ESCORT – UNIVERSITY STUDENTS GUIDANCE SYSTEM BASED ON NLP

Project ID: 2022-179

Project Proposal Report

Deepika Srinivasan

B.Sc. (Hons) Degree in Information Technology

Department of Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

February 2022

ESCORT – A CHAT BOT APPLICATION TO CONTACT UNIVERSITY ADMINISTRATION ABOUT ISSUES

Project ID: 2022-179

Project Proposal Report

Deepika Srinivasan

Supervisor: Ms. Archchana Kugathasan

B.Sc. (Hons) Degree in Information Technology

Department of Software Engineering

Sri Lanka Institute of Information Technology

Sri Lanka

February 2022

Declaration

We declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. `

Name	Student ID	Signature
Deepika. S	IT19148632	

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

…….. ……

Signature of the supervisor: Date

(Ms. Archchana Kugathasan)

Abstract

University students always find a lot of issues related to administration. They find it hard to get a solution to their issues, doubts, or problems University administrations also find difficulties to clarify all these because of time limitation and the higher rate of student count. The communication is affected because of this pandemic situation [1]. Students are unable to contact the administration via phone call because of this. Most of the students are using email as their communication platform and most of the emails never gets a reply.

Chatbots are being used in many fields including Education, Marketing, and Healthcare in the last few years. Currently there is an increase of chat-bots for e-learning to support students learning [2]. This study aims to determine how the above mentioned problems can be minimized by having a trained chat-bot. Specifically, it helps the students to clarify their doubts by asking with their preferred language (English, Tamil or English-Tamil code-mix) and getting a reply in English.

The goal of this research is to create a chat bot application to get clarification about the administrative issues in a university and provide a multi-language option to make communication easier. Having a chat-bot with Machine learning and Natural Language Processing (NLP) helps university students and administration to get and provide a better solution for issues. The proposed system will help to clarify students issues efficiently with the help of a trained system.

Keywords: Machine Learning, Natural Language Processing, Chat-bot, multi-language

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Background & Literature survey

Unable to get proper solution or answer to a question in one of the common problems' university students are facing. One of the reasons for this is, as the number of students in a university is high and institutions do not have enough time to clarify all the questions [4]. So, some students never get reply for the questions they have asked. As shown in the Figure 1.1 more than 85% of the students do not get reply for their questions all the time. The other 14% of the students who got reply form the administration also does not get a confident solution.

As according to the reference university students are facing more than 80% of stress because of several factors [3]. If university administration does not send reply on time, students will get some more stress because of not finding a solution. As show in in Figure 1.2, even if the students get a reply with a solution/ answer it is not clear enough or it in not helping to solve the doubt/ issue they are having. Only approximately 26% of the students are satisfied with the provided solution.

These are many reasons that a student gets wrong information. If the student does not as the question clearly and the administrator do not understand the question well or the answer provided by the administrator is not clear enough to the student. One of the main reasons for the above problem is communication. Sri Lankan students faces difficulties in English communication because it is taught as second language. This communication issue can lead to wrong information transfer. According to Figure 1.3, in the COVID-19 pandemic time students faced a lot of challenges to contact the university administration.

To provide a solution to this developing a chat-bot system to clear students administrative related issues will be a time saver and effective for students to get accurate reply on time. Having a chat-bot with multiple language support will help the students and administration to manage the need of students. They can ask the questions they have with their preferred language, so that the clearness of the question will be easily passed among the student and the chat-bot. And students can ask multiple questions until they find a solution.

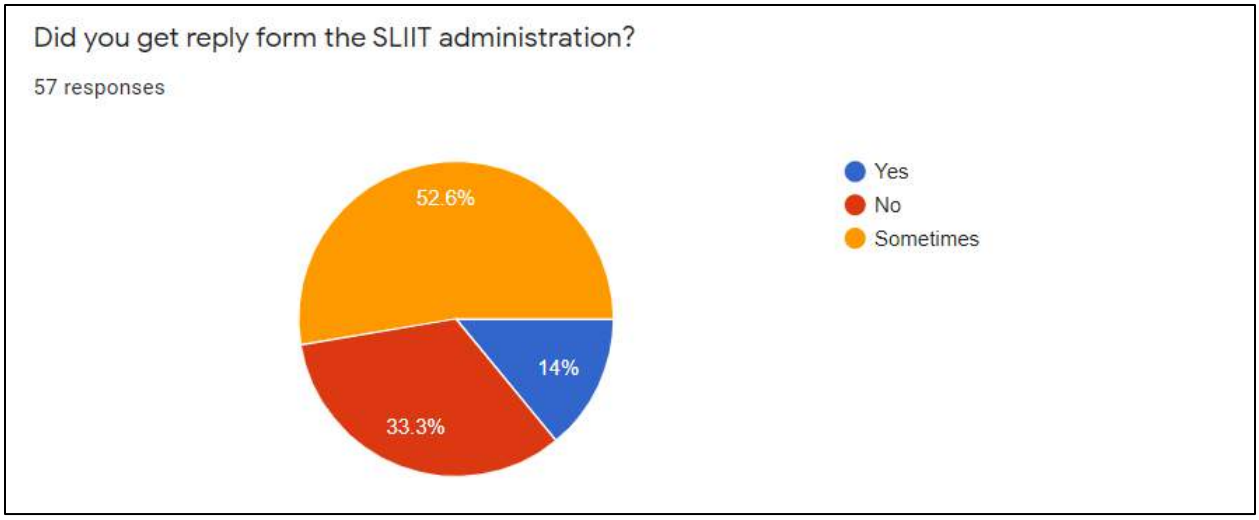

Figure 1. 1 - Summary of responses to know did students receive reply form university administration

