough analysis of the proposed NLP model and speech recognition system. The research's conclusions and insights can guide upcoming changes and advancements in the metro rail service sector, which will eventually boost customer satisfaction and increase operational effectiveness.

Chapter 2

Literature Review

Technological growth can be found in every aspect of our daily lives. In metro rail services, various types of technologies are in use. For this technological solution, a closed-domain question-answering method is proposed to train a voice assistance model. Numerous pieces of research have been studied related to this proposed system and model. The study is elaborated briefly in multiple sections.

2.1 Natural Language Processing

Natural language processing is defined as ”A theoretically motivated range of computational techniques for analyzing and representing naturally occurring texts at the tone or more levels of linguistic analysis to achieve human-like language processing for a variety of tasks or applications [15] .”
As a fundamental field that aims to understand human language, natural language processing (NLP) is situated at the intersection of artificial intelligence, computer science, and computational linguistics. NLP has evolved, moving from rule-based to statistical techniques and, more recently, realizing the promise of deep learning techniques. The ambiguity and imprecision that characterize natural languages provide NLP with fundamental problems that hinder machine understanding. NLP is seen as an ”AI-complete” problem, meaning that it serves as a benchmark for developing artificial intelligence that is comparable to that of humans and allows computers to understand and learn any kind of intellectual work [30]. The literature emphasizes how crucial improved natural language processing (NLP) skills will be in shaping artificial intelligence in the future. This will result in smarter applications that interact with users in natural language and transform human-computer interaction with previously unheard-of accuracy and context awareness [39]. Also, there are many other applications of Natural Language Processing

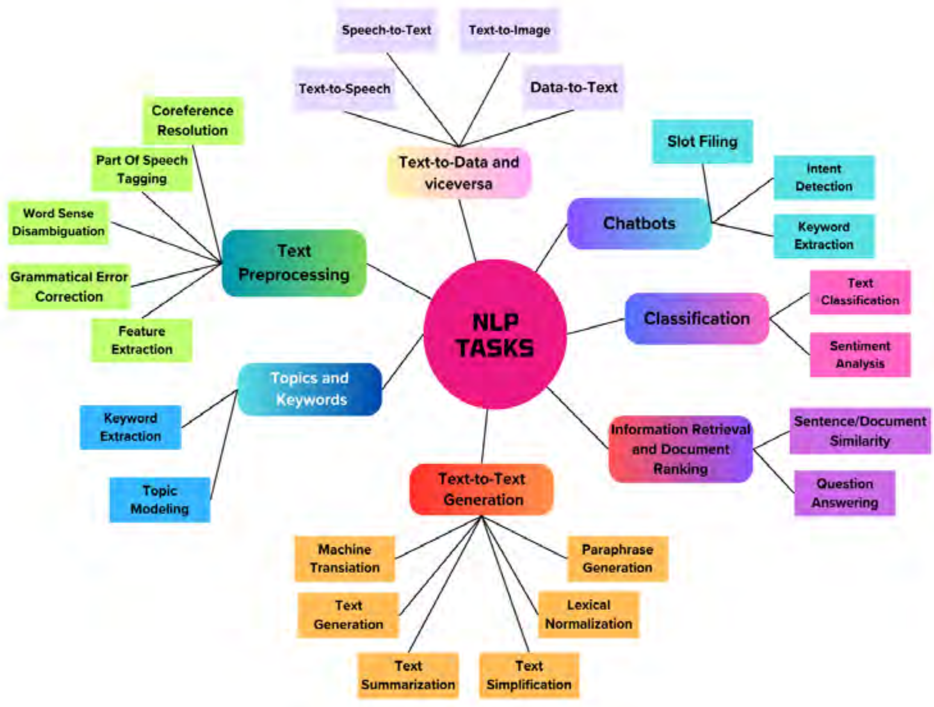


Figure 2.1: Different Applications of Natural Language Processing

2.2 A voice assistant that provides information

A paradigm for a personal assistant has been put out by Bibek Behera to reduce the workload placed on people when they do chores like ordering food or making travel arrangements. Chappie analyses conversations using natural language processing (NLP) to determine the user's purpose. It then conducts a discussion with the user using this data and AIML (Artificial Intelligence Mark-up Language) [17]. The project shows the implementation of a Voice Command System as an Intelligent Personal Assistant (IPA) that can perform a variety of tasks or services for an individual using Raspberry Pi as the main hardware to implement this model, which works on the primary input of a user's voice [22]. Additionally, Othman proposed paper on Voice Controlled Personal Assistant Using Raspberry Pi. A virtual voice-based personal intelligent assistant for visually impaired people was introduced by Aditya Sinha et al. This project will recognize user speech and reply to it effectively and quickly by voice, much as in a conversation. The improvement of conversational agents and speech recognition modules that can understand the Indian accent and work offline was the primary focus of this paper. The authors' strategy was to create an IPA that uses the Java libraries Sphinx-4, MaryTTS, and neural networks to embed learning capabilities [23].

2.3 Closed Domain Question Answering

In terms of the CDQA (Closed Domain Question Answering) method, there are two steps to implement, such as the training phase and the retrieving phase. The retrieval phase contains several processes such as vectorization, word embedding, tokenization, and ASR transcription, which are considered the Pre-processing part of training a CDQA-based model [6]. In this metro service voice assistance model, this CDQA method has been set to categorize the queries of people that can be asked in a metro rail station along with the most possible answers in the Bengali language. The sentence pre-processing methods are also known as the “bag of words” method. This method is commonly used in information retrieval systems [9]. In this model, after receiving a sentence from a user, the internal system breaks the words of the sentences separately, separating the punctuation through word lemmatization to keep the sentence understandable according to human spoken language, which the system has with various meaningful ask-able sentences [9]. There’s another co-occurrence or text pre-processing method available known as the pattern matching method, which is commonly used in conversational assistants and text mining [9] [4] . On the other hand, the question classification scheme can be another method for model training. Among three broad categories of classification, the procedure includes analyzing the syntactic construction of a question, suggesting the pattern of the question, and types of expected answers [12]. Through this methodology, even with a small dataset of questions with parallel answers, the algorithm is able to classify the queries and answer them accordingly. Feature extraction is another method of question classification according to different types of questions divided into different categories. This feature selects different categories and extracts the question and comparative answer from the existing dataset [18].

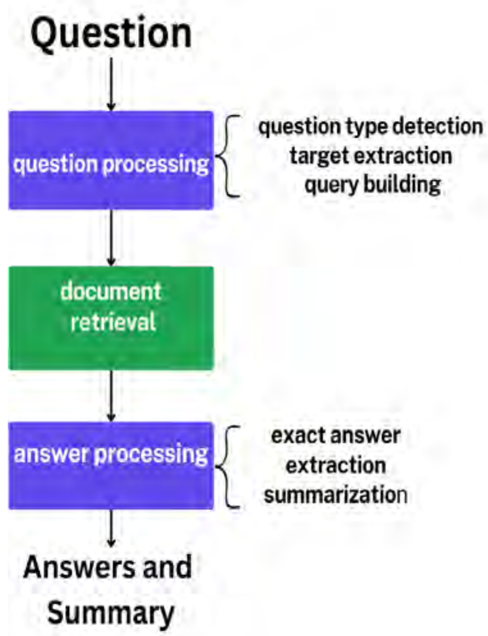


Figure 2.2: Simple Architecture of Closed Domain Question Answering

2.4 Integration with smart devices

In the paper ”Smart Home Using Internet of Things” by Keerthana S and teammates, the use of voice assistants has been elevated to a new level. They have explored how the integration of Internet of Things (IoT) and Wireless Fidelity (Wi-Fi) can result in the creation of a smart home system. They made use of the CC3200MCU, which has temperature sensors and Wi-Fi modules built in. The temperature sensor’s readings are sent to the microcontroller unit (MCU), which posts them to a server. With the help of this information, the state of electrical devices, such as fans and lights, is tracked and managed [19]. Although connected, voice assistants like Siri and Alexa are neither programmable nor adjustable. Natural language programming is also strongly connected to spoken programs, a notion that was formerly derided [5] but is now seeming more and more credible. Smart devices offer the best platform for offering customers of public transit real-time information and navigation services. The best ways to give passengers the most recent information about routes, timetables, delays, and other alternatives have been investigated by researchers. In order to provide individualized travel planning and navigation aid, including real-time bus arrival forecasts and transfer suggestions, Liang et al. (2018) created a smartphone application that combines real-time data from several sources [36]. Leopold and Amber’s [1] (1997) work on keyboardless programming, which was later reviewed by Arnold and Goldthwaite [2] (2001), Begel and Graham [3] (2005), and other researchers, is responsible for the earliest attempts to create speech programming systems. These efforts date back at least two decades. Users of these systems can enter code by speaking into a microphone, however, early speech programming systems had a restricted vocabulary.

2.5 Question answering in non-English languages

Large monolingual corpora, a lack of local speakers, the expense of correctly labelling such datasets, and the difficulty of selecting high-quality annotators make it difficult to curate substantial reading comprehension datasets for low-resource languages. Nevertheless, there has been work on translating English QA datasets to question-answering systems in Arabic [33], Korean [27], and Hindi [31]. They did this by translating English data, such as SQuAD 1.1 [29], to the appropriate languages and then using transformer models for transfer learning to train their own QA systems. We also closely follow these works for training our Bengali systems.
In addition to utilizing translated datasets, efforts have been made to curate sizable question-answering datasets in several other languages, such as French [34], Russian [35], Chinese [24]. These datasets have been used to train benchmark models such as LSTM and BERT [25]. A zero-shot transfer learning approach, in which pre-trained models were evaluated directly in a new language following task-specific training on question answering, has also been attempted on reading comprehension tasks as an alternative to obtaining translated or human-annotated datasets for model training. To our current knowledge, no prior effort at comparable work has been made in Bengali.

2.6 Question answering in Bengali Language

Bengali question-answering research is in its infancy compared to the maturity of English QA research because of a lack of basic natural language processing tools and a data shortage. Attempts to develop factoid-like QA systems employing information retrieval-oriented methods or question categorization systems to be employed as a component of full-fledged question-answering systems comprise the majority of research on QA conducted in Bengali.
To create a system for classifying questions in Bengali, Banerjee and Bandyopadhyay [8] first extracted lexical, syntactic, and semantic features, such as the existence and position of wh-words or interrogative words, question length, end marker, Parts of Speech (POS) tags, etc., and categorized the questions into nine groups. Compared to English, Bengali has a lot more interrogative terms. Interrogatives may also occur in any portion of a sentence. Finally, there aren't as many high-quality NLP tools as there are in Bengali, such as POS taggers, NER systems, and benchmark corpora. In addition, Nirob et al. [20] used uni-grams as features in conjunction with support vector machines to construct a question categorization system employing comparable feature extraction techniques.
To create the first Bengali factual question-answering system, Banerjee et al. [11] developed BFQA, a sophisticated IR system that categorizes questions, finds pertinent phrases, rates them, and extracts the right responses. A multilingual question-answering system called BQAS [16] was also developed. It can create and respond to fact-based queries derived from both English and Bengali materials. A comparable QA system that pulls pertinent responses from many documents and provides targeted responses for queries pertaining to time was also put into place by Islam and Nurul Huda [32].

2.7 Voice Assistance in different Systems

Obtaining the answers to consumer's inquiries is the voice assistant's job. Regarding the user's question, they provide pertinent information. Voice assistants nowadays are capable of processing purchases, responding to inquiries, and carrying out various duties including making phone calls or playing music. There are a tonne of voice assistant-related ventures that are gaining traction and offering consumers excellent choices. In general, it is useful in their place of employment. The world's top voice assistants include Cortana, Siri, Amazon Echo, Bixby, Blackberry Assistant, Google Home, Google Assistant, and Nina [26]. When a user gives a voice command to these devices, the data is converted into text, understood by the devices, and then sent to the speaker, an output device, which responds to the command. Other key technologies in this sector include text-to-speech, voice translators, automated speech recognition, voice activation, and named entity recognition. In smartphones, a voice assistant was released as an app that teaches how to utilize natural language processing for voice-only usage of built-in applications and message sending. Subsequently, it was put to the test for voice-activated event creation and email sending via the calendar. Subsequently, the program was enhanced with additional features including voice-only application launching, reminder setting, and music playback. These days, smart speakers—such as Google Home and Alexa—are utilized to do tasks, which is a terrific invention [38].

