Healthcare Chatbot

A major project report submitted in partial fulfilment of the requirement for the award of degree of

Bachelor of Technology

in

Computer Science & Engineering / Information Technology

Submitted by

Amisha Chauhan (201117)

Naman Puri (201318)

Under the guidance & supervision of

Dr. Aman Sharma

Asst. Prof. (SG)

Department of Computer Science & Engineering and Information Technology

Jaypee University of Information Technology, Waknaghat, Solan - 173234 (India)

CERTIFICATE

This is to certify that the work which is being presented in the project report titled “Healthcare Chatbot using NLP Techniques” in partial fulfilment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering and submitted to the Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat is an authentic record of work carried out by “Amisha Chauhan 201117” and “Naman Puri 201318” during the period from August 2023 to May 2024 under the supervision of Dr. Aman Sharma, Department of Computer Science and Engineering, Jaypee University of Information Technology, Waknaghat.

Amisha Chauhan
Roll No.: 201117

Naman Puri
Roll No.: 201318

The above statement made is correct to the best of my knowledge.

Candidate's Declaration

We hereby declare that the work presented in this report entitled **'Healthcare Chatbot'** in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** in **Computer Science & Engineering / Information Technology** submitted in the Department of Computer Science & Engineering and Information Technology**,** Jaypee University of Information Technology, Waknaghat is an authentic record of my own work carried out over a period from August 2023 to May 2024 under the supervision of **Dr. Aman Sharma** (Asst. Prof. (SG), Department of Computer Science & Engineering and Information Technology).

The matter embodied in the report has not been submitted for the award of any other degree or diploma.

(Student Signature with Date)
Student Name: Amisha Chauhan Roll No.: 201117

(Student Signature with Date)
Student Name: Naman Puri
Roll No.: 201318

This is to certify that the above statement made by the candidate is true to the best of my knowledge.

(Supervisor Signature with Date)
Supervisor Name: Dr. Aman Sharma
Designation: Asst. Prof. (SG)
Department: Computer Science & Engineering and Information Technology Dated:

ACKNOWLEDGEMENT

Firstly, we express our heartiest thanks and gratefulness to almighty God for His divine blessing making it possible to complete the project work successfully.

We are grateful and wish my profound indebtedness to Supervisor **Dr. Aman Sharma, Assistant Professor**, Department of CSE Jaypee University of Information Technology, Waknaghat. Deep Knowledge & keen interest of my supervisor in the field of "**Cloud Computing**" to carry out this project. His proper guidance, encouragement, constant supervision, constructive criticism, valuable advice, reviewing our work and correcting them at all stages have made it possible for partial completion of this project.

We would also generously thank each one of those individuals who have helped us straightforwardly or in a roundabout way in making this project a win. In this unique situation, We might want to thank the various staff individuals, both educating and not instructing, which have developed their convenient help and facilitated our undertaking.

Amisha Chauhan

201117

Computer Science & Engineering and Information Technology

Jaypee University of Information Technology, Waknaghat

Naman Puri

201318

Computer Science & Engineering and Information Technology

Jaypee University of Information Technology, Waknaghat

TABLE OF CONTENT

LIST OF FIGURES

ABSTRACT

Healthcare is rapidly moving toward patient-centred care, and technology has been essential in developing the quality of personalised healthcare. This is a report about an artefact, which is a healthcare chatbot that utilises state-of-the-art natural language processing technology. Chatbot was developed with an aim of enhancing patient participation, provisioning relevant messages in time, as well building linkage between doctors and patients.

It leverages highly advanced NLP systems that can understand the sentiments of conversations and respond back as a real human would, making communication easy and relevant within its context. It can be integrated in existing healthcare platforms and hence made accessible by patients for queries such as health information to specific medical inquiries. Through artificial intelligence or machine learning algorithms, it is possible for the chatbot to keep improving on its performance adjusting according to the ever-changing language and user's preferences.

The healthcare chatbot's key features include symptom analysis, prescription reminders, and health-related FAQs. By adhering to healthcare standards and using encryption mechanisms for sensitive information, the system assures data security and privacy compliance.

This study adds to the expanding body of work in harnessing natural language processing (NLP) for healthcare applications by providing insights into the creation, deployment, and assessment of a chatbot targeted at enhancing patient experiences and optimising healthcare communication channels. The findings highlight the potential for NLP-based chatbots to play an important role in the future of healthcare by enhancing accessibility, efficiency, and patient-centred treatment.

CHAPTER 1: INTRODUCTION

1.1 Introduction

To fulfil the rising need for accessible and efficient patient services in the quickly expanding healthcare scene, the integration of cutting-edge technology has become critical. Natural Language Processing (NLP) has emerged as a transformational force, revolutionising the way people engage with healthcare providers. This study describes a revolutionary healthcare chatbot system that uses NLP techniques to improve patient engagement and communication in the healthcare area.

The healthcare system is by no means perfect. The Indian healthcare system is faced with several challenges being increased healthcare costs, need for nursing, care for senior citizens, poor not being able to afford healthcare and negligence of public health due to lack of utilisation of adequate funds. [1] Only 37% of the total Indian population is covered by health insurance as of 2022, which indicates that close to a billion are still hanging onto mere hope to afford when an emergency may arrive.

The chatbot, designed to improve user experience, employs powerful NLP models to grasp and create contextually relevant replies, mimicking normal human interaction. This intelligent system is intended to help with a variety of healthcare-related tasks, such as symptom assessments and the transmission of relevant health information.

With the healthcare sector moving towards a paradigm that is more patient-centric, this research attempts to add to the increasing amount of information about the use of NLP in healthcare settings. This study aims to highlight the potential influence of natural language processing (NLP) technologies on healthcare communication channel optimisation and, eventually, patient outcome improvement by offering a thorough overview of the healthcare chatbot's development and deployment.

1.2 Problem Statement

An enormous need for easily accessible and effective patient assistance systems exists in the healthcare industry. Frequently, patients need immediate access to medical information and basic diagnostic advice. Using cutting-edge tools like Scikit-Learn, Pandas, and Pyfiglet, the task is to develop a system that elegantly blends these functions.

The need to create a healthcare chatbot that makes use of these technologies to give individualised healthcare support and provide correct medical information is outlined in this issue statement. This technology may address significant difficulties in the healthcare industry and improve the patient experience through accurate information transmission and personalised replies.

1.3 Objectives

- To study the existing technologies for a chatbot
- To propose framework for a chatbot
- Testing and validation of the proposed framework
- We aim to create a platform which provides textual, verbal and visual facilities to the user for self-diagnostics of health.

1.4 Significance and motivation of the project

This initiative bears relevance as it successfully bridges the gulf between sophisticated technology and healthcare, solving major sectoral concerns. The usage of Natural Language Processing (NLP) in healthcare chatbots is spurred by essential reasons, including greater communication, simplified processes, and improved accessibility to medical information, together contributing to the evolution of healthcare services.

- **Enhanced Patient Engagement:** The project's goal is to encourage patients to take a more educated and proactive part in their treatment. Patients may have meaningful conversations about their health, symptoms, and treatment plans with a chatbot that can comprehend and answer natural language inquiries. This fosters a sense of empowerment and engagement in patients' care.

- **Effective Information Distribution:** By providing prompt answers to user questions, the healthcare chatbot acts as a trustworthy information source. This effective information-dissemination process helps to decrease false information and increase user health literacy in addition to increasing the accessibility of healthcare knowledge.

- **Optimised Healthcare Processes:** The chatbot expedites standard healthcare procedures by providing multiple features. By reducing administrative responsibilities and increasing the efficiency with which healthcare resources are used, this optimisation frees up healthcare practitioners to concentrate on more intricate and important activities.

- **24/7 Support and Accessibility:** The chatbot offers consumers a continuous support system and is accessible around-the-clock. This feature is especially important when it comes to medical situations since it guarantees that people may get advice and information whenever they need it, which might lessen the load on emergency services.

- **Data-Driven Insights:** An abundance of useful data on user interactions, preferences, and healthcare queries is produced by the project. Insights into user behaviour, public health trends, and areas where healthcare services may be further enhanced can be gained from analysing this data, which supports evidence-based decision-making in healthcare management.

- **Technology Advancement in Healthcare:** The initiative is in line with the larger trend of technology improvement in healthcare by utilising NLP methods. It demonstrates how cutting-edge technologies may improve patient care, communication, and the provision of healthcare services.

1.5 Organization of Project Report

Chapter 1: An overview of the project that includes information about its purpose, its goals, and how it will benefit users is provided.

Chapter 2: Survey of the literature for the project. This covers the several project reports on the healthcare-focused chatbot programmes and their uses.

Chapter 3: System Development: In this section, we have covered the project's primary design, the connections between its components, the platforms it was built on, and a summary of the algorithms it employed.

Chapter 4: Model testing is completed, ensuring functionality, accuracy, and performance meet specified requirements before deployment in the production environment.

Chapter 5: The model is executed, generating results that are analysed to confirm the expected output and validate performance in real-world scenarios.

Chapter 6: In conclusion, we have discussed the project's results, its future scope, and what more may be done right now.