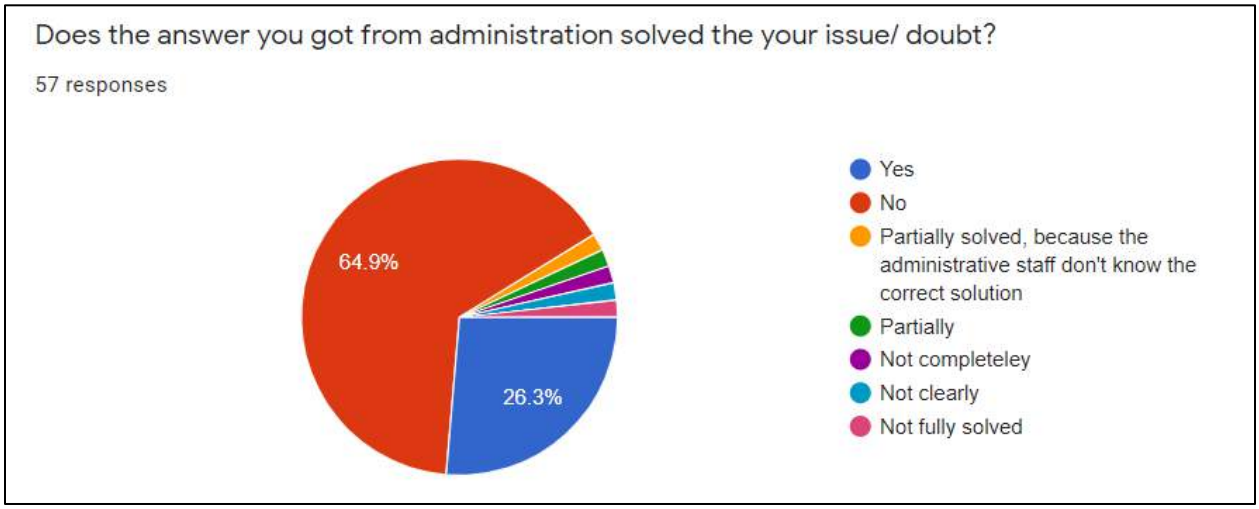

Figure 1. 2 - Summary of how the answer form university administration solved students' issues

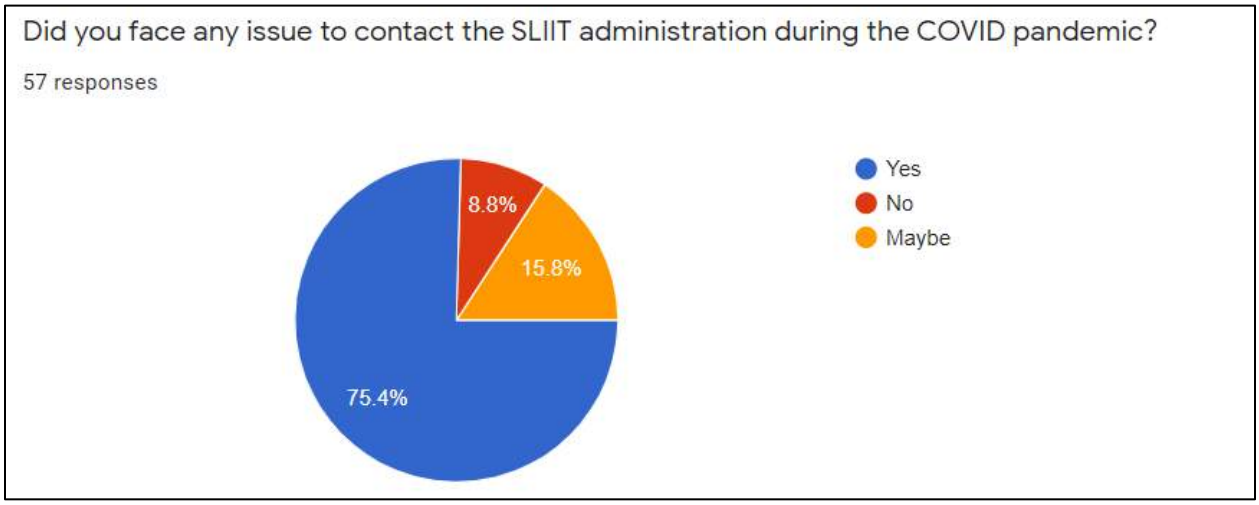

Figure 1. 1 - Summary of responses to know if students faced difficulties to contact administration during COVID pandemic

Many researches have been conducted to create a better chat-bot for university students. According to many authors students are having a chatbot will make students life easier, because they can save a lot of time [1], [2]. According to a research made during the pandemic time, the authors mentioned the interaction between the student and teacher become low and the students find difficulties to get solutions for their doubts [1]. The teachers also found difficulties to answer all the questions asked by each student. Authors have implemented the system based on both voice and text. Having multiple options to ask questions will make the user input easier [1]. In another research also, authors implemented a chatbot which allow both audio and text user input [4].