2.8 TF-IDF

Multiple studies have extracted the TF-IDF characteristic extensively. [7] Pre-processed the dataset, which included 272 questions gathered from various sources. Then, TF-IDF was computed and input into Linear SVM, yielding results that were acceptable for accuracy and precision but unsatisfactory for F-measure and recall. The same TF-IDF feature was used in [10] to expand this study, which included 600 questions in the dataset and tested the impact of categorization with and without stop words. Thus, eliminating stop words has little effect on the outcome. Additionally, [28] made use of the same dataset, but they added an impact factor to the increased TF-IDF. Three classifiers, NB, KNN, and SVM, were then used to manage the classification process; SVM outperformed the other classifiers in terms of performance. To address the problem of keywords overlapping, [21] combined four classifiers—rule-based, SVM, KNN, and NB—using the majority voting algorithm and WordNet similarity. The 100 programming questions in the dataset utilized in this study are split into two sets: a training set of 60 questions and a test set of 40 questions. By reaching a 95% F-measure, the data demonstrate that the ensemble outperforms each classifier.

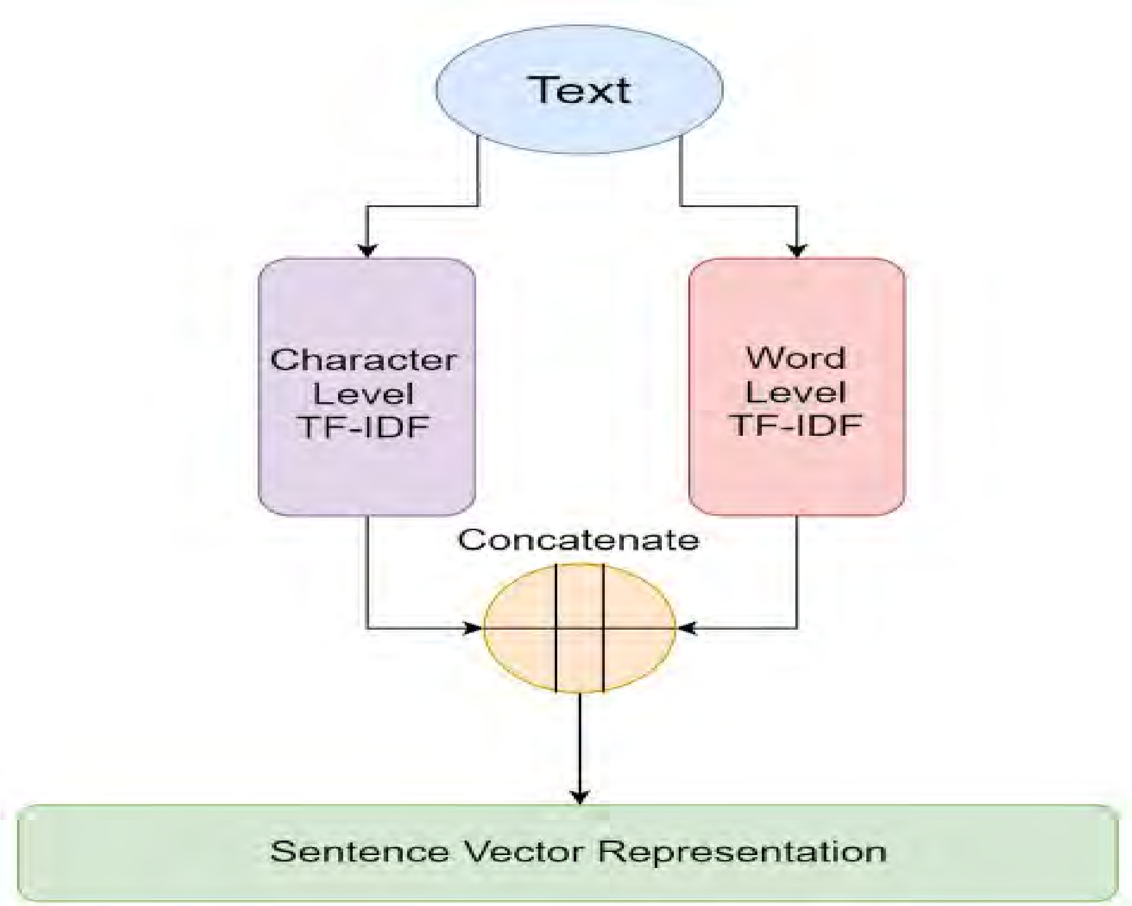


Figure 2.3: TF-IDF vectorization process

Chapter 3

Dataset, Data Analysis, and Data Pre-processing

3.1 Data Collection Process

This section offers a thorough rundown of the dataset collected to establish a voice assistance system specifically designed for use with metro rail services. The collection was carefully selected to meet the unique requirements and questions of the general public at metro train stations. The dataset, which consists of around 866 instances, was diligently collected through online Google Forms questionnaires and on-site surveys at metro rail stations, to foster a smooth and contextually appropriate dialogue.

3.1.1 On-Site Surveys

To gather actual and pressing issues of passengers, surveys were completed at different metro rail stations. Open-ended questions on metro train services were presented to participants, who were also asked to share any comments, complaints, or preferences they may have had in Bengali. The objective of the in-person surveys was to collect current information on frequently asked questions by travelers at metro rail stations.

Figure 3.1: On-Site Survey

3.1.2 Online Surveys via Google Forms

Google Forms was used for managing online surveys in order to gather more data. A series of prepared questions and responses were given to the participants to simulate a normal discussion with the voice assistance system. With a wide variety of situations and user queries, our strategy produced a structured yet flexible dataset.

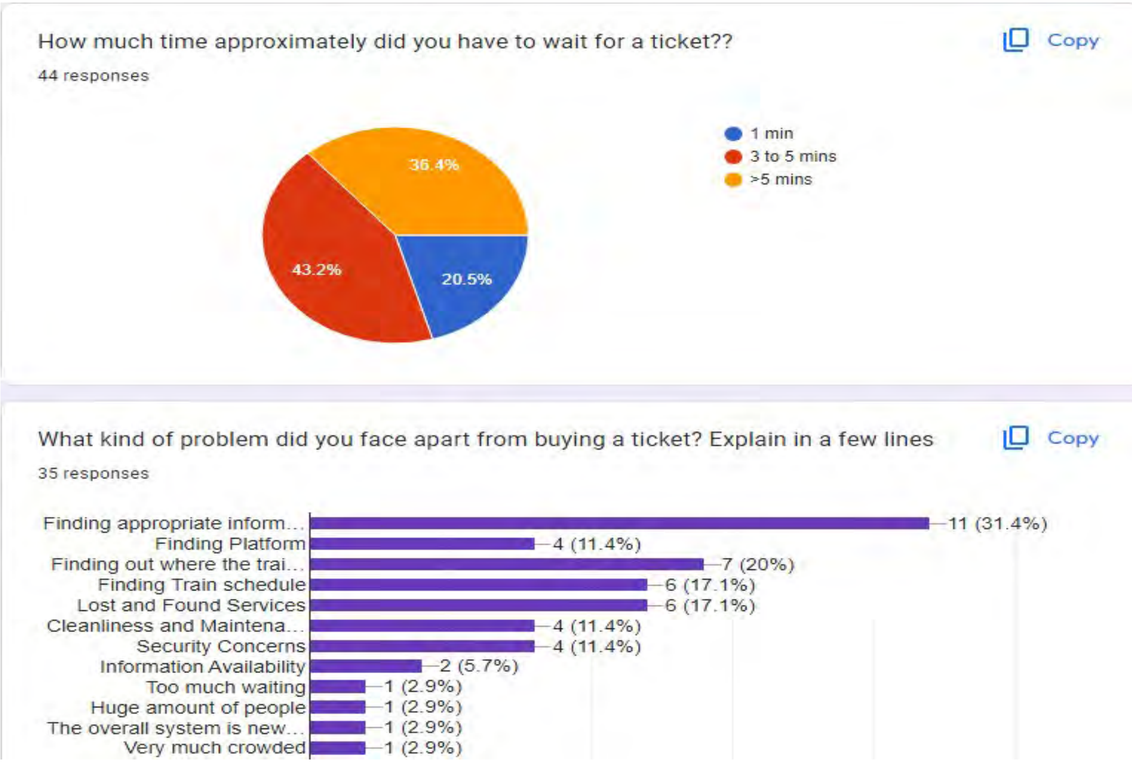


Figure 3.2: Online Surveys

3.2 Dataset Description

The most important part of training an AI model is the dataset. According to the dataset structure, different models can be designed to train a system. According to the proposed project of a voice assistant for metro rail services, the system requires a dataset, that can be trained to interact with humans in metro rail stations. In recent eras, voice assistant technology has marked a paradigm shift in terms of human-computer interactions. The intelligence systems can understand natural language queries along with a responsive feature to interact, which has become an integral part of our daily lives. Following the advancement and necessity of voice assistants in daily life we have formed a unique dataset for training such model specifically in **Bengali Language**. The dataset presented here takes a novel approach while designing which is a text-based interaction (question-answer) in Bengali to train a voice assistant model.This approach not only enlarges the accessibility of voice assistant technology but also addresses challenges specific to linguistic shades in the Bengali language.

Question	Answer	Category
হ্যালো মেট্রো	হ্যালো,আপনাকে কিভাবে সাহায্য করতে পারি?	service
উত্তরা থেকে মতিঝিল যেতে কতক্ষন লাগবে?	হ্যালো মেট্রো	service
মেট্রোতে কোন সময় ভিড় বেশি থাকে?	সকাল ৮টা থেকে ১০টা এবং বিকাল ৫টা থেকে ৭টার মধ্যে মেট্রোতে যাত্রীদের ভিড় সবচেয়ে বেশি থাকে।	travelinfo
এমআরটি পাস হারিয়ে গেলে কি করব	কার্ড হারিয়ে গেলে নিকটস্থ কাস্টমার কেয়ার সেন্টারে উল্লেখ করতে হবে। এই নিয়মে একটি নতুন কার্ড জারি করা হবে।	service
ট্রেন কখন আসবে?	ট্রেন আর মাত্র দশ মিনিট পর আসবে	travelinfo
দুর্ঘটনার ক্ষেত্রে কার সাথে যোগাযোগ করতে হবে?	কোন দুর্ঘটনার ক্ষেত্রে তা□ক্ষণিকভাবে নিকটস্থ মেট্রো কর্মী বা হেল্পলাইন নম্বরে ফোন করতে হবে।	service
মহিলাদের কোচে কতজন বসতে পারবে?	প্রতিটি মহিলাদের জন্য বরাদ্দকৃত কোচে সর্বোচ্চ ৫০ জন করে মোট ১০০ জন বসতে পারবে।	traininfo
কার্ড এর মেয়াদ কত বছর থাকে?	মেট্রো কার্ডগুলো সাধারণত ১০ বছরের জন্য বৈধ। এরপর নতুন করে বানানো লাগবে।	ticket/card
মেট্রোতে ট্রেনের ভিতরে ধূমপান কি শাস্তি আছে?	হ্যাঁ, মেট্রো ট্রেনের ভিতরে সিগারেট ধরা পড়লে জরিমানার সাথে সাথে কারাদণ্ডও রয়েছে।	security
এই স্টেশনে মেট্রো ট্রেনের যাত্রী সংখ্যা কত?	প্রতি দিনে এই স্টেশনে প্রায় ৫০,০০০ থেকে ৬০,০০০ জন মেট্রো ট্রেনের যাত্রী চলছে।	stationinfo
উত্তরা উত্তর থেকে আগারগাঁও যাওয়ার ভাড়া কত ?	ষাট টাকা	fare
পল্লবী থেকে মিরপুর ১১ যাওয়ার ভাড়া কত ?	বিশ টাকা	fare
শেওড়াপাড়া থেকে মতিঝিল যাবো	পঞ্চাশ টাকা	fare
বিজয় সরণি থেকে উত্তরা সেন্টার যাবো	ষাট টাকা	fare
ঢাকা মেট্রো রেলের স্টেশন কতটি?	ঢাকা মেট্রো রেলের প্রথম ধাপে ১৬টি স্টেশন রয়েছে।	traininfo
ট্রেন কতক্ষন প্রতি স্টেশনে দাঁড়ায়	১০ থেকে ১৫ সেকেন্ড	traininfo
আমি মিরপুর যেতে চাই	আপনি কয়টি টিকেট কাটতে চাচ্ছেন?	travelinfo
প্রতিটি ট্রেনে এ কয়টি কোচ?	ট্রেনগুলো ৬ কোচ বিশিষ্ট তবে ভবিষ্যতে ৮ কোচে উন্নীত করা যাবে।	metroinfo
মেট্রোয় বিনামূল্যে কি ভ্রমণ করা যায়?	না, প্রত্যেক যাত্রীকে অবশ্যই টিকিট কিনে ভ্রমণ করতে হবে। বিনামূল্যে ভ্রমণ বেআইনি।	ticket/card

Table 3.1: Bengli Dataset, differentiated with multiple Categories and queries along with respective answers

The dataset consists of around 866 instances. The xlsx dataset-shaped questions are categorized with multiple domains, associated with corresponding answers. These domains differentiate human-askable questions according to human asking styles. There are 8 categories in the dataset.