CHAPTER 2: LITERATURE REVIEW

2.1 Overview of Relevant Literature

In this section we have looked at the past research and studies around chatbot and its implementation. Variety of such researches have been carried out and their target causes have varied as well. We shall also look at the various tools used by them so a comparison can be drawn between our implementation to that of theirs.

Anuj Patil [2] created a healthcare chatbot using AI. In the process, he used the following techniques: NLP, Google API, Rasa NLU and AWS services. The key point in his research was his use of a large dataset to help with better accuracy.

In another research [3], author Greeshma Kurup created an AI Conversational Chatbot for Primary Healthcare Diagnosis. She used NLP and concepts of deep learning. She employed decision tree classifiers in her methodology. Her flaw was her inconsideration of privacy and security measures for the users of her chatbot.

One noble research was based on the pre-screening for female breast cancer [4], the authors here used a variety of tools to help guide their research such as Chatbot Framework, YAML, Character Based Word Segmentation, Logic Adapters. The SAC's assessment scores were compared to the doctors' scores, and it achieved a high correlation accuracy. FBC doctors evaluated based on scaling coefficient and weight factors. However, users giving too much information might not give back any response due to lack of vocabulary.

Another research instance took the opportunity to work on Natural Language Processing based Human Assistive Health Conversational Agent for Multi-Users [5]. The authors Jim Elliot Christopher James; Mahima Saravanan; Deepa Beeta Thiyam; Prasath Alias Surendhar S used Dialog flow application programming interface (API), which is powered by Google's Natural Language Processing algorithm. The results of the paper demonstrate the development of a

health assistant system using Google's Dialog Flow API, which is a natural language processing algorithm. Users can make health requests and receive relevant health suggestions and recommendations through text messages.

In Health Chatbot: Design, Implementation, Acceptance and Usage Motivation [6], the authors A. Softić, J. Husić, Aida Softić, Sabina Baraković made use of the Chat fuel platform to guide their research. It allows users to interact with the chatbot in a conversational manner, helping them identify their symptoms and decide whether or not to seek medical attention.

A simple yet effective system *Chatbot for Healthcare System Using Artificial Intelligence* [7], The application uses a question and answers protocol where it consists of a login page. Query is available in the database or displays similar answers. An expert answering page where experts answer directly to the user's question. Though, it failed to solve complex queries.

In Implementation of interactive healthcare advisor model using chatbot and visualisation [8], Tae-Ho Hwang; JuHui Lee; Se-Min Hyun; KangYoon Lee made a system in which the user's biological signals could be supervised and forwarded to the chatbot for further observation. Limited disease prediction and diagnosis using medical info. It requires artificial intelligence systems.

Now let us look at the various projects already performed on this specific topic and gather key takeaways from it such as tools used, results established and limitations of that paper.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
1.	An AI-Based Medical Chatbot Model for Infectious Disease Prediction	IEEE Innovations in Intelligent Systems and Applications Conference (ASYU), 2023	Using NLP to interact directly with users and give responses.	User-friendly chatbot providing medical info in various languages, offering doctor details, hospital locations, and disease-related information, facilitating easy access.	Takes more time to give accurate results.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
2.	Smart Chatbot Architecture based NLP and Machine learning for health care assistance.	IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2023	This chatbot mainly uses NLP and NLU for its working. It also uses ASR i.e. Automatic Speech Recognition.	The user's inquiries are correctly understood by this system, which then provides relevant answers from the knowledge base in response.	Only inquiries from a closed domain or questions that are defined in the database are answered by the chatbots.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
3.	AI Based Healthcare Chatbot using Natural Language Processing and Pattern Matching	International Journal of Novel Research and Development (2023)	This study has used the following techniques: NLP, Python.	This application combines a language interpretation module, sound-to-information conversion, and voice recognition software	The tool that enables marketers to develop highly personalised customer interactions, improve business response, and address client issues.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
4.	Implementation of an Educational Chatbot using Rasa Framework	International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN, 2022	The primary technology used for building the chatbot is Rasa. Content-based filtering is used for providing personalised recommendations	Users can choose from a list of courses in subjects like Maths, Science, Computer Science, and Aptitude. Users can take quizzes related to the courses they have studied.	The training data has a major impact on the chatbot's accuracy. Insufficient training data may lead to less accurate interpretation of user intents and entities.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
5.	Natural Language Processing based Human Assistive Health Conversational Agent for Multi-Users	Second International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021	In this study, the health assistant system was developed using Dialog Flow application programming interface (API) which is a Google's Natural language processing powered algorithm.	This chatbot acts like an informative and conversational chatbot. This chatbot provides medical knowledge such as disease symptoms and treatments.	Answering a user's inquiry requires extra time. To use live chat, there are fees.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
6.	Online Textual Symptomatic Assessment Chatbot Based on Q&A Weighted Scoring for Female Breast Cancer Pre-screening	IEEE World Conference on Applied Intelligence and Computing (AIC), 2021	Chatterbot Framework [9], YAML, Character Based Word Segmentation, Logic Adapters	The SAC's assessment scores were compared to the doctors' scores, and it achieved a high correlation accuracy. FBC doctors evaluated on the basis of scaling coefficient and weight factors.	Users giving too much information might not give back any response due to lack of vocabulary.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
7.	Health Chatbot: Design, Implementation, Acceptance and Usage Motivation	20th International Symposium INFOTEH-JAHORINA (INFOTEH), 2021	This paper presents the health chatbot application created on the Chat fuel platform [10].	It allows users to interact with the chatbot in a conversational manner, helping them identify their symptoms and decide whether to seek medical attention.	Using medical data to forecast and diagnose diseases is limited. It necessitates the use of artificial intelligence technologies.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
8.	College Enquiry Chatbot using RASA	International Journal of Scientific Research in Computer Science, Engineering and Information Technology ISSN, 2021	The chatbot is implemented using Rasa. Conditional Random Fields is a discriminative model used for predicting sequences.	Users can find contact information for the college, including the website and WhatsApp number.	If the knowledge base is not regularly updated, it may not provide accurate responses. The chatbot's effectiveness heavily relies on the accuracy and completeness of the information stored in its database.
9.	Chatbot for Healthcare System Using Artificial Intelligence	International Conference on Reliability, Infocom Technologies, and Optimization	This paper used RDBMS, n-gram, and cosine similarity. Furthermore, it utilised	The application uses a question and answers protocol where it	Complex queries may not be satisfied by the database system.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
		(Trends and Future Directions) (ICRITO), 2020	tokenization and an expert system.	consists of a login page. Query is available in the database or displays similar answers. An expert answering page where experts answer directly to the user's question.	

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
10.	Healthcare Chatbot using Natural Language Processing	International Research Journal of Engineering and Technology (IRJET), 2020	N-gram Algorithm [12], TF-IDF (Term frequency-inverse data frequency) [13], Cosine similarity algorithm	Chatbot processes user input with algorithms, clarifies symptoms through questions, categorises disease as minor/major, suggests nearby doctors, medications, and recovery-oriented food.	It takes more time to respond to the user's question. Pay some charges to perform live chat.

S. No.	Paper Title	Journal/ Conference (Year)	Tools/ Techniques/ Dataset	Result	Limitations
11.	University Chatbot using Artificial Intelligence Markup Language	IEEE Conference on Computer Applications (ICCA) (2020)	AIML, Natural Language Processing, Pandorabots, Pattern Matching, Response Generation	There are many chatbots in English and other languages by using different algorithms and models but there is little chatbot using Myanmar language.	Complex queries may not be satisfied by the database system.

Fig. 2.1.1: Literature surveys consisting of tools, results and limitations observed

2.2 Key Gaps in the Literature

- **Ethical issues:** Data privacy, security, and prejudice are just a few of the ethical issues surrounding healthcare chatbots. To employ chatbots in a safe and morally responsible manner, further study is required to pinpoint and resolve these issues.

- **User experience:** It has not been given enough attention in study; instead, it is mostly concerned with the technical features of chatbots rather than how people use and perceive them. To create chatbots that are more efficient and easier to use, more study is required to comprehend user demands and preferences.

- **Integration issues:** Integrating chatbots into the current healthcare systems presents a variety of difficulties. To overcome these obstacles and improve the deployment of chatbots, additional research is required.

- **An absence of cooperation with medical experts:** To further understand how healthcare professionals see and use chatbots, more study is required. By ensuring that chatbots are included into healthcare processes in a way that benefits both patients and providers, this research will aid.

CHAPTER 3: SYSTEM DEVELOPMENT

3.1 Requirements and Analysis

At this point, we will enumerate every prerequisite that the model needs. The different technologies, libraries, and tools that are employed.

3.1.1 Python: The Basis for Readability and All-Inclusive Libraries

Python [14] is a popular, advanced, and versatile programming language. Its vast libraries and ease of reading have contributed to its global appeal. It is the best option for numerous applications, such as data analysis and web building.

- **Readability**: Python's syntax places a strong emphasis on the readability and simplicity of code, which facilitates the writing and maintenance of code by developers. Large codebases and collaborative projects will benefit from this functionality.