Normally a student will be unable to contact the administration all the time. There can be holidays and working days issues. But having a chat-bot will help to make the university administration 24/7 [4]. To check the accuracy of the bot many authors have asked the same queries in different ways by changing the working, sentence order and irrelevant orders [4], [5]. Bayu Setiaji and Ferry Wahu Wibowo developed a text based chatbot for Indonesian language [7]. The tokenization, preprocessing and pattern match they have used is similar as used by Guruswami Hiremath, Aishwarya Hajare, Priyanka Bhosale, Rasika Nanaware, and Dr.K.S. Wagh [6].

Developing a chat-bot with providing accurate answers for educational system is important, because students need to get a proper solution for their doubts without wasting their time. A research, which is related to educational domain chat-bots collected around 1500 questions and responsive answers form an educational organization [8]. Having a large amount of data set will help to maintain the accuracy of the answer provide by the chat-bot. In the research made by Sangeetha Kumari, Zaid Naikwadi, Akshay Akole, and Purushottam Darshanjar, the same query is asked in different form by changing the wordings and adding special characters to increase the accuracy of the chat-bot [4].

As per the above-mentioned readings it is crystal clear that the idea of chat-bot with multiple language support will be helpful to university students and university administration. Although there are many researches conducted related to educational chat-bot, most of them don’t have enough accuracy and enough functionality. Also, there are no researches prevailing that have multiple language support for users to communicate with the chat-bot.

1.2 Research Gap

There are a lot of research conducted related to chat-bot for education system where students can get benefit by resolving their issues and save time. But most of those are developed between students and lecturers to ask doubts in a specific subject [1], [2], [8]. But having a platform to ask questions other than modules or subjects is not much researched.

In the research conducted by Sangeeta Kumari, Zaid Naikwadi, Akshay Akole and Purushottam Darshankar, they have implemented a chat-bot for both students and parents to clear the doubts they have related to the university administration [4]. And another research also designed to solve administrative issues using chat-bot [6].

But these researches are mainly focusing on one language. Mostly all the researches are using English as their communication language or one other specific language. The research conducted by Bayu Setiaji and Ferry Wahu Wibowo is developed for Indonesian language only [7]. When it comes to a university, there can be a lot of different students who speak different languages. Having only one specific language for communication is a drawback in these researches.

Only one product is released for real users to test the application to identify the success or failure among them. Other products have not been tested with users.

Products	Designed for university students	Designed to resolve administrative issues	Multiple language selection	Released for real users
Research A [1]	✓	×	×	×
Research B [2]	✓	×	×	✓
Research C [4]	✓	✓	×	×
Research D [6]	✓	✓	×	×
Research E [8]	✓	×	×	×
Escort	✓	✓	✓	✓

Table 1. 1 - Comparison of former researches

The proposed system Escort is designed to overcome the above limitations other researches have. Escort is mainly focusing on the issues university students are having related to university administration. Also, the multi-language selection is also provided where students can choose their preferred language form the provided options. By releasing the final product to real users, we can identify how well the system is useful and what are the improvements or enhancements can be done in the future.

1.3 Research Problem

Students and university administration need to manage their time to be productive. Also, students need to be clarified if they have any doubts related to university administration. From the conducted survey we can identify that most of the students are facing a lot of issues while clarifying the doubts from university administration.

According to Figure 1.4 we can see that approximately 60% of the students think that that the university administration misunderstood their question because of language issue. Understanding the question incorrectly will lead to wrong information sharing and it will be very harmful to both students and administration.