Category	Count
service	35
travel info	61
ticket/card	31
metro info	35
security	33
station info	36
train info	28
fare	606

Table 3.2: Categories in the dataset

On the other hand, a **Json** dataset is also created to train another model to find the accuracy level higher than the previous one containing the same features of categorization with tags as patterns (questions) with respective answers. Therefore, we aim to exaggerate the accuracy and sufficiency of voice assistant models specifically tailored for metro services for the Bengali-speaking population.

3.3 Conversion of XLSX Dataset to JSON Data

During the preparation phase of our research, one of the most important and rigorously carried out procedures was converting the Microsoft Excel (.xlsx) dataset to a structured JSON (JavaScript Object Notation) format. With the use of Python and the pandas library, we were able to import the dataset and convert it into a format that met the complex needs of our advanced NLP-based question-answering system. Each row was repeatedly analyzed by the Python script, which identified the questions as intents and linked them to the appropriate answers to create a hierarchical structure. This method made sure that the data was consistent and also made it possible to provide each intent its own tag, or unique identification. The final JSON file is called "intents bangla.json," and it contains a dictionary of intents, each of which consists of patterns (questions) and replies (answers). The use of tags makes it easier to pinpoint each purpose precisely. The script's clarity and adaptability demonstrate our dedication to a systematic and reliable approach to dataset preparation for later phases of model development and assessment. This approach meets the latest data handling guidelines, paving the way for the deployment of an advanced natural language processing (NLP) system that has the potential to completely transform the way users interact with metro rail systems.

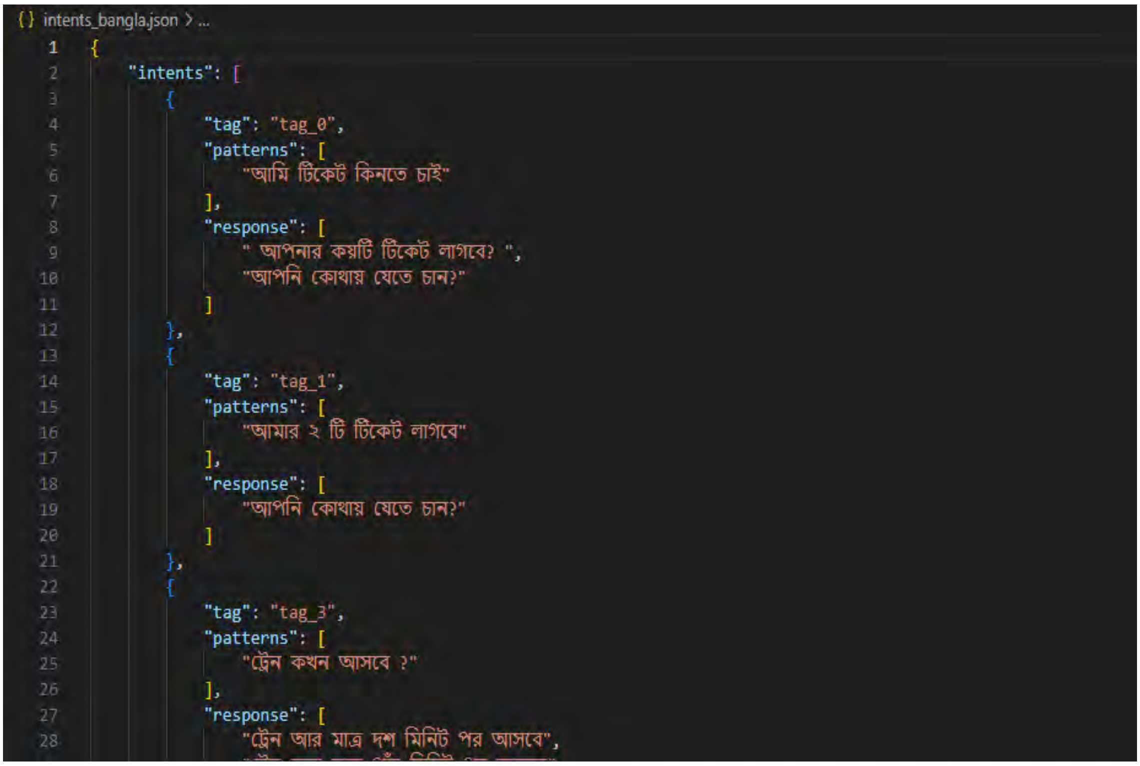


Figure 3.3: The JSON Dataset, differentiated with multiple tags and queries along with respective answers

3.4 Data Pre-processing

The data pre-processing is divided into multiple segments known as data cleaning, data training along with sub-segments.
The first section deals with data cleaning, which is a process of carefully examining and improving raw data by dealing with missing numbers, outliers, and inaccuracies. By doing this, the data's quality and integrity are guaranteed, providing a solid basis for any further analysis. The procedure then continues onto the data training segment, where sub-segments are essential for getting the data ready for model creation. Feature ranges are standardized via feature scaling and normalization, which keeps any one feature from excessively impacting the model. While data augmentation approaches increase the diversity of the training set and are especially useful in situations when there is a lack of data, categorical variable encoding converts non-numeric data into a format that is appropriate for machine learning algorithms. The well-structured methodology, together with its well-crafted constituents, smoothly converts unprocessed data into a refined, standardized, and enhanced format, hence facilitating efficient model training and ensuing analysis. Such architecture promotes transparency and reproducibility in research and analytical initiatives in addition to streamlining the data pre-processing workflow.

3.4.1 Data Cleaning

Data cleaning is a crucial process before training a machine learning model. A well-prepared data ensures the model's accuracy and generalization.As the proposed dataset is a text-based dataset which are basically used for natural language pro-

Figure 3.4: Data Pre-processing methods in our dataset

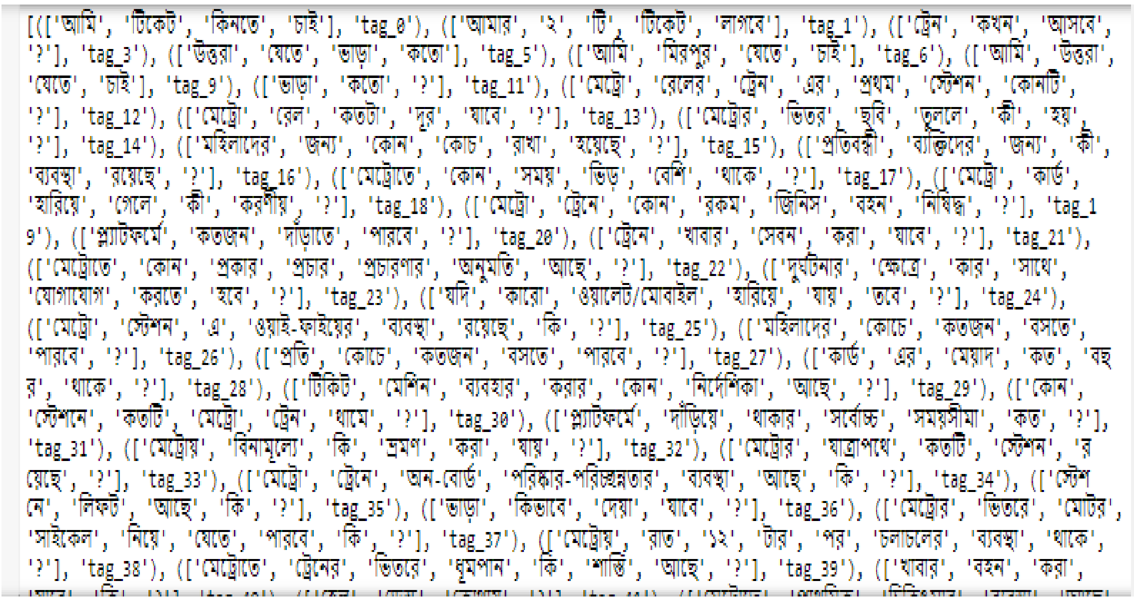

[(['আমি', 'টিকেট', 'কিনতে', 'চাই'], 'tag_0'), (['আমার', '২', 'টি', 'টিকেট', 'লাগবে'], 'tag_1'), (['ট্রেন', 'কখন', 'আসবে', '?'], 'tag_3'), (['উত্তরা', 'যেতে', 'ভাড়া', 'কতো'], 'tag_5'), (['আমি', 'মিরপুর', 'যেতে', 'চাই'], 'tag_6'), (['আমি', 'উত্তরা', 'যেতে', 'চাই'], 'tag_9'), (['ভাড়া', 'কতো', '?'], 'tag_11'), (['মেট্রো', 'রেলের', 'ট্রেন', 'এর', 'প্রথম', 'স্টেশন', 'কোনটি', '?'], 'tag_12'), (['মেট্রো', 'রেল', 'কতটা', 'দূর', 'যাবে', '?'], 'tag_13'), (['মেট্রোর', 'ভিতর', 'ছবি', 'তুললে', 'কী', 'হয়', '?'], 'tag_14'), (['মহিলাদের', 'জন্য', 'কোন', 'কোচ', 'রাখা', 'হয়েছে', '?'], 'tag_15'), (['প্রতিবন্ধী', 'ব্যক্তিদের', 'জন্য', 'কী', 'ব্যবস্থা', 'রয়েছে', '?'], 'tag_16'), (['মেট্রোতে', 'কোন', 'সময়', 'ভিড়', 'বেশি', 'থাকে', '?'], 'tag_17'), (['মেট্রো', 'কার্ড', 'হারিয়ে', 'গেলে', 'কী', 'করণীয়', '?'], 'tag_18'), (['মেট্রো', 'ট্রেনে', 'কোন', 'রকম', 'জিনিস', 'বহন', 'নিষিদ্ধ', '?'], 'tag_1
9'), (['প্ল্যাটফর্মে', 'কতজন', 'দাঁড়াতে', 'পারবে', '?'], 'tag_20'), (['ট্রেনে', 'খাবার', 'সেবন', 'করা', 'যাবে', '?'], 'tag_21'), (['মেট্রোতে', 'কোন', 'প্রকার', 'প্রচার', 'প্রচারণার', 'অনুমতি', 'আছে', '?'], 'tag_22'), (['দুর্ঘটনার', 'ক্ষেত্রে', 'কার', 'সাথে', 'যোগাযোগ', 'করতে', 'হবে', '?'], 'tag_23'), (['যদি', 'কারো', 'ওয়ালেট/মোবাইল', 'হারিয়ে', 'যায়', 'তবে', '?'], 'tag_24'), (['মেট্রো', 'স্টেশন', 'এ', 'ওয়াই-ফাইয়ের', 'ব্যবস্থা', 'রয়েছে', 'কি', '?'], 'tag_25'), (['মহিলাদের', 'কোচে', 'কতজন', 'বসতে', 'পারবে', '?'], 'tag_26'), (['প্রতি', 'কোচে', 'কতজন', 'বসতে', 'পারবে', '?'], 'tag_27'), (['কার্ড', 'এর', 'মেয়াদ', 'কত', 'বছ
র', 'থাকে', '?'], 'tag_28'), (['টিকিট', 'মেশিন', 'ব্যবহার', 'করার', 'কোন', 'নির্দেশিকা', 'আছে', '?'], 'tag_29'), (['কোন', 'স্টেশনে', 'কতটি', 'মেট্রো', 'ট্রেন', 'থামে', '?'], 'tag_30'), (['প্ল্যাটফর্মে', 'দাঁড়িয়ে', 'থাকার', 'সর্বোচ্চ', 'সময়সীমা', 'কত', '?'], 'tag_31'), (['মেট্রোয়', 'বিনামূল্যে', 'কি', 'ভ্রমণ', 'করা', 'যায়', '?'], 'tag_32'), (['মেট্রোর', 'যাত্রাপথে', 'কতটি', 'স্টেশন', 'র
য়েছে', '?'], 'tag_33'), (['মেট্রো', 'ট্রেনে', 'অন-বোর্ড', 'পরিষ্কার-পরিচ্ছন্নতার', 'ব্যবস্থা', 'আছে', 'কি', '?'], 'tag_34'), (['স্টেশ
নে', 'লিফট', 'আছে', 'কি', '?'], 'tag_35'), (['ভাড়া', 'কিভাবে', 'দেয়া', 'যাবে', '?'], 'tag_36'), (['মেট্রোর', 'ভিতরে', 'মোটর', 'সাইকেল', 'নিয়ে', 'যেতে', 'পারবে', 'কি', '?'], 'tag_37'), (['মেট্রোয়', 'রাত', '১২', 'টার', 'পর', 'চলাচলের', 'ব্যবস্থা', 'থাকে', '?'], 'tag_38'), (['মেট্রোতে', 'ট্রেনের', 'ভিতরে', 'ধূমপান', 'কি', 'শাস্তি', 'আছে', '?'], 'tag_39'), (['খাবার', 'বহন', 'করা',