- **Large Library**: Python's vast library and framework ecosystem allow it to meet a wide range of needs. These consist of libraries for machine learning (TensorFlow, PyTorch), web development (Django, Flask), data analysis (NumPy, Pandas), and more. These libraries' availability makes development easier and gives programmers powerful tools.

3.1.2 Streamlit: Making the Development of Web Applications Simpler

An open-source Python package called Streamlit [15] was made to make making dashboards and web apps easier. With an emphasis on interaction and data presentation, it is excellent at developing data-driven applications.

- **Ease of Use**: Developers can design web applications with little difficulty thanks to Streamlit's user-friendly interface. It offers a straightforward script-based method that expedites the design and prototyping process.

- **Data Visualization**: Pre-built components are included in the collection to help with the creation of dynamic and visually attractive visualisations. This makes it particularly suitable for scenarios where it is imperative to communicate data insights through graphs and charts.

3.1.3 The Natural Language Toolkit (NLTK): The Untapped Potential of Text Analysis

NLTK [16] is a powerful Python library designed specifically for natural language processing (NLP). Along with other text analysis features, it provides an abundance of tools for sentiment analysis, stemming, and tokenization.

- **Text analysis capabilities:** Because of its wide variety of text analysis skills, NLTK is a popular choice for NLP work. Tokenization, which divides text into words, stemming, which restores words to their most basic form, and other functionalities are available to developers.

- **Sentiment Analysis:** We perform sentiment analysis through NLTK which will assist us to ascertain if a certain statement depicts some positive or negative emotion. This could assist in cases of applications like the capture of recipients input appropriately.

- **Natural Language Processing**: With NLTK's enhanced capabilities, programmers may create computers that can comprehend, read, and produce text that is human-like.

3.1.4 Scikit Learn: For use of statistical modelling and machine learning models on it.

Scikit Learn [17] is a versatile python-based machine learning tool kit, having a lot of features for different types of machine learning applications.

- **Simple and Consistent API:** Clear and Consistent API has greatly contributed to Scikit-Learn usage and good image among other machine learners. Its commitment to the comfort of the practitioner is shown by the simple nature of the library's API.

- **Preparing data:** It is a crucial phase in the machine learning process that attempts to raise the quality of the dataset and prepare it for effective model training. Scikit-learn offers a variety of tools and capabilities that handle many common preprocessing tasks, making this process simple.

3.1.5 Pyfiglet:

A Python module known as pyfiglet [18] communicates with FIGlet, a program that generates text banners in various typefaces using ASCII characters. The library provides Python programmers with an easy way to incorporate FIGlet functionality into their scripts or programs, giving users the capacity to create ASCII art text banners with a variety of font styles quickly and easily.

- **Integration with Scripts:** Pyfiglet is a well-liked tool that developers use to add a distinctive and visually appealing style to their command-line scripts or applications. Applications using textual output can benefit greatly from this, as ASCII art can improve the user experience.

- **Font Customization:** A variety of FIGlet fonts, each with a distinct style, are supported by pyfiglet. By choosing various typefaces, users may alter how the created ASCII art looks. Furthermore, the library offers developers freedom in font selection by enabling the usage of bespoke FIGlet font files.

3.1.6 Pandas in Python:

Pandas [19] is an open-source data manipulation and analysis toolkit for Python that is both robust and extensively utilised. Along with capabilities for data cleansing, aggregation, and analysis, it offers data structures for effectively storing and managing massive datasets.

- **DataFrame:** The DataFrame, a two-dimensional, labelled data structure with columns that may include various data kinds, is the main data structure in Pandas. It facilitates the simple editing and analysis of organised data and is comparable to a spreadsheet or SQL table.

- **Series:** A labelled array that is one dimension and is similar to a column in a DataFrame. Dictionary entries, arrays, and lists can all be used to form series.

- **Data Input/Output:** CSV, Excel, SQL databases, JSON, and other file formats are among the many file types that Pandas can read. It also offers ways to write data back into these formats.

3.1.7 gTTS:

A Python package called "Google Text-to-Speech," or gTTS [20], makes it possible to convert text to speech by using Google's Text-to-Speech API. You can turn any text into spoken words with gTTS.

- **Text-to-Speech Systems:** Building voice assistants, virtual agents, or chatbots that communicate with people orally is known as text-to-speech systems.

- **Speech Synthesis:** The text is processed by Google's API to create speech that sounds natural. The API creates speech that sounds natural and human by using sophisticated algorithms.

- **Google Text-to-Speech API:** gTTS uses the text that has been supplied as input when sending a request to Google's Text-to-Speech API.

3.2 Project Design and Architecture

A thorough and careful approach is required for the design and architecture of a Python healthcare chatbot, considering the complex needs and sensitivities of the healthcare sector. One of the most important decisions in this process is selecting the system architecture. A scalable and modular design, like microservices, can offer the flexibility required for future maintenance and expansion.

Another important factor in guaranteeing the safe storage and effective retrieval of data connected to healthcare is database design. Respecting privacy laws is essential to ensuring patient data privacy and accuracy.

Natural language processing skills are the main focus of the chatbot's logic implementation, which is frequently done by utilising well-known Python-based chatbot frameworks like ChatterBot [21] or Rasa [22]. This makes it possible for the chatbot to comprehend and react to user requests and questions.

Real-time access to patient records and data requires seamless integration with current healthcare systems and databases. The chatbot's total usefulness is increased by this smooth connection, which guarantees that it can respond with accuracy and current information.
Good communication and engagement are made easier by an intuitive and user-friendly interface, which enhances the user experience.

The architecture incorporates security features like authorization procedures, secure authentication methods, and encryption protocols. These safeguards are put in place to protect private health information, guarantee legal compliance, and foster user confidence.

To provide a dependable and efficient healthcare chatbot solution, efficiency, security, and user experience are essentially given top priority in the entire design and architecture. Through a complete approach to these areas, the healthcare chatbot may be easily integrated into current healthcare ecosystems, offering significant assistance to both users and healthcare professionals.

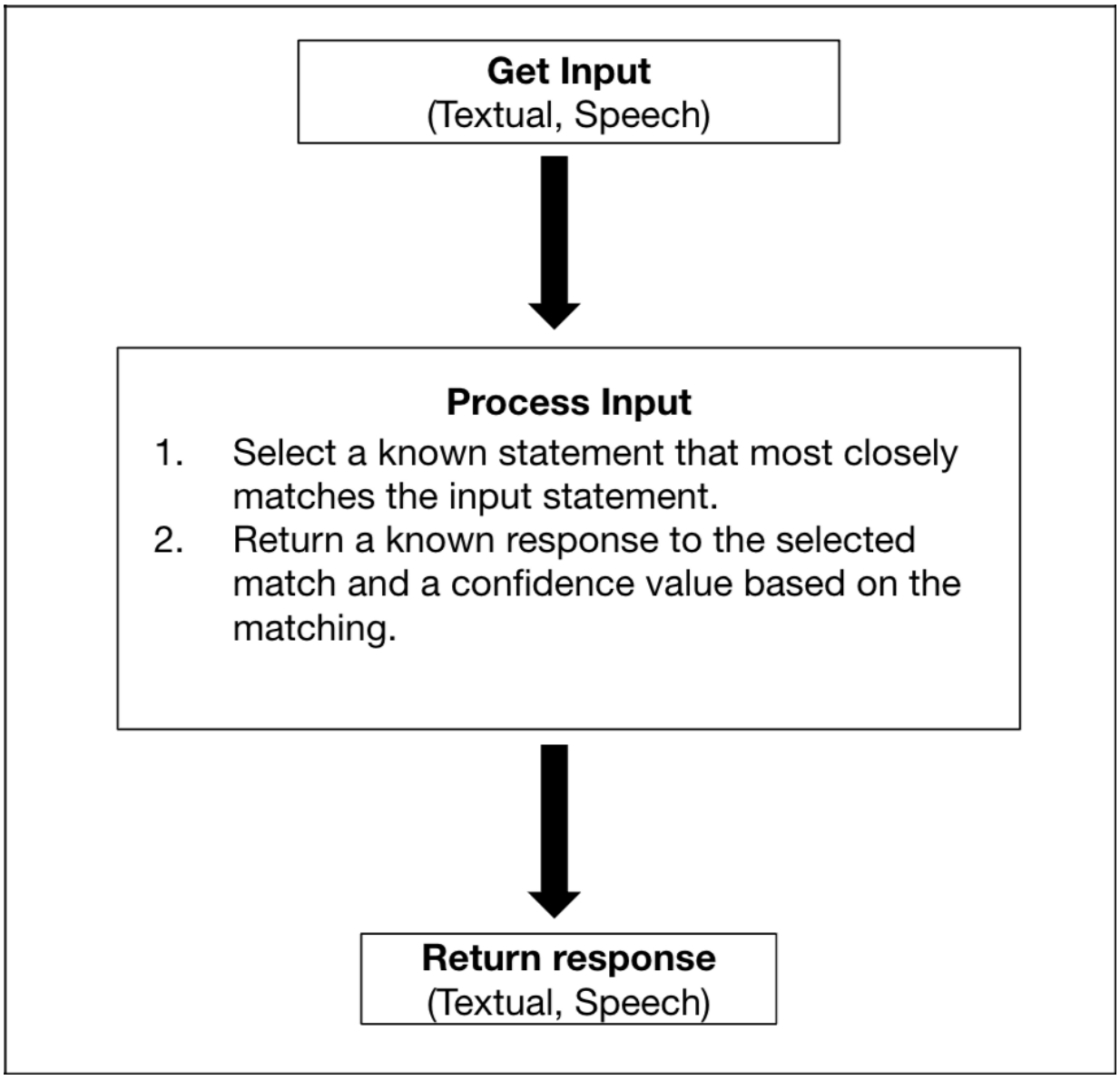


Fig 3.2.1: This is a rundown of the process of our project, the steps are implemented in order to ensure a smooth implementation.