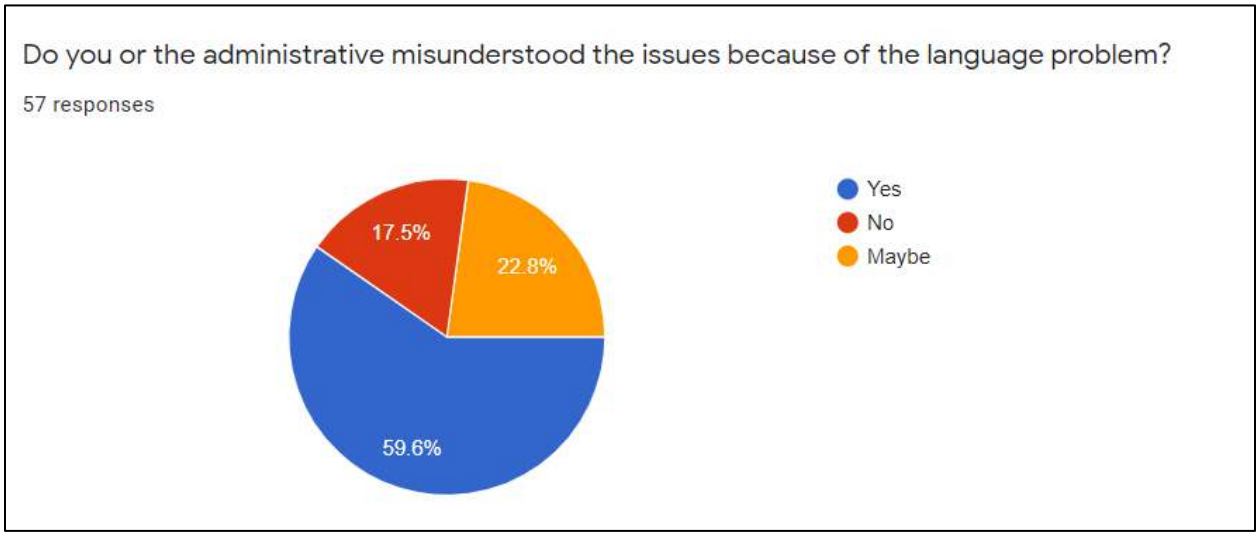

Figure 1. 2 - Summary of response to identify language issue

According to Figure 1.5, most of the students use email and phone as communication media to contact university administration. Because of a lot of emails and calls, university administration is also unable to response to all of them. This leads to having unresolved doubts or questions among university students.

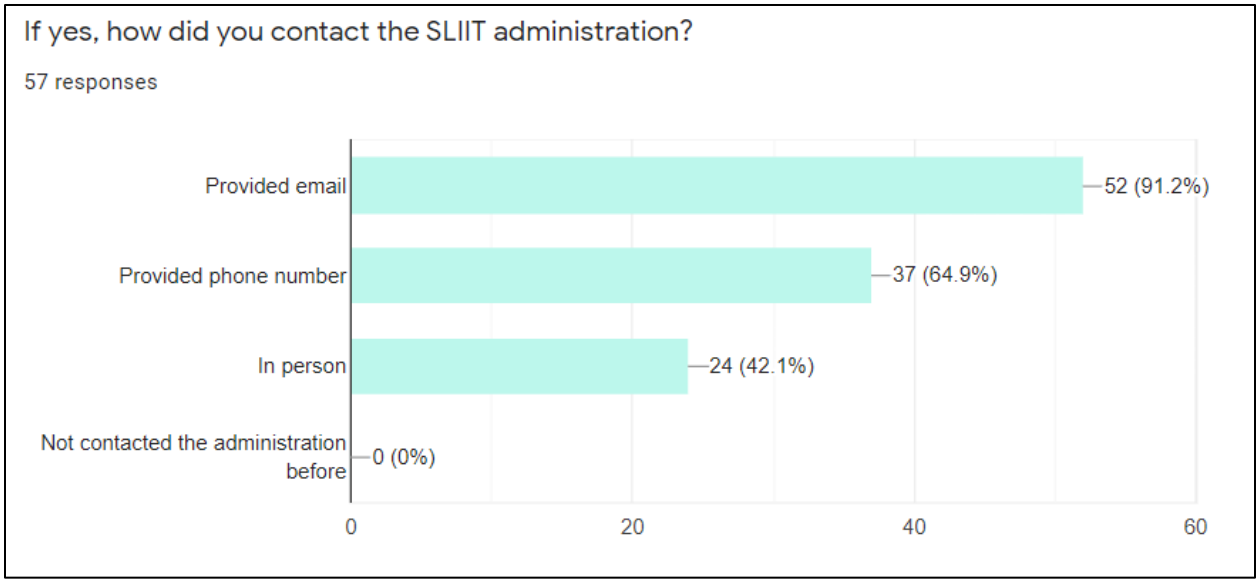

Figure 1. 3 - Summary of response to know communication platform

Also, using a chatbot will reduce the waiting time for asking simple doubts [4]. During the survey, students also expressed that having a chat-bot will help them to reduce time and get accurate response (Figure 1.6).