Figure 3.5: Words from the JSON data with their corresponding tags.

```
[nltk_data] Downloading package wordnet to
[nltk_data]     C:\Users\User\AppData\Roaming\nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
[nltk_data] Downloading package omw-1.4 to
[nltk_data]     C:\Users\User\AppData\Roaming\nltk_data...
[nltk_data]   Package omw-1.4 is already up-to-date!
['অত্যাধুনিক', 'অন-বোর্ড', 'অনুমতি', 'অর্থায়ন', 'আওয়ারে', 'আগারগাঁও', 'আগারাগাঁও', 'আছে', 'আছেন', 'আমার', 'আমি', 'আর', 'আ
রো', 'আলাদা', 'আসছে', 'আসনের', 'আসবে', 'ইতিহাস', 'উত্তর', 'উত্তরা', 'উদ্বোধন', 'ঋণ', 'এ', 'এই', 'এমআরটি-১', 'এমআরটি-৬',
'এয়ার', 'এর', 'এলাকায়', 'ওয়াই-ফাইয়ের', 'ওয়ালেট/মোবাইল', 'কখন', 'কত', 'কতগুলো', 'কতজন', 'কতটা', 'কতটি', 'কতো', 'কবে',
'কমলাপুর', 'কয়টি', 'করছে', 'করণীয়', 'করতে', 'করা', 'করার', 'করে', 'কর্মকর্তা-কর্মচারী', 'কর্মরত', 'কাজ', 'কাজিপাড়া', 'কার', 'কার
ওয়ান', 'কারা', 'কারো', 'কার্ড', 'কার্ডের', 'কি', 'কিনতে', 'কিভাবে', 'কিলোমিটার', 'কী', 'কে', 'কোচ', 'কোচে', 'কোথায়', 'কোথায়',
'কোন', 'কোনটি', 'কোনো', 'ক্ষমতা', 'ক্ষেত্রে', 'কয়টি', 'খাবার', 'খুব', 'খোলা', 'গতি', 'গতির', 'গরম', 'গুলো', 'গুলোর', 'গেলে', 'চ
লমান', 'চলাচলের', 'চাই', 'চার্জ', 'চিকিৎসার', 'চেঞ্জ', 'ছবি', 'ছিল', 'জন', 'জন্য', 'জানা', 'জানাতে', 'জাপানী', 'জিনিস', 'টয়লেট',
'টয়লেটগুলোতে', 'টয়লেটে', 'টয়লেটের', 'টাকা', 'টার', 'টি', 'টিকিট', 'টিকেট', 'টিস্যুর', 'ট্রাফিক', 'ট্রেন', 'ট্রেনে', 'ট্রেনের', 'ডেস্ক',
'ঢাকা', 'তবে', 'তুললে', 'তৈরি', 'তো', 'থাকবে', 'থাকলে', 'থাকার', 'থাকে', 'থামে', 'থেকে', 'দক্ষিণ', 'দরজা', 'দাঁড়াতে', 'দাঁড়িয়ে',
'দাম', 'দিতে', 'দিলে', 'দুর্ঘটনার', 'দূর', 'দূরত্বে', 'দেখা', 'দেয়া', 'দ্বিতীয়', 'দ্রুতগামী', 'ধরনের', 'ধাপ', 'ধাপটি', 'ধাপে', 'ধাপের', 'ধূম
পান', 'নতুন', 'নিয়ন্ত্রক', 'নিয়ে', 'নিরাপত্তার', 'নিরাপদ', 'নির্দেশিকা', 'নির্মাণ', 'নির্মাণকাজের', 'নির্মিত', 'নিষিদ্ধ', 'নেওয়া', 'নেতৃত্বাধীন',
'ন্যাপি', 'পর', 'পরিকল্পনা', 'পরিকল্পনায়', 'পরিচালনা', 'পরিবহন', 'পরিবহনের', 'পরিমাপ', 'পরিষ্কার-পরিচ্ছন্ন', 'পরিষ্কার-পরিচ্ছন্নতার', 'পল্ল
বী', 'পানির', 'পারবে', 'পারি', 'পারেন', 'পার্কিংয়ের', 'পিক', 'পৃথক', 'প্রকার', 'প্রচার', 'প্রচারণার', 'প্রতি', 'প্রতিটি', 'প্রতিটি', 'প্রতিব
ন্ধী', 'প্রতিবন্ধীদের', 'প্রথম', 'প্রাথমিক', 'প্ল্যাটফর্মে', 'ফার্মগেট', 'ফ্রেশনার', 'বছর', 'বসতে', 'বসার', 'বহন', 'বা', 'বাংলাদেশ',
'বাংলাদেশে', 'বাইরে', 'বাচ্চা', 'বাজার', 'বিজয়', 'বিদ্যুতস্থাপনার', 'বিনামূল্যে', 'বিশেষ', 'বিশ্ববিদ্যালয়', 'বিস্তার', 'বেশি', 'ব্যক্তিদের', 'ব্যব
স্থা', 'ব্যবহার', 'ব্যয়ে', 'ব্যাগেজ', 'বড়/', 'ভবিষ্যতে', 'ভাড়া', 'ভাড়া', 'ভিড়', 'ভিতর', 'ভিতরে', 'ভেতরে', 'ভ্রমণ', 'ভ্রমণের', 'মতিঝিল',
'মহিলাদের', 'মিরপুর', 'মেট্রো', 'মেট্রোতে', 'মেট্রোয়', 'মেট্রোর', 'মেট্রোরেলের', 'মেয়াদ', 'মেশিন', 'মেয়াদ', 'মোটর', 'যন্ত্রের', 'যদি', 'যন্ত্র
পাতি', 'যাওয়ার', 'যাচ্ছে', 'যাতায়াত', 'যাত্রা', 'যাত্রাপথে', 'যাত্রী', 'যাবে', 'যায়', 'যুক্ত', 'যেতে', 'যোগাযোগ', 'রকম', 'রয়েছে', 'রাখতে',
'রাখা', 'রাত', 'রুট', 'রেল', 'রেলকে', 'রেলে', 'রেলের', 'রয়েছে', 'লক্ষ্যমাত্রা', 'লম্বা', 'লাইন', 'লাগবে', 'লিফট', 'লেগে', 'লেন', 'শ
ক্তি', 'শান্তি', 'শাহবাগ', 'শিশুর', 'শুরু', 'শেওরাপাড়া', 'শেওড়াপাড়া', 'শেষ', 'সংক্ষিপ্ত', 'সংখ্যা', 'সংখ্যার', 'সংস্থা', 'সচিবালয়', 'সব', 'স
ময়', 'সময়কাল', 'সময়সীমা', 'সময়ে', 'সমস্যা', 'সরণি', 'সরবরাহ', 'সর্বোচ্চ', 'সহ', 'সাইকেল', 'সাইবার', 'সাথে', 'সাবান', 'সাবানের',
'সার্ভিস', 'সাহায্য', 'সিগনাল', 'সিস্টেম', 'সিস্টেমে', 'সুবিধা', 'সেটা', 'সেন্টার', 'সেবন', 'স্টেশন', 'স্টেশনে', 'স্টেশনের', 'হচ্ছে', 'হতে',
'হবে', 'হয়', 'হয়েছে', 'হলে', 'হার', 'হারিয়ে', 'হেল্প', 'হ্যান্ডওয়াশ', 'হয়', '১০', '১১', '১২', '২']
```

Figure 3.6: Total Word list after excluding the Ignore letters.

cessing(NLP) related works, the cleaning process contains multiple segments such as tokenization, Bag of Words(BoW) technique, ignoring punctuations. The "Bag of Words"(BoW) is one of the most common techniques in natural language processing(NLP) and machine learning to represent data. It is a simple way to covert a text into numerical vectors, which machine learning models can use as inputs. This segment contains some sub-sections as internal works such as **tokenization** which is a process that breaks the words individually of a sentence and add a numerical tag corresponding to each of them to be understandable by the machine. Removing stop words to eliminate common words that contains lower weight which means the word doesnot carry much meaning of the sentence. After these sub processes, word lemmatization occurs which reduce words to their base form to make the text standardize. The standardized words are used to find similarity according to their weight with the proposed dataset to generate the respective meaningful answer in contrast to the question.

3.4.2 Handling Noisy Data

Handling noisy data in a text-based dataset removes the HTML Tags and Special Characters, Spell checking and Correction, Remove Stop Words, Tokenization and Lemmatization,Use regular expressions to identify associated patterns to filter out the noisy data.
As in this research CDQA(closed domain question answer) method is used,some more segments are important other than the general terms such as removing irrelevant sentences, filtering outliers,standardize terminology to ensure the consistency in terminology by standardizing synonyms of words which will refer to the same concept or entity, user feedback analysis which refers to analyze user feedback and update the dataset according to their suggestions, if available.
Acknowledging the distinct difficulties presented by this particular domain, the preprocessing pipeline is expanded to include specialized sections designed for maximum efficiency.

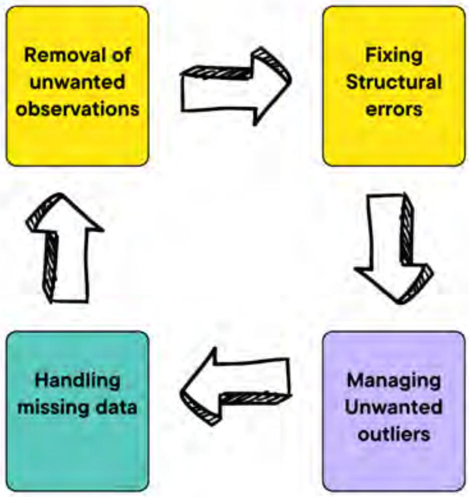


Figure 3.7: Data Cleaning

To guarantee that the model only considers relevant data pertaining to metro train announcements, one essential component is the elimination of unnecessary sentences. By streamlining the dataset, this filtering procedure improves the model's ability to recognize and answer pertinent queries.
Finding and eliminating outliers in the speech data is another crucial factor to take into account. The accuracy of the model might be negatively impacted by outliers, which can appear as anomalies or unexpected deviations in speech patterns. The dataset can be improved to exclude abnormal occurrences by implementing outlier detection techniques, guaranteeing a more stable and dependable model.
The part on standardizing terminology is noteworthy since it attempts to attain linguistic uniformity among various metro rail announcements. To provide a consistent and coherent dataset, this entails standardizing term usage in addition to eliminating synonyms. The model can read and react to user inquiries more efficiently by taking into account variances in language and terminology.
Analysis of user input adds a dynamic component to the preprocessing procedure. In order to update the dataset iteratively, this section entails examining user input and adding pertinent suggestions. By using an adaptive approach, the model is guaranteed to change over time and become more sensitive to the preferences and demands of the user, which will improve its ability to comprehend and answer to a variety of inquiries.

3.5 Data Training

This is the crucial part after the data pre-processing and the major part to train new models. It involves transforming raw data in an effective format and cleaning to use it for training. Data cleaning process basically handles the missing values to decide whether to remove missing values or to fill them using different methods like mean, median or interpolation. Moreover, removing duplicate entities from the dataset also considered as pre processing.

3.5.1 Data Normalization

Normalizing data is to ensure that the dataset becomes more cleaner and standardized to facilitate effective training of a voice assistant. Specific steps may vary based on the characteristics of the proposed dataset and also on the requirements of a voice assistant application. Normalization usually means scaling numerical features within a uniform range to avoid any one feature having an undue impact on the training process. Normalization can involve a variety of methods, especially for voice assistant programs, which frequently handle a variety of audio inputs. Feature scaling is a popular technique that entails converting numerical information to a standard scale, including min-max scaling or z-score normalization. This makes sure that elements with disparate scales don't interfere too much with learning.

Moreover, extra normalizing operations could be required due to the properties of audio data. For this research amplitude and frequency changes are common in audio communications. To normalize these variances and make the model resilient to varying input volumes and frequencies, methods like dynamic range compression and signal normalization are used.

Chapter 4

Methodology, Architecture, and Model Specification

This vital phase of development, creation, and refinement highlights how important technological innovation is to change how people interact with metro trains. It represents our steadfast dedication to pushing the boundaries of traditional transportation systems and our goal of creating a smooth, user-centered experience inside the intricate realm of metro rail operations. This section covers the planning and development of the question-answering system as well as the fine-tuning of the models.