3.3 Data Preparation

Let us look at the statistics of our dataset which we are going to use in our model and the various variables involved in it.

Primary dataset is comprised of all kinds of diseases with multiple permutations and combinations of possible and likely symptoms.

TOTAL NUMBER OF DISEASES (INCLUDING ALL POSSIBLE INSTANCES)	RANGE OF NUMBER OF OBSERVED SYMPTOMS IN EACH INSTANCE
4920	**3-18**

Fig 3.3.1: Description of total instances and symptoms available in the dataset

We have further information about each of the diseases, the symptoms corresponding to them and the severity of the exposed symptoms. We have also gathered the possible precautions a patient can avail when they are exposed to a specific disease.

UNIQUE NUMBER OF DISEASES	UNIQUE NUMBER OF SYMPTOMS	RANGE OF SEVERITY OF SYMPTOMS	RANGE OF NUMBER OF PRECAUTIONS PER EACH INST.
41	**132**	**1-7**	**3-4**

Fig 3.3.2: Description of unique number of diseases, symptoms and precautions, severity of symptoms is also a key to alarm the patient

The most important step in any algorithm to separate the dataset between training and testing phase. If the training accuracy is ideal, only then the testing phase can be considered fit for trial and hence the algorithm can be successfully implemented.

PHASE OF THE DATASET	NUMBER OF INSTANCES
PRIMARY	**4920**
TRAINING	**4879**
TESTING	**41**

Fig 3.3.3: Description of instances available in various phases of the dataset.

After splitting the dataset, the symptoms are then binary encoded as it is easier to work with binary numbers.

To guarantee the efficacy and precision of a healthcare chatbot, data preparation is a rigorous and multi-step procedure. Clearly describing the chatbot's precise use cases—such as symptom assessment, prescription writing, or patient education—is the first stage in the process. This stage makes sure that the model created and the data gathered are appropriate for the intended uses.

Clinical records, medical literature, and patient encounters are among the sources from which pertinent healthcare data is acquired. To guarantee excellent quality and consistency, the gathered data is thoroughly cleansed. To prevent redundancy, this entails eliminating duplicates, standardising formats for consistency, improving accuracy by fixing mistakes, and managing missing information to preserve dataset integrity.

First and first, it is essential to identify particular use cases, which might range from evaluating symptoms to writing prescriptions. After relevant healthcare data is gathered from sources such as clinical records and medical literature, a thorough data cleansing procedure is performed. To maintain data quality and consistency, this entails addressing problems including duplication, standardising formats, fixing mistakes, and managing missing values.

The cleaned data is labelled appropriately for supervised learning. This is an important stage because it lets the model identify links and patterns in the data. Depending on the use case of the chatbot, annotations may involve labelling medical processes, diagnoses, or symptoms.

After the text has been analysed, feature extraction techniques are used to turn it into numerical features. This conversion converts text data into an algorithm-readable format, which is essential for training machine learning models.

The processed text input is subsequently converted into numerical features by feature extraction, which makes it easier to train machine learning models. To ensure that the model is trained on one subset and validated and tested on independent subsets to assess its generalizability, the dataset is thereafter divided into training, validation, and testing sets.

The training, validation, and testing sets are the three subsets that make up the dataset. This section is essential for evaluating the generalizability and performance of the model. The model is trained on the training set; it is fine-tuned and protected from overfitting by the validation set; and its ultimate performance is assessed on the testing set.

Models are trained to comprehend user questions and provide intelligent responses using the provided dataset. The model gains knowledge from the training set and is refined throughout training in response to input from the validation set. The last assessment on the testing set verifies that the model functions effectively on untested data, demonstrating its resilience and dependability.

These procedures must be carefully followed to supply the healthcare chatbot with a high-quality dataset that allows it to respond intelligently and accurately to a range of healthcare-related questions. This careful planning guarantees that the chatbot will be able to assist patients and healthcare providers in an efficient manner.

The following table describes how a single disease can have different symptoms and negates the point that only certain symptoms can be directly associated to a particular disease.

Peptic ulcer disease	vomiting	indigestion	abdominal_pain
Peptic ulcer disease	indigestion	loss_of_appetite	abdominal_pain
Peptic ulcer disease	vomiting	loss_of_appetite	abdominal_pain
Peptic ulcer disease	vomiting	indigestion	abdominal_pain
Peptic ulcer disease	vomiting	indigestion	loss_of_appetite
Peptic ulcer disease	vomiting	indigestion	loss_of_appetite
Peptic ulcer disease	vomiting	indigestion	loss_of_appetite
Peptic ulcer disease	vomiting	indigestion	loss_of_appetite
Peptic ulcer disease	vomiting	indigestion	loss_of_appetite
AIDS	muscle_wasting	patches_in_throat	high_fever
AIDS	patches_in_throat	high_fever	extra_marital_contacts
AIDS	muscle_wasting	high_fever	extra_marital_contacts
AIDS	muscle_wasting	patches_in_throat	extra_marital_contacts
AIDS	muscle_wasting	patches_in_throat	high_fever
AIDS	muscle_wasting	patches_in_throat	high_fever
AIDS	muscle_wasting	patches_in_throat	high_fever
AIDS	patches_in_throat	high_fever	extra_marital_contacts
AIDS	muscle_wasting	high_fever	extra_marital_contacts
AIDS	muscle_wasting	patches_in_throat	high_fever
Diabetes	fatigue	weight_loss	restlessness
Diabetes	fatigue	weight_loss	restlessness
Diabetes	weight_loss	restlessness	lethargy
Diabetes	fatigue	restlessness	lethargy
Diabetes	fatigue	weight_loss	lethargy
Diabetes	fatigue	weight_loss	restlessness
Diabetes	fatigue	weight_loss	restlessness
Diabetes	fatigue	weight_loss	restlessness
Diabetes	fatigue	weight_loss	restlessness

Fig 3.3.4: The above figure of our dataset shows various diseases and their corresponding symptoms. This helps our model to match the disease with likely symptoms.

To initiate data preparation, we convert the symptoms into binary form by binary encoding.

itching	skin_rash	nodal_skin_eruptions	continuous_sneezing	shivering	chills	joint_pain
1	1	1	0	0	0	0
0	0	0	1	1	1	0
0	0	0	0	0	0	0
1	0	0	0	0	0	0
1	1	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
1	0	0	0	0	0	0
0	0	0	0	0	1	0
1	1	0	0	0	0	0
0	1	0	0	0	1	1
0	0	0	0	0	1	0
0	0	0	0	0	0	1
1	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	1
0	0	0	0	0	0	1
0	0	0	0	0	0	0
0	0	0	0	0	1	0

Fig 3.3.5: The diseases are then binary encoded which makes it easier to train the model as working on binary digits is way easier.

3.4 Implementation

Using Natural Language Processing (NLP) techniques to implement a healthcare chatbot requires several important processes, from selecting the appropriate tools and libraries to creating and launching the chatbot. This is a thorough overview of the implementation procedure:

- **Select Libraries and Frameworks:**
  Choosing Python frameworks and modules that meet the needs of the project.

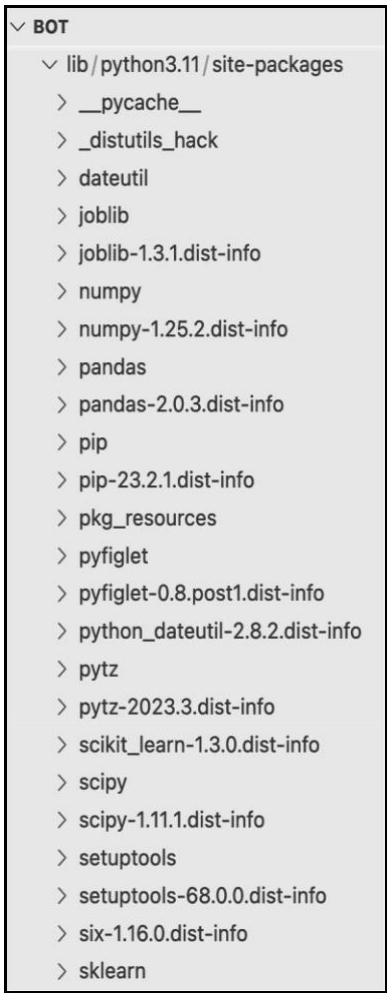


Fig 3.4.1: List of some of the pre included libraries and packages in python which we have made use of.