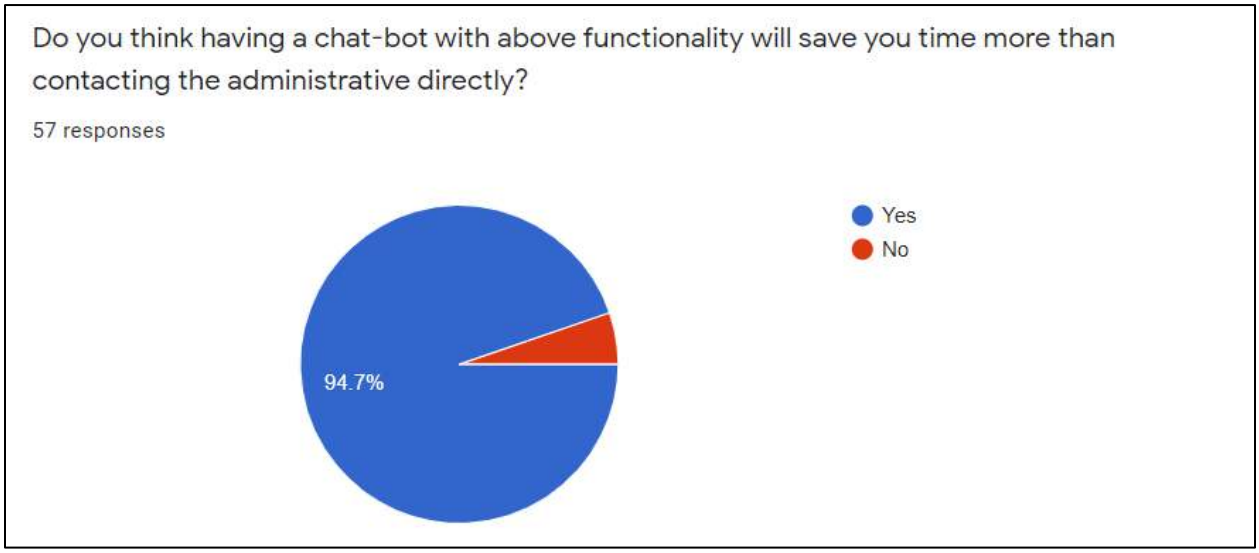

Figure 1. 4 - Summary of response to know if chat-bot is useful among students

Electronic devices and trending technologies had become the sensation of young generation, and most of them prefer to have online platform to solve every issue. This research proposes 'Escort' as a solution to support young generation university students.

2 Objectives

2.1 Main Objectives

The main objective of implementing Escort administrative chat-bot system is to make university life easier for both students and university administration, by making the communication between them more effective. Given the fact the administration is available on-campus will help to achieve the above easily. But still students are facing issues to communicate with the administrator because of language issue. The questions asked by the student is not clear for the administrator or the solution provided by the administrator is not clear to students. Which lead to wrong information passing.

Since the world is adopting new technologies like Machine learning, NLP, and AI it is better to have a similar software solution to overcome the daily life issues. Because of the number of students in a university, the administration will not be able to solve each student's administrative related issues individually. Building a chat-bot, which can give a proper solution for students' issue will be easier and effective.

Students who are unable to explain the issues/ doubts they are having in a specific language are getting un-clear solutions. However, providing a chat-bot with multiple language support like English, Tamil, and English-Tamil code-mix will help them to choose the language they are comfortable with. The methodology of administrative issue related chat-bot is explained further in the document.

2.2 Specific Objectives

In order to achieve the main objective, the specific objectives that needs to be clarified is as follows,

1. Identify the main categories of administration which students are having more issues/ doubts or needed more clarifications.
2. Get proper data set which will give an accurate solution.
3. Implement a chat-bot which will support multiple language input from students (English, Tamil and English-Tamil code-mix)
4. Implement a website which can be accessed by all technical devices.

3 Methodology

The proposed “Escort” administration related chat-bot has the capability of,

- Allow students to choose their preferred language to ask questions.
- Identify the issue student asked in their preferred language.
- Give an accurate solution for the asked question.

The system will identify the question asked by the student in the given language choice and it will preprocess the sentence and identify the proper solution. Students can access the system via a web application and there they can access the cat-bot. They can choose their preferred language mode (English, Tamil, English-Tamil code-mix) and start asking questions related to administration.

The system will identify the input and prepare the text data by commencing preprocessing, [10] which will use various steps like removing punctuations, tokenization and removing URLs according to our dataset. Text preprocessing will happen using four sub parts which are,

1. Sentence splitting,
2. Tokenization,
3. POS tagging
4. Stemming.

First the paragraph user entered will be divide into sentences using sentence splitting. The tokenization will be used to divide each sentence into words/ tokens. Each word/ token will be gone under POS tagging to categorize them as verbs, nouns and pronouns. After POS tagging, stemming will be used to remove affixes and suffixes.

Once the texts are prepared the prebuild model will identify them and analysis those with the knowledge resource we have. Then the system will reply with a suitable answer. Students can continue this process until they have received a proper solution for the questions they have.