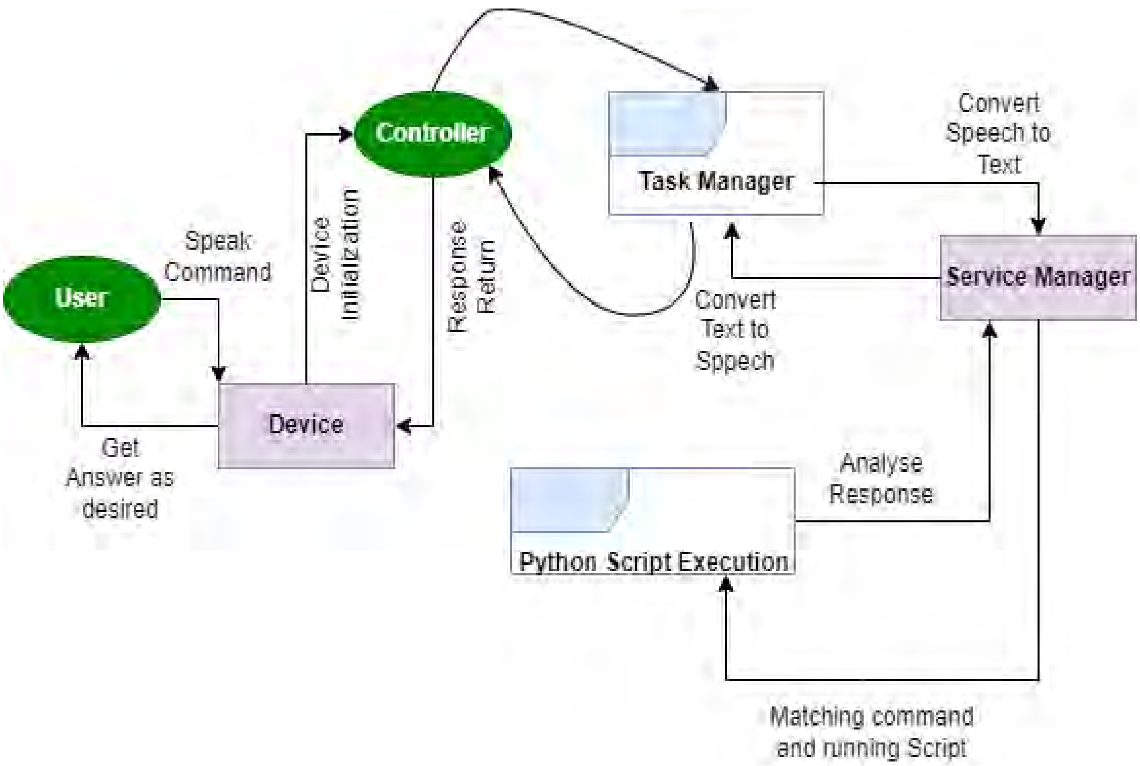


Figure 4.1: Working Method

4.1 Technologies Used

The process of creating an effective question-answering system for metro rail user engagement required careful consideration of many different technologies. Each technology was specifically selected to enhance the system's efficacy, scalability, and resilience.

Python: Our system's basic language for both traditional and scientific computing is Python. Python, well known for its readability and flexibility, allowed for easy integration with several frameworks and modules that are essential for data manipulation, machine learning, and natural language processing. The whole development life cycle benefited greatly from the language's vast ecosystem and effectiveness.

Node.js: Node.js is a strong back-end programming language that was developed using the V8 JavaScript engine in Chrome. Our system found great value in its event-driven, non-blocking design for managing several requests at once. The system's responsiveness and scalability were enhanced by using Node.js's advantages, ensuring peak use situations and excellent performance.

React: The question-answering system's user-friendly and responsive front-end was created in large part thanks to the JavaScript user interface framework React. Its component-based design made it easier to create dynamic, modular user interfaces, which improved the user experience as a whole. The declarative methodology of React simplified the rendering process, allowing for effective system updates and interactions.

Flask: The development of the system's web-based user interface was made possible using Flask, a Python micro web application framework. A scalable and maintainable interface may be created thanks to its lightweight and modular architecture. Python and Flask worked together seamlessly to improve the system's overall coherence and flexibility to accommodate changing needs.

Postman: A collaborative tool that has become essential for streamlining the creation and assessment of the REST-API, a fundamental part of our system, is Postman. Its easy-to-use interface and strong testing capabilities made API development more efficient and guaranteed smooth communication between various system components. Postman's teamwork-oriented features expedited the API development process and enabled efficient team communication.

gTTS: Google's text-to-speech packages, or gTTS, translate spoken inquiries into text. gTTS converts the answer from the look-up function you write to get the answer to the query or command into an audio format. This package provides an API interface for Google Translate.

Jupyter Notebook: Our models were developed and presented using Jupyter Notebook, an interactive computing environment. Because of its adaptability, shared documents with live code, mathematics, and visualizations. Jupyter Notebook promoted communication and guaranteed the accuracy of our research results by enabling an open and repeatable model-building process.

4.2 Proposed System

Using natural language processing (NLP) for communication on any closed domain system, our suggested method can identify speech. It built an online application that leverages the most recent technological advancements. Instead of searching and navigating the website, users may use speech to communicate with the online application. The chatbot may respond to any inquiries about the hospitality industry, the hotel web application (such as where to acquire it or how to use it), or the business domain. The system has been taught to respond to inquiries pertaining to the hotel industry. Any data collection including information about the hotel website is used to train the algorithm. It is designed in a manner that makes it simple to retrain using any collection of data. The following modules are needed for the suggested system: a voice chatbot that may be hosted on a hotel website. Web applications must be able to record voice input and translate it into text. There are many ways to handle this. The two most popular speech to text converters for online applications are listed below.

4.3 Speech Recognition API

Voice to text conversion in online apps is supported via the online Speech API, which was introduced in Chrome version 25 in recent years. It is becoming more and more common and will dominate speech recognition in the future. Through the Speech Recognition interface, the browser makes the speech recognition capability available. This interface may identify speech context from an audio input (usually via the built-in speech recognition service on the device) and react accordingly. To recognise when speech is entered via the device's microphone, we must construct a Speech Recognition object in JavaScript with several event handlers. Webkit Speech Recognition, which is included in the browser window object, may also be used to verify the browser's compatibility. A specific set of grammar that our software must recognise is included in the Speech Grammar interface. JSpeech Grammar Format is used to specify grammar.

4.4 Text to Voice Conversion

It will be necessary to transform the text output that the server produces into voice. Therefore, we need to find a productive technique to turn this text into speech. The consumers will get an engaging sense thanks to the voice response. In text to speech conversion, there are several solutions accessible, such as standalone implementations employing trained models like gTTS or a variety of cloud-hosted APIs. In this case, we examined two options: Web Speech API and gTTS.

4.5 Speech Synthesis API

An interface for the Web Speech library API in browsers is called speech synthesis. Programs may read aloud their text content by using the Speech Synthesis interface, a text-to-speech component that provides access to voice synthesis. We are able to adjust the synthesis voice's pace, pitch, and variety of voice kinds.

4.6 Voice Assistant

”Metro” is the name of the voice assistant that was created. When the user taps the ”Mic” button, voice input will be listened to until the user stops speaking. The text field below will display the voice input converted to text using the Web Recognition API. To start the model-hosted server call, click ”Send.” Take note that the user’s speech transcription was also shown on the screen as text. The audio answer will appear below the text area as indicated once the result is received.

Figure 4.2: Voice Assistant UI screen

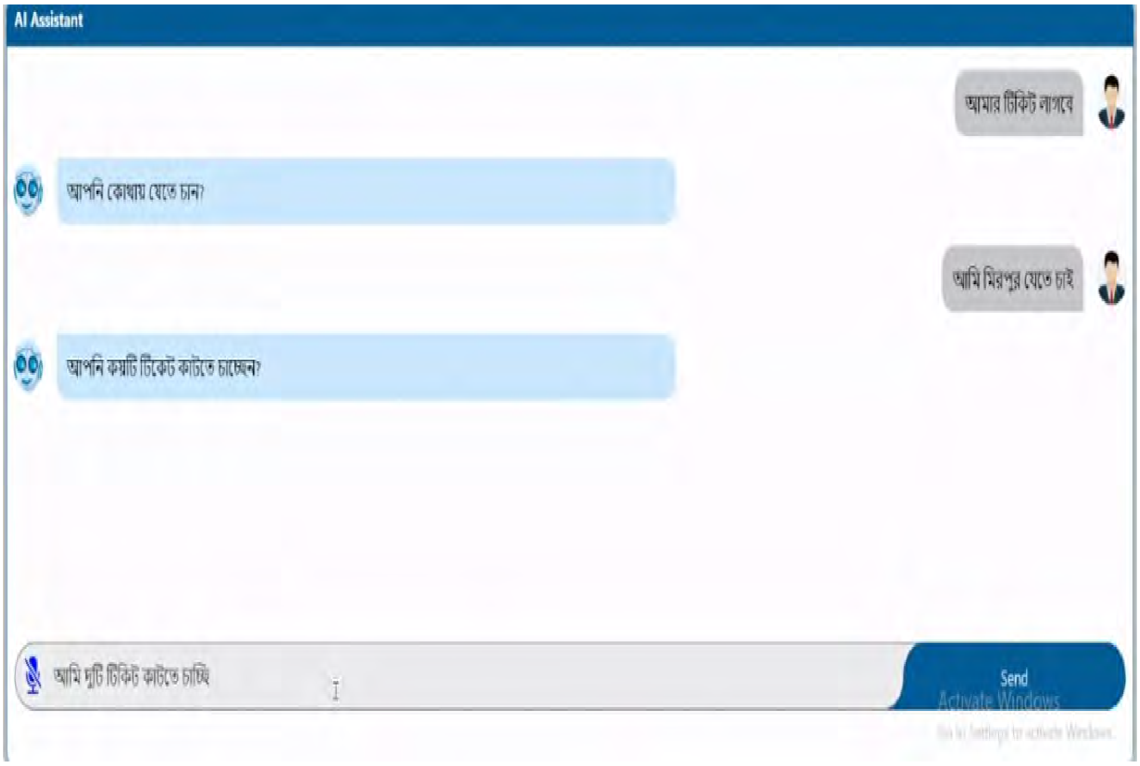


Figure 4.3: The voice and text responses of the Voice Assistant

4.7 Methodology

The process starts with gathering a wide dataset of metro rail queries that captures speaker accents, background noise fluctuations, and real-world settings. After that, a strict preprocessing pipeline is implemented, which includes tokenization, spell-checking, and noise reduction. TF-IDF is used in feature extraction to convert the preprocessed text into numerical representations that serve as the foundation for further analysis. These TF-IDF vectors are then used to train the sequential model, which identifies sequential dependencies in the data. Performance criteria such as accuracy and precision are used to evaluate the model. Robustness is guaranteed by cross-validation, responsiveness is assessed by simulated user interactions, and flexibility is improved by a feedback mechanism. The models are validated by statistical analysis, which also offers insights into how well they work for voice recognition in the dynamic metro rail environment. The process ends with the setup of a chatbot model.

The procedure is extended to include user-initiated speech recordings relating to metro rail information in order to integrate consumer interactions particular to metro rail requests. By pressing the microphone button and starting the recording process, customers start the conversation. The chatbot uses a speech-to-text conversion technology utilizing NLP to record spoken utterances and turn them into text.

After transcribing, the voice assistant uses similarity search techniques to locate pertinent data in the metro rail domain and then provides a suitable response. In the event that the voice assistant is unable to understand the customer's question, it will actively initiate an explanation conversation with them. This entails asking the user to clarify or add more information to their query until the assistant correctly determines that it is about metro rail. This iterative procedure is intended to guarantee efficient communication—especially for issues pertaining to metro rail—between the user and the voice assistant. The voice assistant improves the system's overall functionality by enabling a dynamic interaction model and giving users accurate and pertinent information about metro rail services.