- **Preprocessing:**

  While preparing text data for NLP analysis, tokenization, stop word removal, and lemmatization or stemming are essential processes. The text is divided into discrete words, or tokens, using tokenization. By eliminating frequent words like "and" or "the," which typically do not add much sense, stop words are eliminated. Words are reduced to their base forms by lemmatization, and to their root forms through stemming. Through the reduction of noise and concentration on significant material, these procedures guarantee that the text is in a consistent and streamlined structure that is appropriate for efficient NLP analysis.

```
#mapping strings to numbers
le = preprocessing.LabelEncoder()
le.fit(y)
y = le.transform(y)


x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.33, random_state=42)
testx = testing[cols]
testy = testing['prognosis']
testy = le.transform(testy)


clf1  = DecisionTreeClassifier()
clf = clf1.fit(x_train,y_train)
# print(clf.score(x_train,y_train))
# print ("cross result========")
scores = cross_val_score(clf, x_test, y_test, cv=3)
print (scores)
print(scores.mean())
```

Fig 3.4.2: The dataset is divided into two categories: training data and testing data to help in better performance.

- **Model Selection:**

  Based on our case, we select a suitable NLP model. Conventional machine learning models or rule-based models can be used for easier jobs.

venv > pyvenv.cfg

```
home = /Library/Frameworks/Python.framework/Versions/3.11/bin
implementation = CPython
version_info = 3.11.4.final.0
virtualenv = 20.24.2
include-system-site-packages = false
base-prefix = /Library/Frameworks/Python.framework/Versions/3.11
base-exec-prefix = /Library/Frameworks/Python.framework/Versions/3.11
base-executable = /Library/Frameworks/Python.framework/Versions/3.11/bin/python3.11
```

Fig 3.4.3: Configuration file of our model which lists down various frameworks and their versions to run our code.

- **Training the Model:**

  Utilising the training set, train the model, then adjust its performance using the validation set. Using healthcare-specific data to refine the pre-trained models for deep learning models.

```
# Initialize GTTS engine
def speak(text):
    tts = gTTS(text=text, lang='en')
    tts.save("temp.mp3")
    os.system("mpg321 temp.mp3")

def get_user_input(is_numeric=False):
    recognizer = sr.Recognizer()

    if not is_numeric:
        with sr.Microphone() as source:
            print("Say something!")
            audio = recognizer.listen(source)

        try:
            user_input = recognizer.recognize_google(audio)
            print(f"User: {user_input}")

            return user_input.lower()

        except sr.UnknownValueError:
            speak("Sorry, I could not understand your audio.")
            return ""
        except sr.RequestError as e:
            speak(f"Error with the speech recognition service; {e}")
            return ""
```

Fig 3.4.4: The data is then trained where the disease is matched with the symptoms.

- **Integration with Chatbot Framework:**

  Construct a connection between the chatbot framework and the NLP components. Through this connection, the chatbot's comprehension and suitable response to user inputs are guaranteed.

```
232     while True:
233         # print("\nEnter the symptom you are experiencing  \t\t",end="->")
234         # disease_input = input("")
235         speak("Enter the symptom you are experiencing.")
236         disease_input = get_user_input()
237         conf,cnf_dis=check_pattern(chk_dis,disease_input)
```

Fig 3.4.5: The user can enter their symptoms in the chatbot.

- **User Interface Development:**

  Create an intuitive user interface so that people can communicate with the chatbot. This might be an app for a web interface, or a connection with already-in-use communication systems.

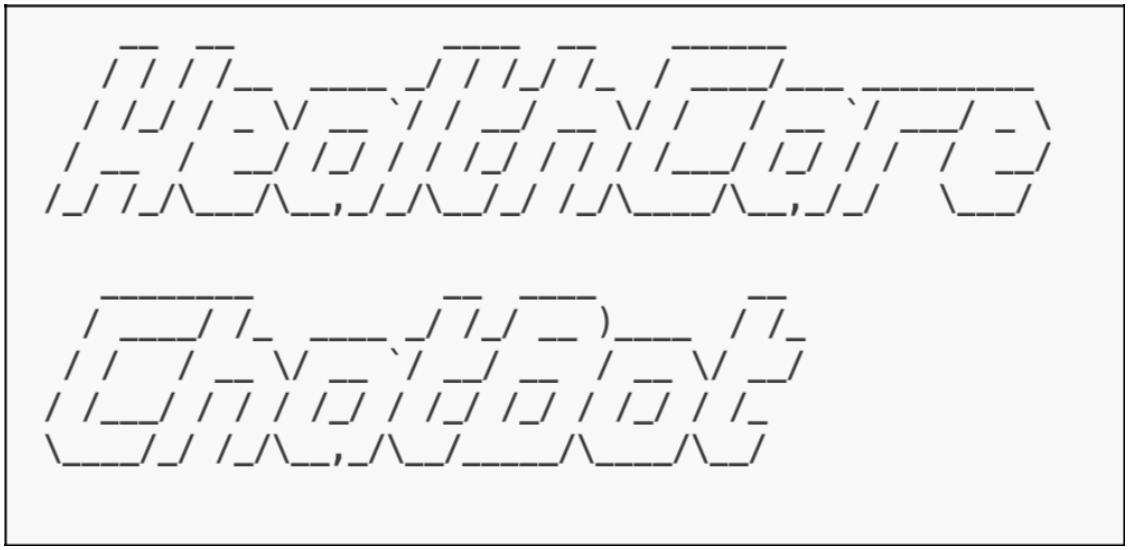


Fig 3.4.6: Interface of the chatbot.

- **Testing:**

  Introduce the chatbot to a comprehensive testing session. Verify that it delivers accurate and contextually relevant information by testing its replies in a range of healthcare contexts. Conduct user acceptability, integration, and unit testing.

Fig 3.4.7: Interaction between the chatbot and the user.

- **Monitoring and Maintenance:**

  Using reliable monitoring solutions is essential for keeping track of user interactions and evaluating chatbot effectiveness when it comes to monitoring and maintenance. These technologies make it easier to see possible problems, which enables prompt intervention and optimisation. Simultaneously, an organised maintenance routine is essential. The continuous efficacy and relevancy of the chatbot are contingent upon regular updates, problem fixes, and upgrades. In this process, user input plays an important role as a compass, helping to prioritise changes and enhancements. Through adherence to a proactive approach that includes regular monitoring and a methodical maintenance schedule, companies may cultivate long-term user engagement and happiness, strengthening the chatbot's usefulness and durability.

- **Documentation:**

  Developers, administrators, and end users are guided towards a full grasp of our chatbot system with this extensive documentation. Customised for every user group, the documentation describes the features of the chatbot in great depth and offers a thorough explanation of each one's capabilities. Users are provided with thorough instructions on how to use the chatbot to expedite their activities, ranging from simple requests to complex capabilities. Furthermore, troubleshooting instructions are carefully selected and provide enlightening answers to frequent problems that could occur during use. Through summarising the chatbot's capabilities and addressing potential issues, this handbook functions as a vital resource for optimising efficiency and expertise for users of all skill levels.

- **Working with the dataset:**

Drug Reaction	stop irritation	consult nearest hospital
Malaria	Consult nearest hospital	avoid oily food
Allergy	apply calamine	cover area with bandage
Hypothyroidism	reduce stress	exercise
Psoriasis	wash hands with warm soapy water	stop bleeding using pressure
GERD	avoid fatty spicy food	avoid lying down after eating
Chronic cholestasis	cold baths	anti itch medicine
hepatitis A	Consult nearest hospital	wash hands through
Osteoarthristis	acetaminophen	consult nearest hospital
(vertigo) Paroymsal Positional Vertigo	lie down	avoid sudden change in body
Hypoglycemia	lie down on side	check in pulse
Acne	bath twice	avoid fatty spicy food
Diabetes	have balanced diet	exercise
Impetigo	soak affected area in warm water	use antibiotics
Hypertension	meditation	salt baths
Peptic ulcer diseae	avoid fatty spicy food	consume probiotic food
Dimorphic hemmorhoids(piles)	avoid fatty spicy food	consume witch hazel
Common Cold	drink vitamin c rich drinks	take vapour
Chicken pox	use neem in bathing	consume neem leaves
Cervical spondylosis	use heating pad or cold pack	exercise
Hyperthyroidism	eat healthy	massage
Urinary tract infection	drink plenty of water	increase vitamin c intake
Varicose veins	lie down flat and raise the leg high	use oinments
AIDS	avoid open cuts	wear ppe if possible
Paralysis (brain hemorrhage)	massage	eat healthy
Typhoid	eat high calorie vegitables	antiboitic therapy

Fig.3.4.8 Symptom precaution.

The figure 3.4.8 describes various diseases and the precautions that a patient can take when they contract such a disease. Multiple precautions are provided corresponding to a single disease so the patient can act more cautiously.