3.1 System Architecture

The system architecture is shown in the Figure 3.1

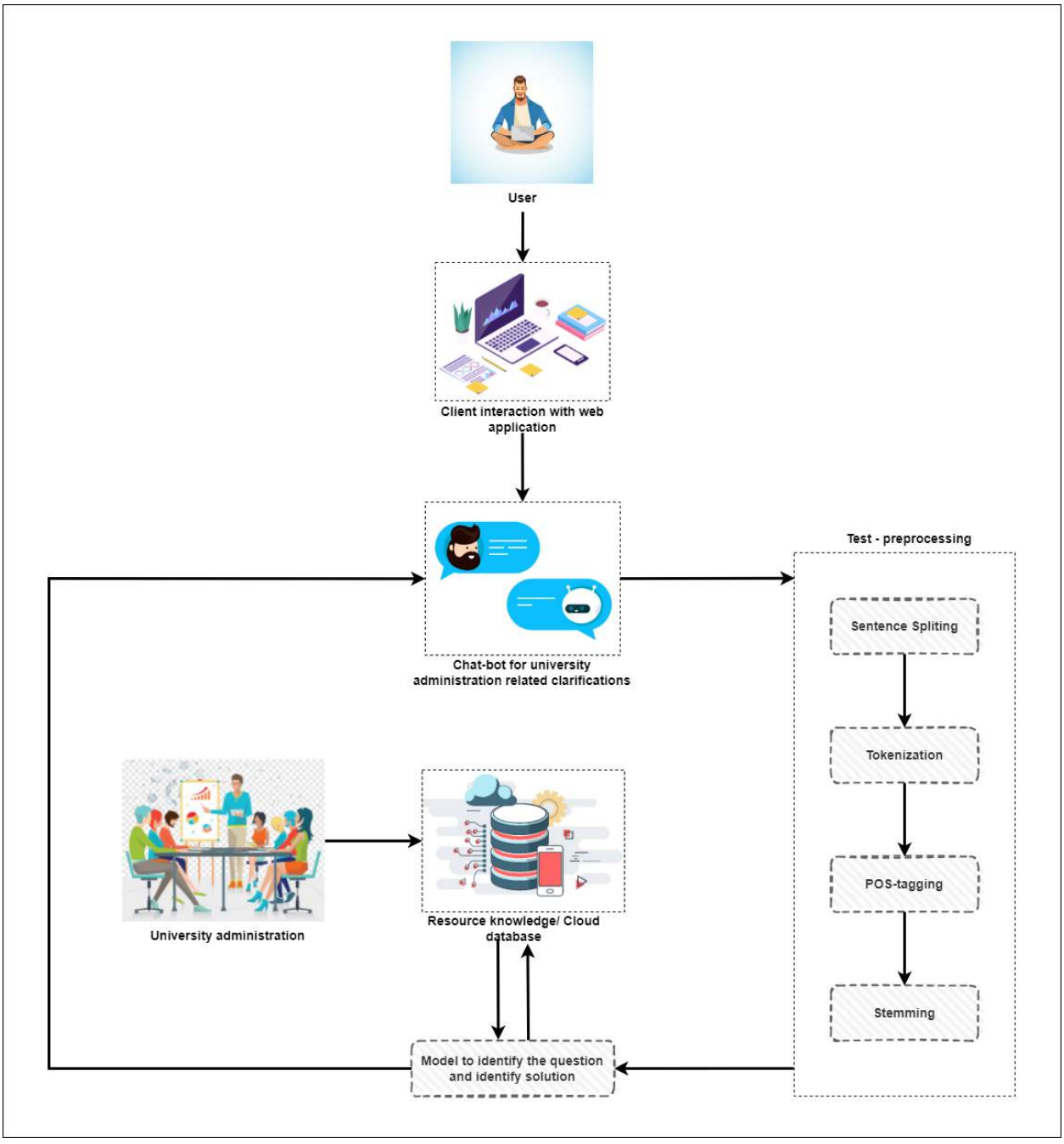

Figure 3. 1 - High Level Architectural Diagram

3.2 Software Solution

The agile methodology has been chosen as the software development life cycle and scrum will be the methodology that will be followed under agile. Scrum is a framework that helps teams work together and adjust the project with several phases. In each phase we can easily improve our project. This framework will help to manage the complexity of the research project. The Figure 3.2 clearly shows that the framework will help us to manage the project with multiple sprints.

Figure 3. 2 - Agile methodology

1. Requirement gathering and analysis

   The first phase of software development life cycle is requirement gathering. The time we spent on the requirement gathering via questionnaire has helped a lot to identify what kind of administrative issues students are having, how the administration is providing solutions and how that solution satisfied the students.

2. Feasibility Study

   This phase considers all the relevant possibilities including technical, financial, operational and the time to complete the project. From this phase we were able to identify answers for the followings,

   - Whether the project can be done
   - Whether the final product will benefit the users
   - What are the alternative solutions?
   - Whether there is a preferred alternative

3. Design and Development

   In this phase we were able to transform the software requirements into architectural design and low-level design. Once we develop a low-level design it will be easier to develop the software. Also, we can collect the relevant data set and identify the software to store these data and then start the development of the project.

4. Testing

   Once the development of a sprint is finished, we will start doing the relevant testing of the developed software. This process will help us to identify the issues we have and solve them in the next phase.

5. Deployment

   After completing all the sprints and the final testing we can deploy the project into cloud web services. As machine learning models are easier to deploy in web services, we will use a relevant cloud web service.