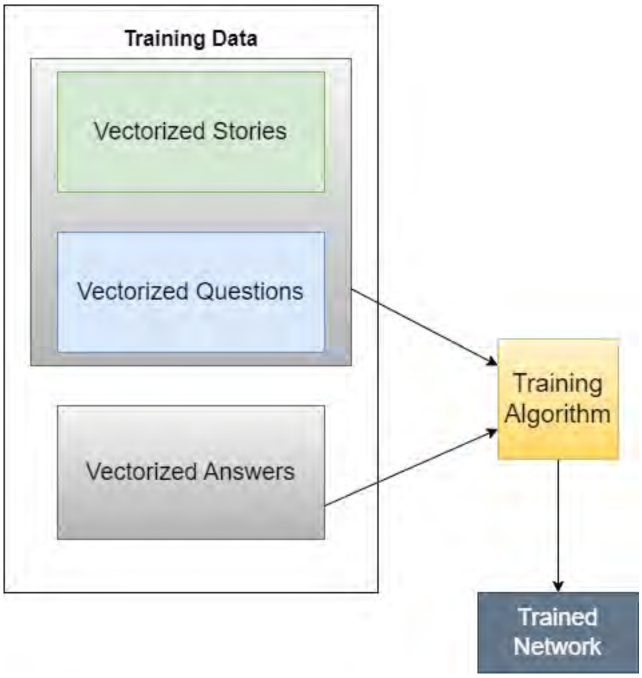


Figure 4.4: Training Algorithm

4.8 Model Specification

This thesis was implemented through a number of phases to investigate different NLP approaches.
At first, TF-IDF, or Term Frequency-Inverse Document Frequency, is employed. This numerical metric is used in information retrieval and natural language processing to evaluate the importance of terms in a text in comparison to a document corpus. By integrating Term Frequency (TF) and Inverse Document Frequency (IDF), it focuses on analyzing word frequency and assessing the significance of phrases in a collection of documents.
On the other hand, the Sequential Model, which is built using the Keras library, is crucial to voice recognition for metro train systems. The sequential data processing and analysis capabilities of this paradigm are essential for comprehending the input pieces' order. The model's resistance is increased by additional dropout layers, and complex patterns are identified using a hidden layer with multiple units and ReLU activation. The output layer creates probability distributions for precise classification in metro rail voice recognition using softmax activation. This model is specifically designed to represent sequential dependencies and successfully tackle the difficulties caused by the different acoustic environments found in metro train cars.

4.8.1 TF-IDF

In natural language processing and information retrieval, TF-IDF (Term Frequency-Inverse Text Frequency) is a numerical statistic that is used to evaluate a term's significance in a text against a corpus of documents. Text analysis, document similarity, and information retrieval are three popular applications for the TF-IDF model.
1. It analyses the frequency with which a word or phrase occurs in a text. The formula for calculating it is to divide the total number of phrases in a text by the number of times a particular phrase occurs.

$$\mathrm{TF}(t, d) = \frac{\text{Number of times term } t \text{ appears in document } d}{\text{Total number of terms in document } d} \tag{4.1}$$

2. It evaluates the importance of a phrase in a collection of documents. It is computed using smoothing to prevent division by zero and the logarithm of the total number of documents divided by the number of documents containing the phrase.

$$\mathrm{IDF}(t, D) = \log\left(\frac{\text{Total number of documents in the corpus } N}{\text{Number of documents containing term } t + 1}\right) + 1 \tag{4.2}$$

3. It is a combination of IDF and TF. It highlights phrases that are exclusive to a certain document by giving particular terms that appear often in that text a high weight even when they are rare across the corpus.

$$\text{TF-IDF}(t, d, D) = \mathrm{TF}(t, d) \times \mathrm{IDF}(t, D) \tag{4.3}$$

4.8.2 Sequential Neural Network Model

The sequential model sets the stage for understanding its fundamental role in machine learning. Sequential models represent a class of neural networks designed to process and analyze sequential data, where the order of input elements is significant. These models have found widespread application in various domains, including natural language processing, time-series analysis, and speech recognition, due to their capacity to capture temporal dependencies within the data. As opposed to traditional feedforward neural networks, sequential models introduce a dynamic element by incorporating recurrent connections, enabling them to consider the context of previous inputs.
The implementation of a sequential model assumes particular significance in this research on the concept of speech recognition for metro rail systems. When handling the temporal dynamics present in spoken language, the Sequential Model—which is implemented using the Keras library—becomes an invaluable instrument for processing sequential data. The model's construction is specifically designed to tackle the special difficulties brought about by the acoustic environments found in metro train compartments.
Rectified Linear Unit (ReLU) activation function and 128 units make up the Dense layer design of the input layer, which is the first layer. This layer is designed to accommodate the depiction of a bag of words, which is especially vital for capturing the crucial elements incorporated into the sequential structure of speech. By adding non-linearity, the ReLU activation function enables the model to see and understand intricate correlations in the input data.

A dropout layer with a dropout rate of 0.5 comes after the input layer. When it comes to voice recognition jobs, overfitting is a prevalent risk that this layer is vital in preventing. During training, 50 percent of the connections between neurons are randomly dropped, which encourages the model to generalize the patterns it has learnt and ensures that it can adapt to a variety of auditory situations found in metro rail settings.

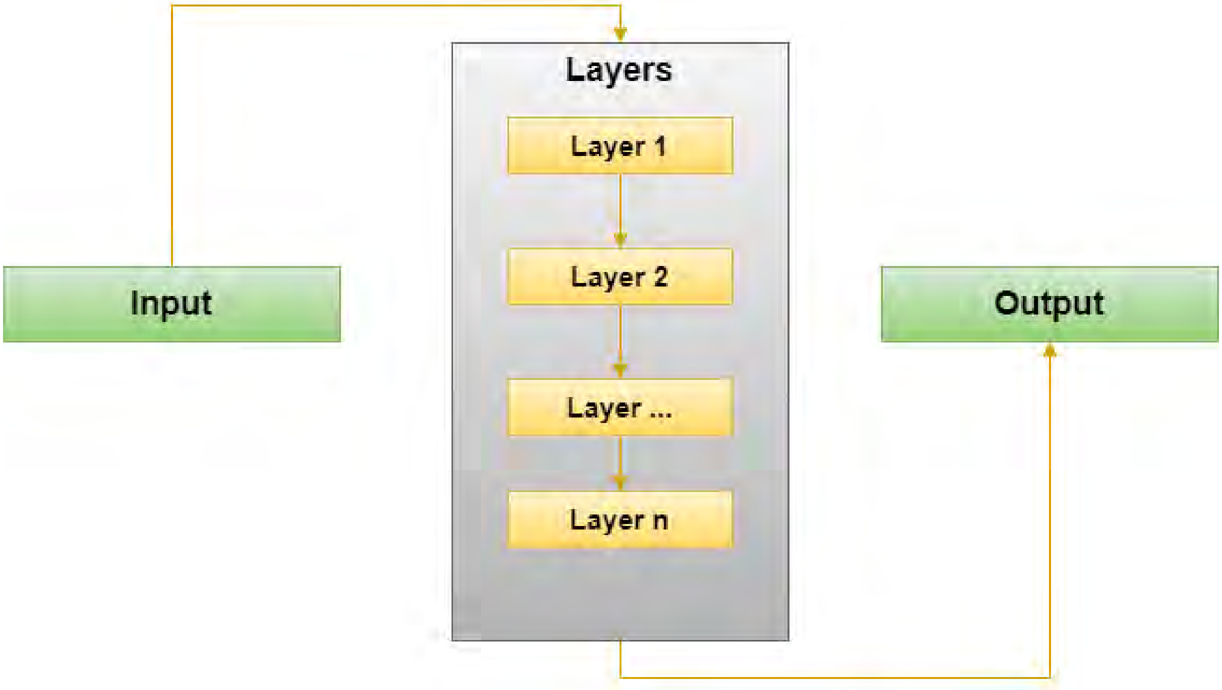


Figure 4.5: Sequential Model Flow

The next Hidden layer is positioned to identify more complex patterns in the data. It is a Dense layer with 64 units and a ReLU activation function. This layer is crucial for speech recognition because it can identify subtle aspects that are incorporated into spoken words. This helps the model understand the intricacy of speech signals that are influenced by the unique acoustics of metro rail stations.
Another Dropout layer with a dropout rate of 0.5 is added to further combat overfitting and improve the model's resilience. The model's ability to generalize is strengthened by the redundancy in dropout layers, which keeps it from depending too much on certain patterns seen in the training set that would not translate well to speech circumstances in metro train compartments in the real world.
Using the softmax activation function, the output layer—the last layer—takes the shape of a dense layer with as many units as there are different classes. This setup makes it easier to generate probability distributions across the classes, which helps the model perform well when classifying several classes. This layer serves as the finishing touch in the context of metro rail speech recognition. It generates predictions based on the sequential patterns that have been learned, which eventually helps to improve accuracy and reliability in classifying spoken commands or queries within the particular acoustic conditions of metro rail environments.

Chapter 5

Result Analysis

5.1 Performance Evaluation Metrics

The proposed voice assistant model for metro rail services has become effective throughout the training process and has been proven throughout the experiment on the CDQA(closed domain question answer) dataset. It's efficiency throughout the training phase from TF-IDF and Sequential model, the sequential model has come out with the better result with a larger number of accuracy percentage.

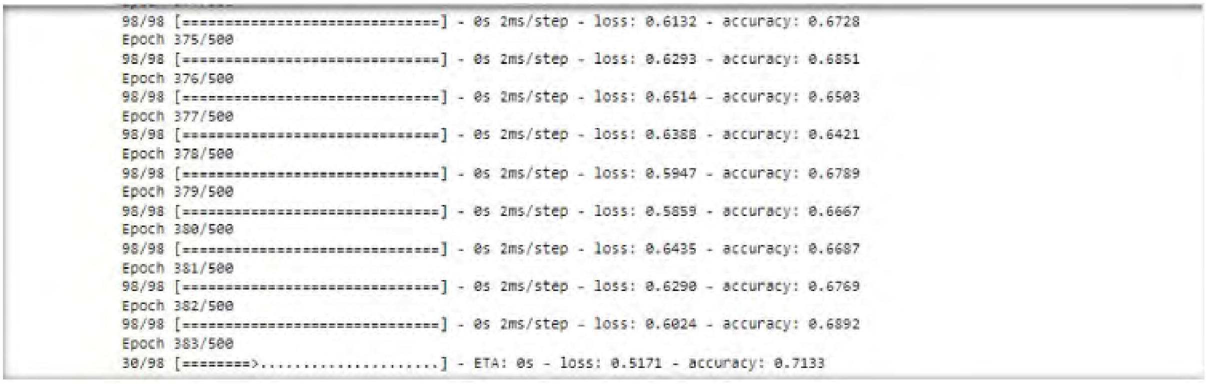
Accuracy: The accuracy graph displays the percentage of all the samples that are identified correctly according to the shaped dataset and categories.
Training and Loss Curves: These are the training and loss curves values generated for all the attributes from expecting the correct answers. Here the probabilities can be shown that for a particular sentence if some words are changed or removed the probability then changes accordingly.

5.2 Experimental Result Analysis

The research's comparative investigation of TF-IDF and sequential models in the context of metro rail voice recognition indicates an interesting tendency. Although TF-IDF has been traditionally effective at capturing document-term relationships, it seems to perform weakly in the particular job of voice recognition on metro trains. This observation highlights the distinct issues that arise from the dynamic and time-sensitive nature of announcements aboard metro trains.

Epochs	Accuracy in percentage
489/500	0.6327
491/500	0.65529
493/500	0.6127
495/500	0.68209
497/500	0.6769
499/500	0.6892
500/500	0.7133

Table 5.1: Model Training Accuracy Tabular Format



```
98/98 [==============================] - 0s 2ms/step - loss: 0.6132 - accuracy: 0.6728
Epoch 375/500
98/98 [==============================] - 0s 2ms/step - loss: 0.6293 - accuracy: 0.6851
Epoch 376/500
98/98 [==============================] - 0s 2ms/step - loss: 0.6514 - accuracy: 0.6503
Epoch 377/500
98/98 [==============================] - 0s 2ms/step - loss: 0.6388 - accuracy: 0.6421
Epoch 378/500
98/98 [==============================] - 0s 2ms/step - loss: 0.5947 - accuracy: 0.6789
Epoch 379/500
98/98 [==============================] - 0s 2ms/step - loss: 0.5859 - accuracy: 0.6667
Epoch 380/500
98/98 [==============================] - 0s 2ms/step - loss: 0.6435 - accuracy: 0.6687
Epoch 381/500
98/98 [==============================] - 0s 2ms/step - loss: 0.6290 - accuracy: 0.6769
Epoch 382/500
98/98 [==============================] - 0s 2ms/step - loss: 0.6024 - accuracy: 0.6892
Epoch 383/500
30/98 [========>.....................] - ETA: 0s - loss: 0.5171 - accuracy: 0.7133
```