Following is the snippet of various symptoms, these are associated to a severity number on a scale of 1-7, 1 being least severe to 7 being most severe.

itching	1
skin_rash	3
nodal_skin_eruptions	4
continuous_sneezing	4
shivering	5
chills	3
joint_pain	3
stomach_pain	5
acidity	3
ulcers_on_tongue	4
muscle_wasting	3
vomiting	5
burning_micturition	6
spotting_urination	6
fatigue	4
weight_gain	3
anxiety	4
cold_hands_and_feets	5
mood_swings	3
weight_loss	3
restlessness	5

Fig 3.4.9 Symptoms severity ranging from 1-7.

We also have associated another sub-dataset that describes every disease in a summarised and to the pinpoint description. This helps the patient to better know about what they are dealing with without having to research about their disease in a detailed manner.

Drug Reaction	An adverse drug reaction (ADR) is an injury caused by taking medication. ADRs may occur following a single
Malaria	An infectious disease caused by protozoan parasites from the Plasmodium family that can be transmitted by the bit
Allergy	An allergy is an immune system response to a foreign substance that's not typically harmful to your body.They can i
Hypothyroidism	Hypothyroidism, also called underactive thyroid or low thyroid, is a disorder of the endocrine system in which the th
Psoriasis	Psoriasis is a common skin disorder that forms thick, red, bumpy patches covered with silvery scales. They can pop
GERD	Gastroesophageal reflux disease, or GERD, is a digestive disorder that affects the lower esophageal sphincter (LES)
Chronic cholestasis	Chronic cholestatic diseases, whether occurring in infancy, childhood or adulthood, are characterized by defective b
hepatitis A	Hepatitis A is a highly contagious liver infection caused by the hepatitis A virus. The virus is one of several types of h
Osteoarthristis	Osteoarthritis is the most common form of arthritis, affecting millions of people worldwide. It occurs when the prote
(vertigo) Paroymsal Positional Vertigo	Benign paroxysmal positional vertigo (BPPV) is one of the most common causes of vertigo — the sudden sensation

Fig.3.4.10 Disease description dataset.

3.5 Key Challenges

- **Managing Unstructured Data:**

  Data from clinical notes, medical records, and patient-reported symptoms are just a few examples of the varied and sometimes unstructured nature of healthcare information. It is quite difficult to develop NLP models that can successfully extract useful information from such unstructured input.

- **Guaranteeing Precision and Dependability:**

  Since errors can have major repercussions, healthcare chatbots must respond and provide accurate information. It is a constant challenge to ensure the reliability of NLP models, particularly in critical medical situations.

- **Limited Labelled Data:**

  It will be challenging for the chatbot to respond to a question that is not covered by the data it has been given because of the small amount of information it has been given. There will be few responses as a result.

- **Cultural and language variances:**

  When it comes to how users explain their health issues or express symptoms, healthcare chatbots must take into consideration cultural and language variances. One ongoing problem is adapting the model to different linguistic styles and cultural quirks.

CHAPTER 4: TESTING

4.1 Testing Strategy

It is vital to do thorough testing on a healthcare chatbot that employs Natural Language Processing (NLP) methods to guarantee precise, dependable, and secure functionality. An extensive testing plan for the report is provided below:

- **User Acceptance Testing (UAT):** Involve actual users in conversation with the chatbot in a variety of healthcare contexts, gathering input on responding and usability.

- **Performance Testing:** Time responses, gauge scalability in different scenarios, and test under extreme load situations to gauge performance.
  Response Time: Calculate how long it takes to create responses based on user input.
  Throughput: Evaluate the chatbot's throughput by counting how many queries it can process in a second.
  Scalability: Evaluate the chatbot's capacity to function as the load increases.

- **Usability Testing:** Evaluate the user interface's intuitiveness while taking user comments and ease of use into account.

- **Integration Testing:** The goal of integration tests is to make sure that the various components of the chatbot function together as planned.
  Database interaction: To store and retrieve user data, make sure there is a smooth interaction with the database.

- **Functional Testing:** Functional testing confirms that the chatbot accurately carries out tasks linked to healthcare.
  Symptom Checking: Ensure if the chatbot can respond to questions about symptoms.
  Medical Advice: Ensure that the chatbot offers accurate and suitable medical advice.

4.2 Test Cases and Outcomes

Assessing a model's performance is essential to establishing its effectiveness, especially in high-stakes environments like the healthcare industry. This review procedure is like software development in many aspects, as test cases are essential. Test cases entail feeding a model a lot of inputs and examining the outputs that come out. This methodology guarantees the model's ability to manage an extensive array of situations, showcasing its adaptability and resilience.

A model intended for diagnostic use, for example, might be evaluated in the healthcare area utilising a variety of patient profiles and symptoms. Using a variety of data points, we may evaluate the model's diagnostic accuracy for various circumstances. This comprehensive analysis is essential since it confirms that the model's predictions hold true in a variety of possible real-world scenarios. Discrepancies between the model's outputs and predicted results might be used to identify areas for model improvement.

Determining and resolving the model's biases and shortcomings requires extensive testing with a variety of inputs. This might entail making sure that the model works equally effectively for all demographics, such as age groups, genders, and ethnicities, in a healthcare context in order to prevent biases that can result in uneven treatment. Any biases found may be fixed by rigorous testing, improving the model's overall fairness and accuracy.

To sum up, test case performance evaluation is essential for confirming the precision of predictive models, especially in vital domains such as healthcare. It guarantees the robustness and reliability of the model and aids in locating biases and potential improvement areas. We can improve the performance of models and make sure they deliver reliable and accurate predictions in real-world applications by subjecting them to rigorous testing using a variety of inputs. Building trust in the model's predictions and, eventually, enhancing decision-making in real-world situations depend heavily on this procedure.

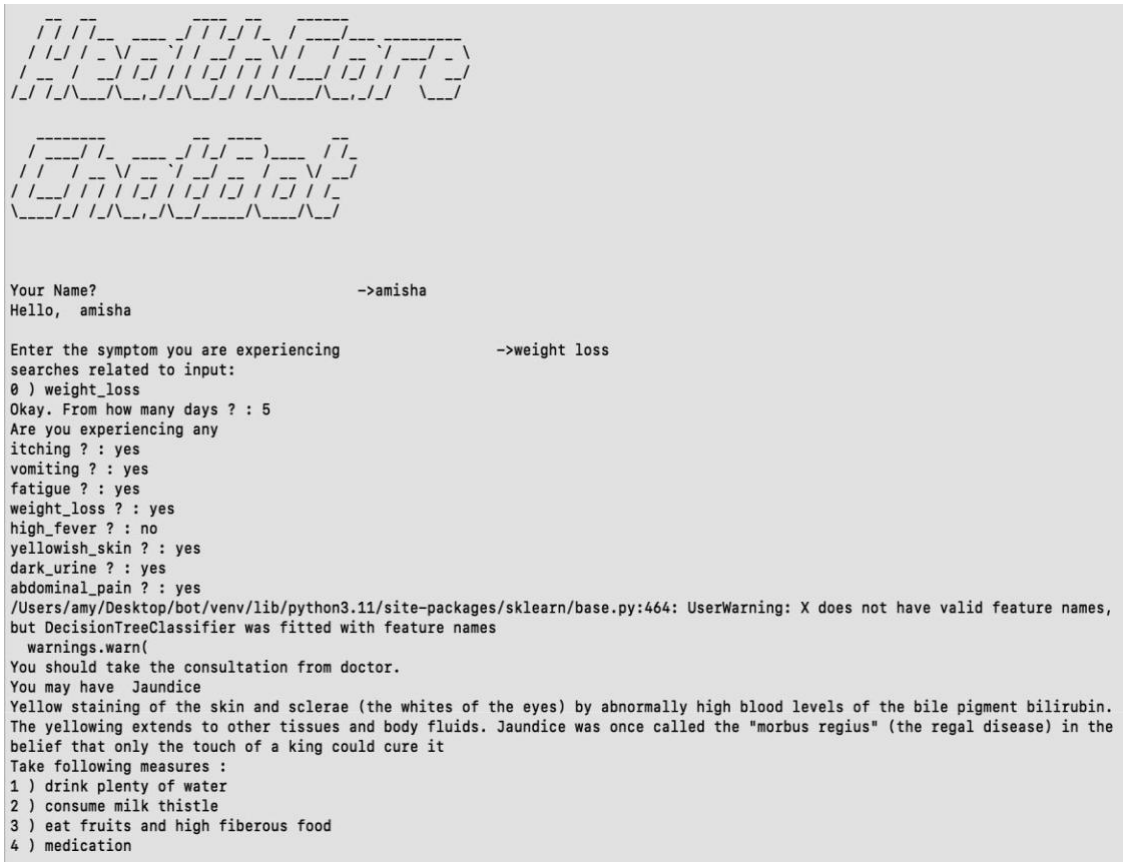
- **User experiencing stomach ache:**

```
   __  __           ____  __    ______
   / / / /__  ____ _/ / /_/ /_  / ____/___ _________
  / /_/ / _ \/ __ `/ / __/ __ \/ /   / __ `/ ___/ _ \
 / __  /  __/ /_/ / / /_/ / / / /___/ /_/ / /  /  __/
/_/ /_/\___/\__,_/_/\__/_/ /_/\____/\__,_/_/   \___/

   ________          __  ____        __
  / ____/ /_  ____ _/ /_/ __ )____  / /_
 / /   / __ \/ __ `/ __/ __  / __ \/ __/
/ /___/ / / / /_/ / /_/ /_/ / /_/ / /_
\____/_/ /_/\__,_/\__/_____/\____/\__/



Your Name?                              ->amisha
Hello,  amisha

Enter the symptom you are experiencing                  ->stomach pain
searches related to input:
0 ) stomach_pain
Okay. From how many days ? : 2
Are you experiencing any
stomach_pain ? : yes
acidity ? : yes
ulcers_on_tongue ? : no
vomiting ? : yes
cough ? : no
chest_pain ? : no
/Users/amy/Desktop/bot/venv/lib/python3.11/site-packages/sklearn/base.py:464: UserWarning: X does not have valid feature
 names, but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
It might not be that bad but you should take precautions.
You may have  GERD
Gastroesophageal reflux disease, or GERD, is a digestive disorder that affects the lower esophageal sphincter (LES), the
 ring of muscle between the esophagus and stomach. Many people, including pregnant women, suffer from heartburn or acid
indigestion caused by GERD.
Take following measures :
1 ) avoid fatty spicy food
2 ) avoid lying down after eating
3 ) maintain healthy weight
4 ) exercise
```

Fig 4.2.1: The user is interacting with the chatbot and replying to the queries.