4 Project Requirements

4.1 Functional Requirements

1. Analysis students' questions in multiple language
2. Analysis on solution/ answer for student's question

4.2 Non-functional Requirements

1. Give accurate solution/ answer for students' questions.
2. Secure students' questions and maintain privacy.
3. Better performing website.
4. Reliability of the software

4.3 User requirements

1. PC, Laptop or Portable electronic devices
2. Internet connection

4.4 System Requirements

1. Windows 10 or above
2. Intel Core i7 Processor
3. 8 GB RAM
4. Software to develop the system

5 Gantt chart

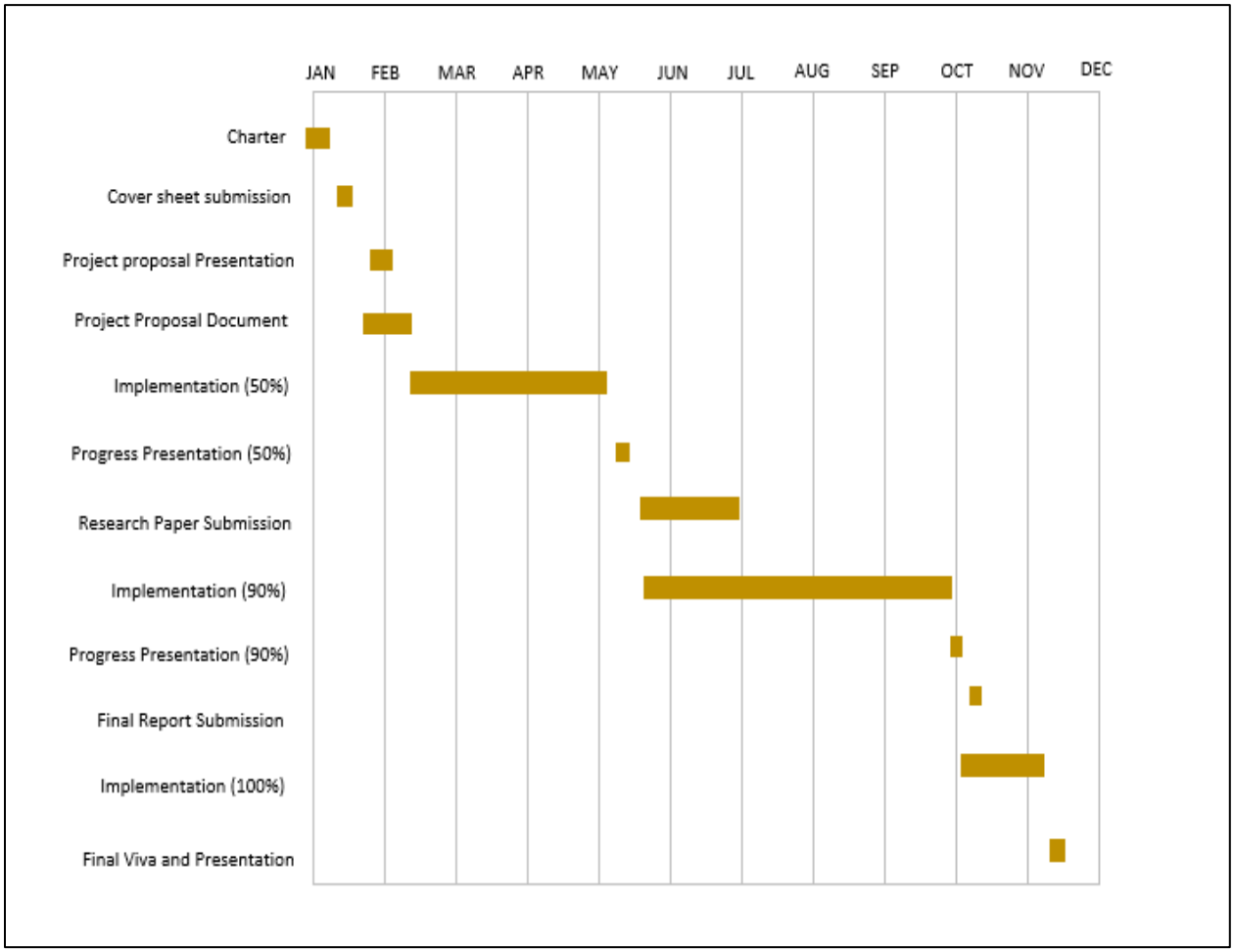

Figure 5. 1- Gantt Chart

6 Work Breakdown Chart

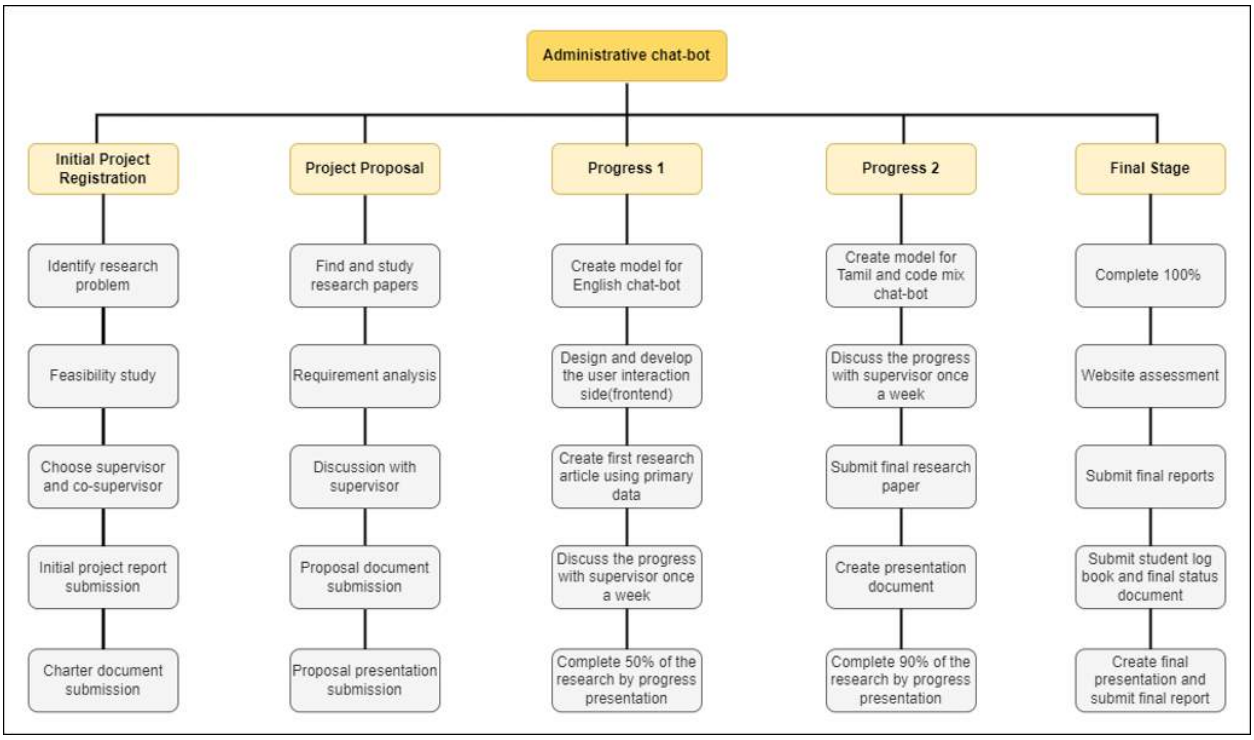

Figure 6. 1 - Work Breakdown Chart

7 Description of Personal Facilities

Member	Component	Task
Deepika. S	A chat-bot application to contact university administration about issues/ doubts or clarifications	Requirement gathering and analysis System design/ UI design Database designing Implementation of function • Multiple language selection • Text preprocessing on user input • Accurate input • NLP techniques Developing web application Documentation Testing Presentation

8 Budget and Budget Justification

The budget allocation for various resource types used in the research is given below

Resource type	Amount (LKR)
Preparation of report and printing	Rs. 1500
Internet usage for researching	Rs. 5000
Domain Name Registration (annual)	Rs. 1500
Hosting (annual)	Rs. 10000
Other costs	Rs. 2000
Total	**Rs. 20000**

Table 8. 1 - Budget

Reference list

[1] E. Kasthuri and Dr.S. Balaji, "A Chatbot for Changing Lifestyle in Education", Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks, ICICV 2021.