Figure 5.1: Model Training Accuracy

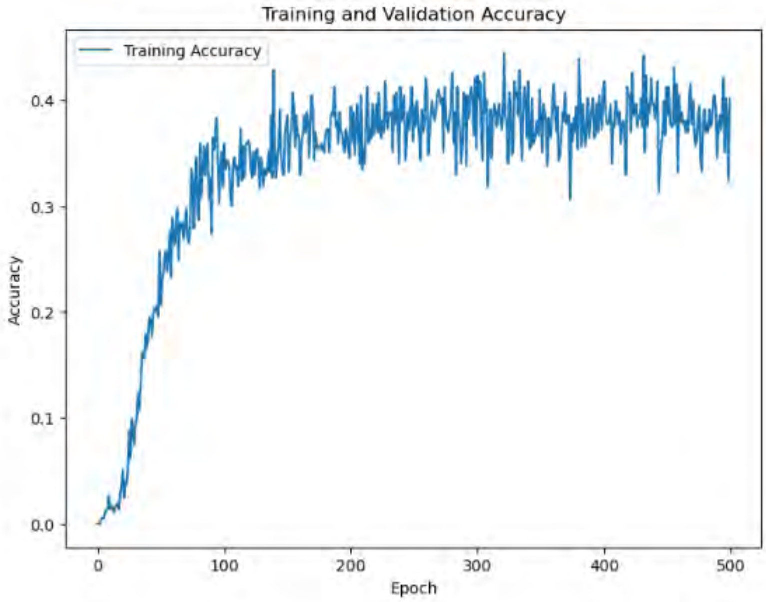


Figure 5.2: Training and Validation Accuracy in Sequential Model

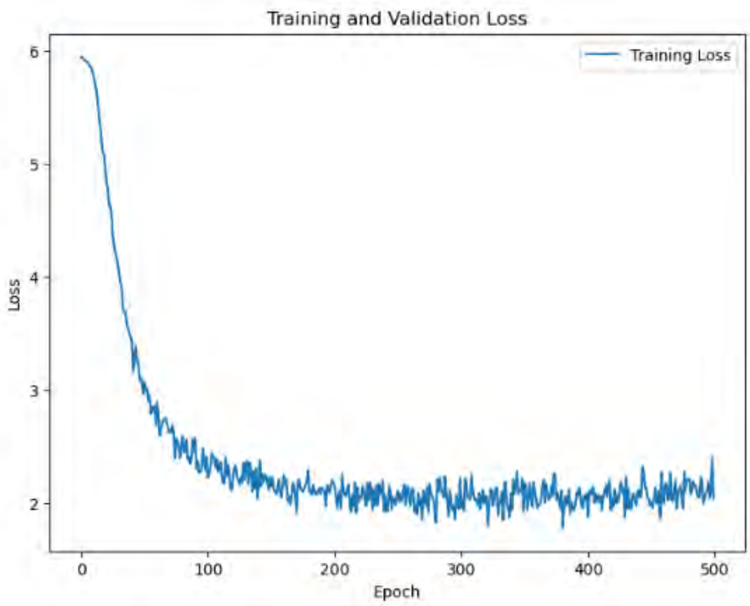


Figure 5.3: Training and validation loss in Sequential Model

On the other hand, sequential models have shown remarkable performance in this field. In identifying speech patterns linked to metro rail announcements, these models perform better than TF-IDF by utilizing their capacity to identify and exploit sequential dependencies in data. Spoken announcements' complex temporal linkages, which are essential for providing accurate and timely information in the metro rail setting, complement their natural advantages.

```
Go! Bot is running!
মেট্রো রেলের দ্বিতীয় ধাপে কত কিলোমিটার রুট যুক্ত হতে যাচ্ছে?
1/1 [==============================] - 0s 228ms/step
[{'intent': 'tag_101', 'probability': '1.0'}]
প্রায় ৪৮ কিলোমিটার নতুন রুট যুক্ত করার লক্ষ্য রয়েছে মেট্রোর ২য় ধাপে।
মেট্রো রেলের দ্বিতীয় ধাপে কত কিলোমিটার যুক্ত হতে যাচ্ছে?
1/1 [==============================] - 0s 37ms/step
[{'intent': 'tag_101', 'probability': '1.0'}]
প্রায় ৪৮ কিলোমিটার নতুন রুট যুক্ত করার লক্ষ্য রয়েছে মেট্রোর ২য় ধাপে।
মেট্রো রেলের দ্বিতীয় ধাপে কত কিলোমিটার রুট যুক্ত?
1/1 [==============================] - 0s 36ms/step
[{'intent': 'tag_101', 'probability': '0.9999918'}]
প্রায় ৪৮ কিলোমিটার নতুন রুট যুক্ত করার লক্ষ্য রয়েছে মেট্রোর ২য় ধাপে।
```

Figure 5.4: Manual Testing of Sequential Model Data

In addition to the quantitative evaluations, this research also carried out manual testing to confirm the sequential model's performance in metro rail speech recognition. This required a careful analysis of the model's sensitivity to changes in sentence length, an important consideration considering the real-world diversity in announcement lengths in metro rail scenarios. Through a methodical manipulation of sentence length and an assessment of the corresponding probability outputs, it is possible to obtain important understandings regarding the model's flexibility in various speech circumstances.

The results show that the sequential model performs consistently and reliably in phrases of different durations. This resilience is especially important in metro train settings where announcements might be anything from brief instructions to detailed information sharing. The accuracy with which the model can determine the likelihood of a sentence, highlights its adaptability and fit for the dynamic nature of metro rail voice data, independent of its length.

In addition to strengthening the quantitative results, this manual testing raises the degree of confidence in the sequential model's applicability. These combined quantitative and qualitative evaluations help to navigate the complexities of metro rail speech recognition and provide a thorough grasp of the model's capabilities, leading to a more methodical and trustworthy approach to information extraction from spoken announcements in public transportation settings.

5.3 Manual Testing of Models

We performed manual testing of our voice assistant system by asking our voice assistant questions directly in both the TF-IDF and Sequential Neural Network models. We avoided formal procedures to provide a genuine testing environment for the TF-IDF model by interacting with a wide range of real-world user inquiries and circumstances. This made it possible for us to assess the accuracy and responsiveness of the model in a user-centered setting. The efficacy of the TF-IDF model in providing users with relevant information while navigating the intricacies of metro rail services was thoroughly evaluated via the use of user input and subjective judgments.
In the same way, we replicated real user interactions by asking the voice assistant questions directly during our manual testing of the Sequential Neural Network model. This unstructured approach aimed to emulate the variety and spontaneity seen in real-world user chats. To evaluate the flexibility and efficacy of the Neural Network, we examined its answers to a range of language subtleties and user inquiries. A more natural assessment of the model's performance in a situation that was rather similar to real-world use was made possible by the lack of rigorous procedures during manual testing. The results showed that the Sequential Neural Network model (71.6%) outperformed the TF-IDF model (62.4%) in terms of accuracy. Adjusting hyperparameters such as layer configurations and dropout rates resulted in improved accuracy in the Neural Network model. Qualitative evaluations confirmed its superiority in real-world circumstances by highlighting its ability to handle complex questions and offer contextually appropriate solutions.

Model	Accuracy in percentage
Sequential NN model	71.6%
TF-IDF	62.4%

Table 5.2: Manual Testing Accuracy in Tabular Forma

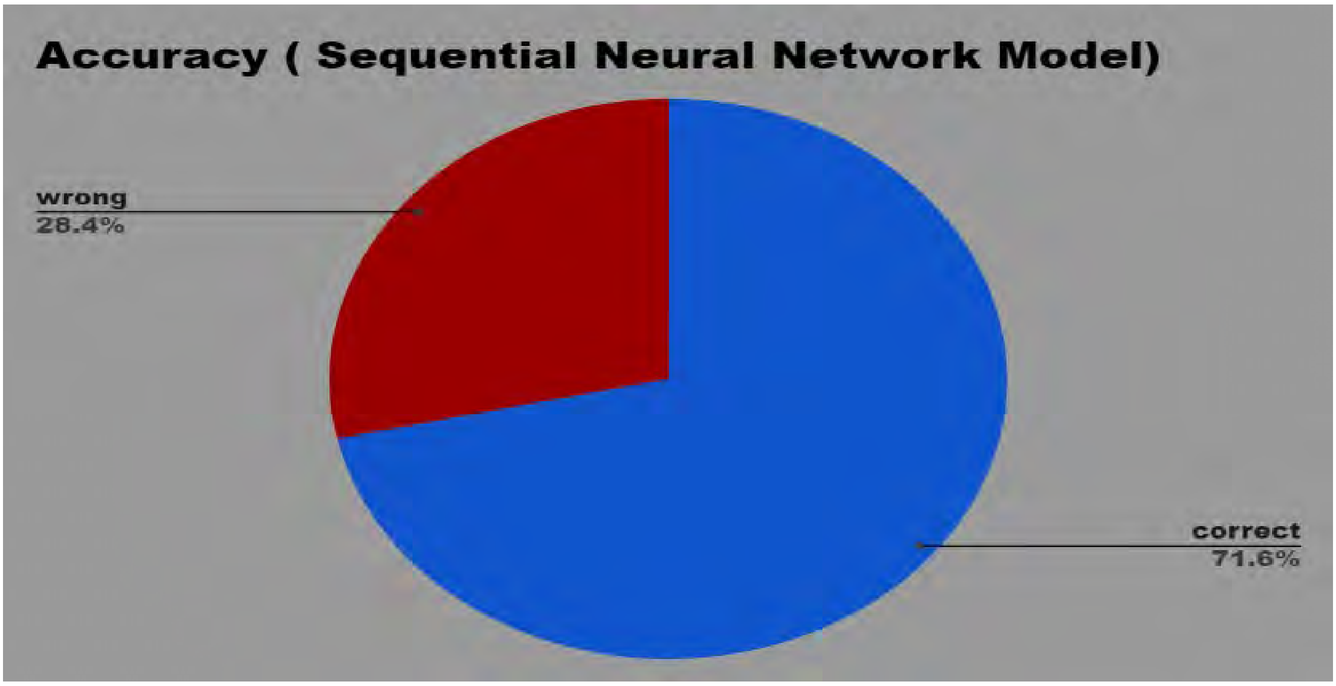


Figure 5.5: Manual Testing Accuracy of Sequential Model

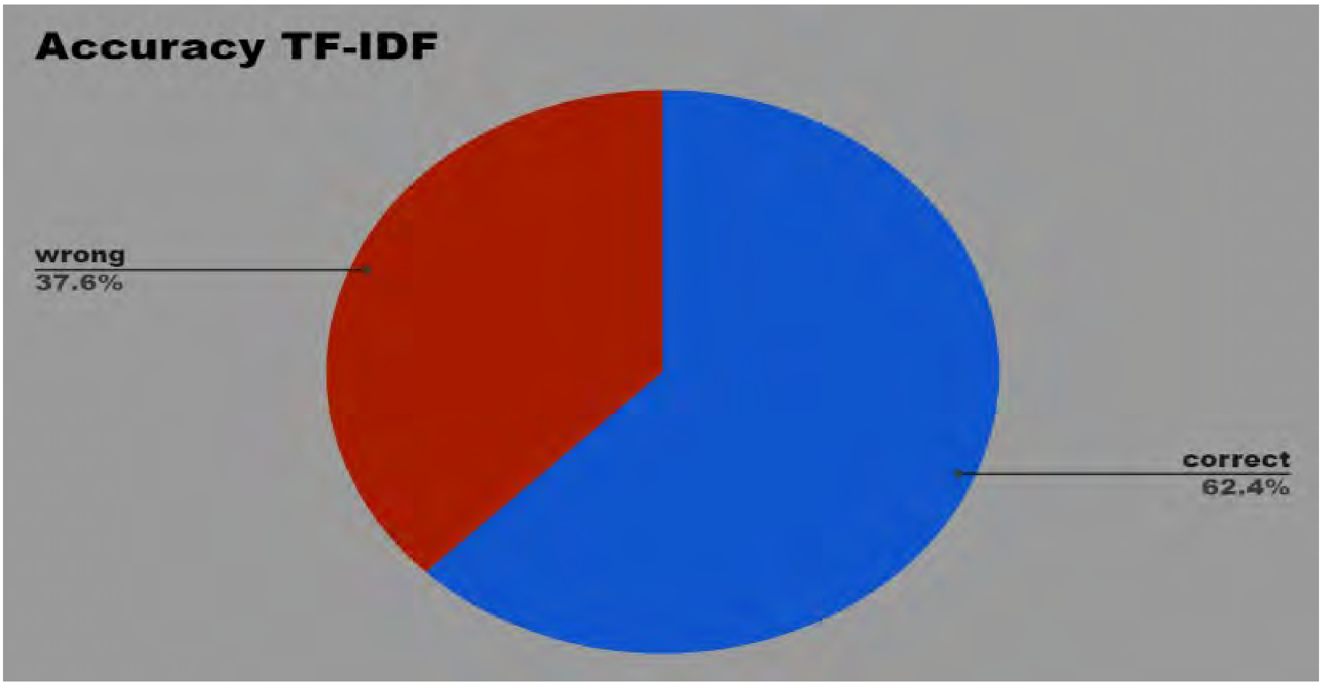


Figure 5.6: Manual Testing Accuracy of TF-IDF Model

Chapter 6

Conclusion

The proposed voice assistant improves the entire user experience by resolving the difficulties commuters have, especially when it comes to getting tickets and navigating metro stations. The rigorous collection of relevant information guarantees that the voice assistant is adequately prepared to address inquiries from more people. The voice assistant's ability to comprehend and reply to user inquiries is made more reliable and effective by the algorithms. This ensures that users of the metro-rail system may simply communicate with it using natural language, even if they are not used to using standard ticket vending machines. It also looks into the possibility of making voice assistants available to the general public, which would make the transportation system more accessible and user-friendly. This shows that by lowering dependency on traditional interfaces and enabling dynamic interactions via natural language, voice assistance technologies may greatly improve the user experience in metro train services. The results open the door to more research and the use of comparable solutions in global public transport networks, promoting technical developments advantageous to both transit organizations and passengers.
The recommended system's capabilities also enable the study of commuter data, improving services and fostering thoughtful decision-making. By implementing this approach successfully, metro rail services may anticipate higher commuter satisfaction and use. The results of this study and the created prototype show the viability and potential benefits of this strategy, indicating the possibility of its application to other metro rail systems across the world. Overall, integrating an AI model with a voice assistance system that uses NLP has the potential to revolutionize metro rail services, improve the customer experience overall, and increase operational effectiveness.