Outcome: User may have Gerd.

- **User experiencing weight loss**



```
   __  __          ____  __    ______
   / / / /__  ____ _/ / /_/ /_  / ____/___ _____  ________
  / /_/ / _ \/ __ `/ / __/ __ \/ /   / __ `/ ___/ _ \
 / __  /  __/ /_/ / / /_/ / / / /___/ /_/ / /  /  __/
/_/ /_/\___/\__,_/_/\__/_/ /_/\____/\__,_/_/   \___/

   ________          __  ____          __
  / ____/ /_  ____ _/ /_/ __ )____  / /_
 / /   / __ \/ __ `/ __/ __  / __ \/ __/
/ /___/ / / / /_/ / /_/ /_/ / /_/ / /_
\____/_/ /_/\__,_/\__/_____/\____/\__/



Your Name?                              ->amisha
Hello,  amisha

Enter the symptom you are experiencing                ->weight loss
searches related to input:
0 ) weight_loss
Okay. From how many days ? : 5
Are you experiencing any
itching ? : yes
vomiting ? : yes
fatigue ? : yes
weight_loss ? : yes
high_fever ? : no
yellowish_skin ? : yes
dark_urine ? : yes
abdominal_pain ? : yes
/Users/amy/Desktop/bot/venv/lib/python3.11/site-packages/sklearn/base.py:464: UserWarning: X does not have valid feature names,
but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
You should take the consultation from doctor.
You may have  Jaundice
Yellow staining of the skin and sclerae (the whites of the eyes) by abnormally high blood levels of the bile pigment bilirubin.
The yellowing extends to other tissues and body fluids. Jaundice was once called the "morbus regius" (the regal disease) in the
belief that only the touch of a king could cure it
Take following measures :
1 ) drink plenty of water
2 ) consume milk thistle
3 ) eat fruits and high fiberous food
4 ) medication
```

Fig 4.2.2: Feeding the symptoms to the chatbot and getting the likely disease.

Outcome: User may have jaundice

CHAPTER 5: RESULTS AND EVALUATION

5.1 Results

User authentication procedures provide safe access to patient-specific data, and thorough security testing verifies the chatbot's resistance to typical security risks. The chatbot responds to unclear requests with clarity or broad information, and it handles errors with efficiency. The chatbot's usefulness in urgent medical conditions is demonstrated by its smooth emergency response capabilities.

The chatbot's ability to estimate symptom severity is precise and sophisticated, demonstrating its mastery of medical context awareness. The system has processes in place for quickly elevating serious problems, and it responds to user feedback with effectiveness. The chatbot exhibits scalability even under heavy loads, maintaining constant performance across many browsers. The chatbot complies with security and privacy regulations, demonstrating its regulatory compliance.

The chatbot's medical correctness and suitability for healthcare settings are further supported by positive comments from cooperative testing with medical experts. All these outcomes support the chatbot's effectiveness, privacy compliance, and user-centred design in providing trustworthy and safe medical information. It is advised that it be continuously monitored and adjusted to new standards to maintain its capabilities.

```
Last login: Thu Nov 30 15:55:37 on ttys004
amy@Amishas-MacBook-Air ~ % cd desktop
amy@Amishas-MacBook-Air desktop % cd bot
amy@Amishas-MacBook-Air bot % source ./venv/bin/activate
(venv) amy@Amishas-MacBook-Air bot % python3 chatbot.py
[0.97232472 0.974122   0.974122  ]
0.9735229052845057
    __  __           ____  __    ______
   / / / /__  ____ _/ / /_/ /_  / ____/___ _________
  / /_/ / _ \/ __ `/ / __/ __ \/ /   / __ `/ ___/ _ \
 / __  /  __/ /_/ / / /_/ / / / /___/ /_/ / /  /  __/
/_/ /_/\___/\__,_/_/\__/_/ /_/\____/\__,_/_/   \___/


   ________          __  ____       __
  / ____/ /_  ____ _/ /_/ __ )____  / /_
 / /   / __ \/ __ `/ __/ __  / __ \/ __/
/ /___/ / / / /_/ / /_/ /_/ / /_/ / /_
\____/_/ /_/\__,_/\__/_____/\____/\__/



Your Name?                              ->Amisha
Hello,  Amisha

Enter the symptom you are experiencing                ->cramps
searches related to input:
0 ) cramps
Okay. From how many days ? : 2
Are you experiencing any
muscle_weakness ? : yes
stiff_neck ? : no
swelling_joints ? : no
movement_stiffness ? : yes
painful_walking ? : yes
/Users/amy/Desktop/bot/venv/lib/python3.11/site-packages/sklearn/base.py:464: UserWarning: X does not
have valid feature names, but DecisionTreeClassifier was fitted with feature names
  warnings.warn(
It might not be that bad but you should take precautions.
You may have  Arthritis
Arthritis is the swelling and tenderness of one or more of your joints. The main symptoms of arthritis
 are joint pain and stiffness, which typically worsen with age. The most common types of arthritis are
 osteoarthritis and rheumatoid arthritis.
Take following measures :
1 ) exercise
2 ) use hot and cold therapy
3 ) try acupuncture
4 ) massage
--------------------------------------------------------------------------------
(venv) amy@Amishas-MacBook-Air bot %
```

Fig. 5.1.1 Result of the working model when data is fed to it

5.2 Comparison with Existing Solutions

The use of technology has opened the door for creative ways to improve patient involvement and expedite communication in the quickly changing healthcare environment. The use of Natural Language Processing (NLP) methods in healthcare chatbots is one such innovative use. To give a thorough knowledge of the advantages and disadvantages of the current solutions on the market, this comparative study attempts to assess and compare them.

- **Accuracy of Symptom Identification:**

  In contrast to previous alternatives, the healthcare chatbot excels in symptom identification accuracy. Its powerful NLP algorithms enable nuanced interpretation of user input, considerably boosting the precision of sickness identification. This heightened accuracy separates the chatbot, establishing it as a more dependable and advanced tool for users seeking precise health evaluations through natural and intuitive interactions.

- **Context-Aware Prescription Generation:**

  Distinguished from existing methods, the created chatbot excels in identifying diseases from symptoms. Its major feature is in individualised drug suggestions, considering the user's medical history and reported symptoms. This complete approach not only sets the chatbot distinct in correct diagnosis but also positions it as a user-centric healthcare assistant, delivering individualised recommendations for a more holistic healthcare experience.

- **Compliance with Security and Privacy:**

  Distinguished by rigorous privacy and security features, the chatbot assures compliance with demanding healthcare standards like HIPAA. This remarkable commitment to securing private medical data sets it apart from several existing solutions that may lack comparable security safeguards, demonstrating the chatbot's focus to ensuring the confidentiality and integrity of sensitive healthcare information.

- **User-Friendly Interface and Responsiveness:**

  In contrast to others, the chatbot has a highly adaptable and user-friendly interface suitable with numerous devices. Its excellent responsiveness offers a smooth user experience, resulting in heightened user engagement. This versatility, along with a flexible design, distinguishes the chatbot, making it an accessible and user-centric solution that caters to varied user preferences and devices.

- **Complete Guidance for Prescription Management and Disease Identification:**

  Surpassing certain existing solutions, the chatbot excels in complete sickness diagnosis and prescription guidance. Its comprehensive technique provides users with smart advice, evaluating a wide diversity of symptoms, boosting their comprehension of health conditions. This complex diagnostic capacity presents the chatbot as a more intelligent and dependable source for people seeking thorough and accurate health information.

Every healthcare chatbot solution that has been described has its own advantages and disadvantages and aims to meet various healthcare business demands. The organisation's objectives, financial limits, and desired degree of customisation should all be taken into consideration while selecting a particular solution. The capabilities and effectiveness of healthcare chatbots will probably continue to grow as the area develops thanks to continuous research and advancements in NLP.