[2] Fabio Clarizia, Francesco Colace, Marco Lombardi, Francesco Pascale, and Domenico Santaniello, "Chatbot: An Education Support System for Student", Springer Nature Switzerland, AG 2018.

[3] Saba Asif, Azka Mudassar, Talala Zainab Shahzad, Mobeen Raouf, and Tehmina Pervaiz, "Frequency of depression, anxiety and stress among university students", [online] Available at: Frequency of depression, anxiety and stress among university students (nih.gov) [2020].

[4] Sangeeta Kumari, Zaid Naikwadi, Akshay Akole, and Purushottam Darshankar. "Enhancing College Chat Bot Assistant with the Help of Richer Human Computer Interaction and Speech Recognition", International Conference on Electronics and Sustainable Communication System (ICESC), pp.427-433. IEEE, 2020.

[5] György Molnár and Szűts Zoltán, "The Role of Chatbots in Formal Education", IEEE 16th International Symposium on Intelligent Systems and Informatics, 2018.

[6] Guruswami Hiremath, Aishwarya Hajare, Priyanka Bhosale, Rasika Nanaware, and Dr.K.S. Wagh, "Chatbot for education system", International Journal of Advance Research, Ideas and Innovations in Technology, 2020.

[7] Bayu Setiaji and Ferry Wahyu Wibowo, "Chatbot Using A Knowledge in Database", 2016 7th International Conference on Intelligent Systems, Modelling and Simulation.

[8] Anupam Mondal, Monalisa Dey, Dipankar Das, Sachit Nagpal, Kevin Garda, "Chatbot: An automated conversation system for the educational domain", IEEE, 2018.

[9] Deepanshi, "Text Preprocessing in NLP with Python codes", [online] Available at: https://www.analyticsvidhya.com/blog/2021/06/text-preprocessing-in-nlp-with-python-codes/ [2021].

[10] Maria Joao Pereira, Luisa Coheur, Pedro Fialho, Ricardo Ribeiro, "Chatbots' Greetings to Human-Computer Communication", 2016.