6.1 Challenges

The challenges encountered in various research areas demonstrate how sincerely this work is committed to offering a dependable and culturally relevant solution to embark on the innovative journey of enhancing metro rail services through the integration of NLP technologies. One of the primary challenges the research faces is creating a specific dataset for the Bengali language, which has its own linguistic variations and cultural traits. Unlike widely spoken languages, a comprehensive procedure of data collection, annotation, and preparation was necessary due to the absence of a pre-made dataset. This custom dataset creation process presented many challenges, including the need to convey cultural context, the lack of pre-existing resources, and the wide range of language variations. The difficulties of representing regional languages, including cultural variances, and translating and standardizing Bengali are only a few of the issues covered in this section on the challenges of building a dataset in the language. In addition to the comprehensive approach to overcoming contextual, linguistic, and linguistic barriers with the goal of a successful implementation, the careful consideration of these issues demonstrates the commitment to ensuring the effectiveness and acceptance of the NLP model for Bengali metro rail services. Although this research is an innovative investigation project on metro train tickets, it was challenging to locate relevant data since there hasn't been any prior research on the usage of speech recognition technology in this context. As this was a new area, much caution had to be taken in gathering relevant information. Since there is no established context, the purpose of this research is to highlight novel aspects of metro rail systems. In order to navigate this unexplored area, a thorough evaluation of research covering ticketing systems, transportation technologies, and advancements in the field as a whole was required. This study is influenced by perspectives from adjacent subjects and domains. To navigate the difficulties involved with innovative research and provide new insights and solutions to the metro rail industry, a flexible and inventive technique is needed.

6.2 Limitations

The lack of an existing pre-built dataset necessitated the development of a custom dataset for the concerned system. The dataset is significantly condensed and limited due to the Bengali language environment, which affects the overall accuracy of the system.

In the Bengali language, there are many different ways that questions might appear. But the dataset isn't diverse enough in the questions it asks, which makes it less accurate than it might be. Still, it is expected that this problem may be resolved with more time for training and dataset tuning.

The non-conversational character itself is a considerable barrier. Because there are few datasets accessible, the system can't support long-form conversations when it operates just in Bengali. Although a more comprehensive dataset might have been used to construct an open-domain question-answering approach, time restrictions and language hurdles made this impractical within the period allotted.

The many accents found in the Bengali language provide other challenges since speakers from various areas tend to pronounce words differently. At the moment, the technology only supports mainstream Bengali and has trouble understanding accents from rural areas. However, careful training and enough time invested might potentially lessen this constraint and improve the system's capacity to adjust to various language subtleties.

6.3 Future Work

There are certain limits to our present system, but as the dataset expands in the future, there is a great chance that a smooth and error-free discussion will take place. As this system develops, one may expect it to be incorporated into tangible objects like vending machines, completely changing the way customers purchase tickets in person. In addition, the system may be enhanced with other functionalities, such as the ability to process payments and transactions via online payment gateways or QR code readers.

Expanding the dataset has the potential to improve open-domain question answering and facilitate more seamless user-system interactions. This enhanced conversational capacity has the potential to greatly improve the user experience in general. In the future, a progressive idea would be to integrate this cutting-edge technology into a robot, providing users with unmatched accessibility and an improved ticketing experience. Imagine a robotic presence moving between stations, helping people, and making the environment more effective and user-friendly.

Bibliography

[1] J. L. Leopold **and** A. L. Ambler, **?**Keyboardless visual programming using voice, handwriting, and gesture**?**, **in** *Proceedings. 1997 IEEE Symposium on Visual Languages (Cat. No. 97TB100180)*, IEEE, 1997, **pages** 28–35.

[2] S. C. Arnold, L. Mark **and** J. Goldthwaite, **?**Programming by voice, Vocal-Programming**?**, **in** *Proceedings of the fourth international ACM conference on Assistive technologies*, 2000, **pages** 149–155.

[3] A. Begel **and** S. L. Graham, **?**Spoken programs**?**, **in** *2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)*, IEEE, 2005, **pages** 99–106.

[4] J.-H. Chiang **and** H.-C. Yu, **?**Literature extraction of protein functions using sentence pattern mining**?**, *IEEE Transactions on Knowledge and Data Engineering*, **jourvol** 17, **number** 8, **pages** 1088–1098, 2005.

[5] E. W. Dijkstra, **?**On the foolishness of "natural language programming"**?**, *Program Construction: International Summer School*, **pages** 51–53, 2005.

[6] Y. Li, D. McLean, Z. A. Bandar, J. D. O'shea **and** K. Crockett, **?**Sentence similarity based on semantic nets and corpus statistics**?**, *IEEE transactions on knowledge and data engineering*, **jourvol** 18, **number** 8, **pages** 1138–1150, 2006.

[7] A. A. Yahya **and** A. Osman, **?**Automatic classification of questions into Bloom's cognitive levels using support vector machines**?**, 2011.

[8] S. Banerjee **and** S. Bandyopadhyay, **?**Bengali question classification: Towards developing qa system**?**, **in** *Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing*, 2012, **pages** 25–40.

[9] A. Oludele **andothers**, **?**Information Retrieval: An Overview.**?**, *International Journal of Advanced Research in Computer Science*, **jourvol** 3, **number** 5, 2012.

[10] A. A. Yahya, Z. Toukal **and** A. Osman, **?**Bloom's taxonomy–based classification for item bank questions using support vector machines**?**, **in** *Modern advances in intelligent systems and tools*, Springer, 2012, **pages** 135–140.

[11] S. Banerjee, S. K. Naskar **and** S. Bandyopadhyay, **?**Bfqa: A bengali factoid question answering system**?**, **in** *Text, Speech and Dialogue: 17th International Conference, TSD 2014, Brno, Czech Republic, September 8-12, 2014. Proceedings 17*, Springer, 2014, **pages** 217–224.

[12] P. Biswas, A. Sharan **and** R. Kumar, ?Question Classification using syntactic and rule based approach?, **in** *2014 international conference on advances in computing, communications and informatics (ICACCI)*, IEEE, 2014, **pages** 1033–1038.

[13] P. Daniels, ?Using Web Speech technology with language learning applications.?, 2, **volume** 11, ERIC, **august** 2015, **pages** 177–187.

[14] J. Hirschberg **and** C. D. Manning, ?Advances in natural language processing?, *Science*, **jourvol** 349, **number** 6245, **pages** 261–266, 2015.

[15] ——, ?Advances in natural language processing?, *Science*, **jourvol** 349, **number** 6245, **pages** 261–266, 2015.

[16] S. Hoque, M. S. Arefin **and** M. M. Hoque, ?BQAS: A bilingual question answering system?, **in** *2015 2nd international conference on electrical information and communication technologies (EICT)*, IEEE, 2015, **pages** 586–591.

[17] B. Behera, ?Chappie-a semi-automatic intelligent chatbot?, *cse. iitb. ac. in*, 2016.

[18] N. Van-Tu **and** L. Anh-Cuong, ?Improving question classification by feature extraction and selection?, *Indian Journal of Science and Technology*, **jourvol** 9, **number** 17, **pages** 1–8, 2016.

[19] A. B. Keerthana S Meghana H, ?Smart Home Using Internet of Things?, *Perspectives in Communication, Embedded -systems and signal processing*, 2017.

[20] S. M. H. Nirob, M. K. Nayeem **and** M. S. Islam, ?Question classification using support vector machine with hybrid feature extraction method?, **in** *2017 20th international conference of computer and information technology (ICCIT)*, IEEE, 2017, **pages** 1–6.

[21] K. Osadi, M. Fernando, W. Welgama **andothers**, ?Ensemble classifier based approach for classification of examination questions into Bloom's taxonomy cognitive levels?, *International Journal of Computer Applications*, **jourvol** 162, **number** 4, **pages** 1–6, 2017.

[22] E. S. Othman, ?Voice Controlled Personal Assistant Using Raspberry P?, *International Journal of Scientific and Engineering Research*, **jourvol** 8, **november** 2017.

[23] A. Sinha, G. Garg, G. Rajwani **and** S. Tayal, ?Intelligent Personal Assistant?, *International Journal of Informative & Futuristic Research, Volune*, **jourvol** 4, 2017.

[24] Y. Cui, T. Liu, W. Che, L. Xiao, Z. Chen, W. Ma, S. Wang **and** G. Hu, ?A span-extraction dataset for Chinese machine reading comprehension?, *arXiv preprint arXiv:1810.07366*, 2018.

[25] J. Devlin, M.-W. Chang, K. Lee **and** K. Toutanova, ?Bert: Pre-training of deep bidirectional transformers for language understanding?, *arXiv preprint arXiv:1810.04805*, 2018.

[26] V. Kepuska **and** G. Bohouta, ?Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home)?, **in** *2018 IEEE 8th annual computing and communication workshop and conference (CCWC)*, IEEE, 2018, **pages** 99–103.

[27] K. Lee, K. Yoon, S. Park **and** S.-w. Hwang, ?Semi-supervised training data generation for multilingual question answering?, **in** *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, 2018.

[28] M. Mohammed **and** N. Omar, ?Question classification based on bloom's taxonomy using enhanced tf-idf?, *Int J Adv Sci Eng Inf Technol*, **jourvol** 8, **pages** 1679–1685, 2018.

[29] P. Rajpurkar, R. Jia **and** P. Liang, ?Know what you don't know: Unanswerable questions for SQuAD?, *arXiv preprint arXiv:1806.03822*, 2018.

[30] T. Young, D. Hazarika, S. Poria **and** E. Cambria, ?Recent trends in deep learning based natural language processing?, *ieee Computational intelligenCe magazine*, **jourvol** 13, **number** 3, **pages** 55–75, 2018.

[31] D. Gupta, A. Ekbal **and** P. Bhattacharyya, ?A deep neural network framework for English Hindi question answering?, *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, **jourvol** 19, **number** 2, **pages** 1–22, 2019.

[32] S. T. Islam **and** M. N. Huda, ?Design and development of question answering system in Bangla language from multiple documents?, **in** *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, IEEE, 2019, **pages** 1–4.

[33] H. Mozannar, K. E. Hajal, E. Maamary **and** H. Hajj, ?Neural Arabic question answering?, *arXiv preprint arXiv:1906.05394*, 2019.

[34] M. d'Hoffschmidt, W. Belblidia, T. Brendlé, Q. Heinrich **and** M. Vidal, ?FQuAD: French question answering dataset?, *arXiv preprint arXiv:2002.06071*, 2020.

[35] P. Efimov, A. Chertok, L. Boytsov **and** P. Braslavski, ?Sberquad–russian reading comprehension dataset: Description and analysis?, **in** *Experimental IR Meets Multilinguality, Multimodality, and Interaction: 11th International Conference of the CLEF Association, CLEF 2020, Thessaloniki, Greece, September 22–25, 2020, Proceedings 11*, Springer, 2020, **pages** 3–15.

[36] L. L. Wang, K. Yao, X. Q. **and** Y., ?A smartphone-based bus arrival prediction method with GPS data.?, *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, **jourvol** 24(1), **pages** 51–64, 2020.

[37] T. Desk, ?A brief look at metro rails around the world?, *Dhaka Tribune*, **december** 2022. **url**: https://www.dhakatribune.com/around-the-web/301426/a-brief-look-at-metro-rails-around-the-world.

[38] L. Kumar, ?VOICE ASSISTANT USING NATURAL LANGUAGE PROCESSING (NLP)?,

[39] R. Poonguzhali **and** D. LAKSHMI, ?Analysis on the Performance of Some Standard Deep Learning Network Models for Question Answering Task?, *Networks*, **jourvol** 7, **number** 14, **page** 2020,