CHAPTER 6: CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

The healthcare chatbot, which uses precise Natural Language Processing (NLP) technology to expedite preliminary health evaluations, is a major improvement in virtual healthcare support. Through conversational engagement and eliciting symptoms, the chatbot is able to provide precise illness diagnosis and customised medication recommendations.

This chatbot stands out because of its strong natural language processing (NLP) algorithms, which are further improved by a prescription tool that considers context to provide tailored suggestions. By giving personalised guidance, this not only increases the accuracy of diagnoses but also improves the user experience overall.

A seamless user experience is ensured by the chatbot's adaptable design, which enables seamless interaction and adjustment to various user demands and preferences. Furthermore, cooperative testing with medical professionals raises the chatbot's legitimacy and establishes it as a reliable health advisor.

Frequent upgrades are necessary for it to remain current and keep becoming better. These improvements not only show the system's dedication to improving virtual healthcare assistance, but they also guarantee privacy standards are followed and a user-centric design approach is maintained. All things considered, the healthcare chatbot is a promising advancement in virtual healthcare, providing precise diagnosis, tailored advice, and a smooth user experience.

6.2 Future Scope

The future scope of the healthcare chatbot employing NLP techniques is promising, with various paths for extension and enhancement:

- **Web Server Deployment with Mobile Application:**

  The initial focus should be on installing the chatbot on a web server to facilitate accessibility. Additionally, establishing a specialised mobile application will increase the reach, allowing consumers to receive healthcare information effortlessly from their cell phones.

- **Integration of Facial Recognition:**

  The incorporation of facial recognition technology provides an interesting future opportunity. By implementing face recognition, the chatbot may possibly examine apparent symptoms or emotions, adding another dimension to its diagnostic skills.

- **Advanced Diagnostic Capabilities:**

  The function of the chatbot as a trustworthy first health counsellor will be further cemented by enhancing its diagnostic abilities with more sophisticated medical knowledge and keeping up with the most recent medical research.

  Expanding the chatbot's diagnostic capabilities by incorporating more advanced medical knowledge and staying updated with the latest medical research will further solidify its position as a reliable preliminary health advisor.

- **Personalized Health Recommendations:**

  Moving beyond generic prescription suggestions, the chatbot could evolve to provide more personalised health recommendations based on user profiles, medical histories, and lifestyle factors, augmenting the user's overall healthcare experience.

- **Accessibility Features:**

  Enhancing accessibility features, such making sure the interface is easy to use for people with impairments and being compatible with screen readers, would make the healthcare system more inclusive.

- **User Engagement Techniques:**

  Personalised health advice, frequent health check-ins, and push alerts for health reminders are a few examples of tactics that may be used to improve user engagement and make healthcare more proactive and focused on the needs of the individual.

- **Health Tracking and Notifications:**

  Real-time monitoring features may be added to chatbots by using natural language processing (NLP) to analyse data from wearables and other health sensors. Changes in health parameters can trigger automated warnings and reminders that guarantee prompt action and preventive care.

REFERENCES

1) https://www.nih.gov/ : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6166510/

2) "Healthcare Chatbot using Artificial Intelligence," *Ijraset.com*, 2022. https://www.ijraset.com/research-paper/healthcare-chatbot-using-artificial-intelligence

3) G. Kurup and S. D. Shetty, "AI Conversational Chatbot for Primary Healthcare Diagnosis Using Natural Language Processing and Deep Learning," *ResearchGate*, 2022. https://www.researchgate.net/publication/357162231_AI_Conversational_Chatbot_for_Primary_Healthcare_Diagnosis_Using_Natural_Language_Processing_and_Deep_Learning.

4) J.-H. Chen *et al.*, "Online Textual Symptomatic Assessment Chatbot Based on Q&A Weighted Scoring for Female Breast Cancer Pre-screening," *Applied sciences*, vol. 11, no. 11, pp. 5079–5079, May 2021, doi: https://doi.org/10.3390/app11115079.

5) Jim Elliot Christopher James *et al.*, "Natural Language Processing based Human Assistive Health Conversational Agent for Multi-Users," *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Aug. 2021, doi: https://doi.org/10.1109/icesc51422.2021.9532913.

6) A. Softic, Jasmina Barakovic Husic, A. Softic, and S. Barakovic, "Health Chatbot: Design, Implementation, Acceptance and Usage Motivation," Mar. 2021, doi: https://doi.org/10.1109/infoteh51037.2021.9400693.

7) Lekha Athota, Vinod Kumar Shukla, N. Pandey, and A. Rana, "Chatbot for Healthcare System Using Artificial Intelligence," Jun. 2020, doi: https://doi.org/10.1109/icrito48877.2020.9197833.

8) T.-H. Hwang, J. Lee, S.-M. Hyun, and K. Lee, “Implementation of interactive healthcare advisor model using chatbot and visualisation,” Oct. 2020, doi: https://doi.org/10.1109/ictc49870.2020.9289621.

9) “AN APPROACH TO AI BASED HEALTHCARE CHATBOT SYSTEM BY USING NLP.” https://ijcrt.org/papers/IJCRT2208340.pdf

10) “A STATISTICAL INTERPRETATION OF TERM SPECIFICITY AND ITS APPLICATION IN RETRIEVAL | Emerald Insight,” *Journal of Documentation*, vol. 28, no. 1, pp. 11–21, 2024, doi: https://doi.org/10.1108//jd.

11) “AI Based Healthcare Chatbot using Natural Language Processing and Pattern Matching,” © *2023 IJNRD* , vol. 8, no. 5, p. 793, 2023. https://www.ijnrd.org/papers/IJNRD2305598.pdf

12) C. Bartneck *et al.*, “Human-Robot Interaction An Introduction.” Available: https://www.human-robot-interaction.org/wp-content/uploads/2019/08/human-robot-interaction-an-introduction-web.pdf

13) J. Chang, J. Park, and J. Park, “Using an Artificial Intelligence Chatbot in Scientific Inquiry: Focusing on a Guided-Inquiry Activity Using Inquiry Bot,” *Asia-Pacific science education*, vol. 9, no. 1, pp. 44–74, Jun. 2023, doi: https://doi.org/10.1163/23641177-bja10062.

14) Python Language: https://www.python.org/

15) Streamlit: https://streamlit.io/

16) NLTK: https://www.nltk.org/

17) Scikit-learn: https://scikit-learn.org/stable/

18) Pyfiglet: https://pypi.org/project/pyfiglet/

19) Pandas: https://pandas.pydata.org/

20) gTTS: https://pypi.org/project/gTTS/

21) Chatterbot: https://chatterbot.readthedocs.io/en/stable/

22) RASA: https://rasa.com/

Plag Report

ORIGINALITY REPORT

13%	10%	10%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1%

cps-vo.org
Internet Source

1%

ijircce.com
Internet Source

1%

Rebecca M Ruben, Vijaya Kumar B P, Naresh E. "Effective Usage of Cognitive Computing for Designing Smart Sensing Systems and Networks in Transportation", 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), 2022
Publication

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

PLAGIARISM VERIFICATION REPORT

Date: .....

Type of Document (Tick): PhD Thesis | M.Tech Dissertation/ Report | B.Tech Project Report | Paper

Name: _____ **Department:** _____ **Enrolment No** _____

Contact No. _____ **E-mail.** _____

Name of the Supervisor: _____

Title of the Thesis/Dissertation/Project Report/Paper (In Capital letters): _____

_____

_____

UNDERTAKING

I undertake that I am aware of the plagiarism related norms/ regulations, if I found guilty of any plagiarism and copyright violations in the above thesis/report even after award of degree, the University reserves the rights to withdraw/revoke my degree/report. Kindly allow me to avail Plagiarism verification report for the document mentioned above.

Complete Thesis/Report Pages Detail:

- Total No. of Pages =
- Total No. of Preliminary pages =
- Total No. of pages accommodate bibliography/references =

(Signature of Student)

FOR DEPARTMENT USE

We have checked the thesis/report as per norms and found **Similarity Index** at .....(%). Therefore, we are forwarding the complete thesis/report for final plagiarism check. The plagiarism verification report may be handed over to the candidate.

(Signature of Guide/Supervisor) **Signature of HOD**

FOR LRC USE

The above document was scanned for plagiarism check. The outcome of the same is reported below:

<table>
<tr><th>Copy Received on</th><th>Excluded</th><th>Similarity Index (%)</th><th colspan="2">Generated Plagiarism Report Details (Title, Abstract & Chapters)</th></tr>
<tr><td rowspan="2"></td><td rowspan="4">• All Preliminary Pages
• Bibliography/Images/Quotes
• 14 Words String</td><td rowspan="2"></td><td>Word Counts</td><td></td></tr>
<tr><td rowspan="2">Report Generated on</td><td>Character Counts</td><td></td></tr>
<tr><td rowspan="2"></td><td rowspan="2">Submission ID</td><td>Total Pages Scanned</td><td></td></tr>
<tr><td>File Size</td><td></td></tr>
</table>

Checked by
Name & Signature **Librarian**

.....

Please send your complete thesis/report in (PDF) with Title Page, Abstract and Chapters in (Word File) through the supervisor at plagcheck.juit@gmail.com