ESCORT – NATURAL LANGUAGE PROCESSING BASED UNIVERSITY STUDENTS' GUIDANCE SYSTEM

Project ID: 2022-179

Final Report - Individual

Deepika Srinivasan
IT19148632

B.Sc. (Hons) Degree in Information Technology
Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

September 2022

ESCORT – A CHATBOT APPLICATION TO CONTACT UNIVERSITY ADMINISTRATION ABOUT ISSUES

Project ID: 2022-179

Final Report - Individual

Deepika Srinivasan
IT19148632

Supervisor
Ms. Hansi De Silva

Co-Supervisor
Mr. Dharshana Kasthurirathna

Department of Software Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

September 2022

DECLARATION

I declare that this is my own work, and this proposal does not incorporate without acknowledgment 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 acknowledgment is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce dan distribute my dissertation, in whole or in part print, electronic or other medium. I retrain the right to use this content in whole or part in future works (such as articles, or books).

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. Hansi De Silva)

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 it challenging to clarify all these because of time limitations and the higher rate of student count. Communication is affected because of this pandemic situation [1]. Students are unable to contact the administration via phone call because of this. Most students use email as their communication platform, and most emails never get 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 chatbots 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 chatbot. Specifically, it helps the students to clarify their doubts by asking in their preferred language (English, Tamil, or Thanglish) and getting a reply in the same language.

This research aims to create a chatbot application to clarify administrative issues in a university and provide a multi-language option to make communication easier. Having a chatbot 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

ACKNOWLEDGMENT

First of all, my deepest gratitude should offer to our supervisor Ms. Hansi De Silva, the main person who gave great support to this research. Without her proper guidance and intensity, this research would not have been a success. Also, our co-supervisor Mr. Dharshana Kasthurirathna was the other person who guided us and helped us to achieve our research goals. Then I would like to offer my appreciative thanks to my research team for supporting me and finishing the research successfully. Finally, I would like to acknowledge all the lecturers and colleagues of the Sri Lanka Institute of Information Technology who have supported us from the beginning of this research. Without any of their single guidance and suggestions, it would have been complicated to achieve our targets.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBERIVIATIONS

NLP	Natural Language Processing
NLTK	Natural Language Took Kit
ReLU	Rectified Linear Unit
PKL	Pickle file
SLIIT	Sri Lanka Institute of Information Technology
API	Application Programming Interface

1 INTRODUCTION

Most university students are facing difficulties to get solutions to administration-related issues. Students have numerous doubts and questions regarding payment, sports, exams, parking, etc. But most of the students don't get reply form administration on time. The main reasons for these are, as the number of students in a university is high and institutions do not have enough time to clarify all the questions in a limited time range [2]. There are many reasons for a student to get incorrect information and late replies such as if the student does not ask the question clearly and the administrator does not understand the question correctly or the answer provided by the administrator is not clear enough for the student. According to the reference, more than 80% of university students face stress because of several factors [4]. A chatbot system is developed to provide solutions to administration-related issues. The model for the chatbot is developed with NLTK. Tokenization and lemmatization are used under text-preprocessing. This will be an effective time-saver for students to come to an accurate reply at any time. Having a chatbot with multiple languages will support the students and administration in managing the needs of students. From this, the students will be delivered the correct and accurate answers in the selected language until finding a solution.

1.1 Background & Literature survey

Unable to get a proper solution or answer to a question is 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 [5]. So, some students never get replies for the questions they have asked. As shown in Figure 1 more than 85% of the students do not get replies to their questions all the time. The other 14% of the students who got reply form the administration also does not get a confident solution.

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 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 question/ 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 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 chatbot 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.

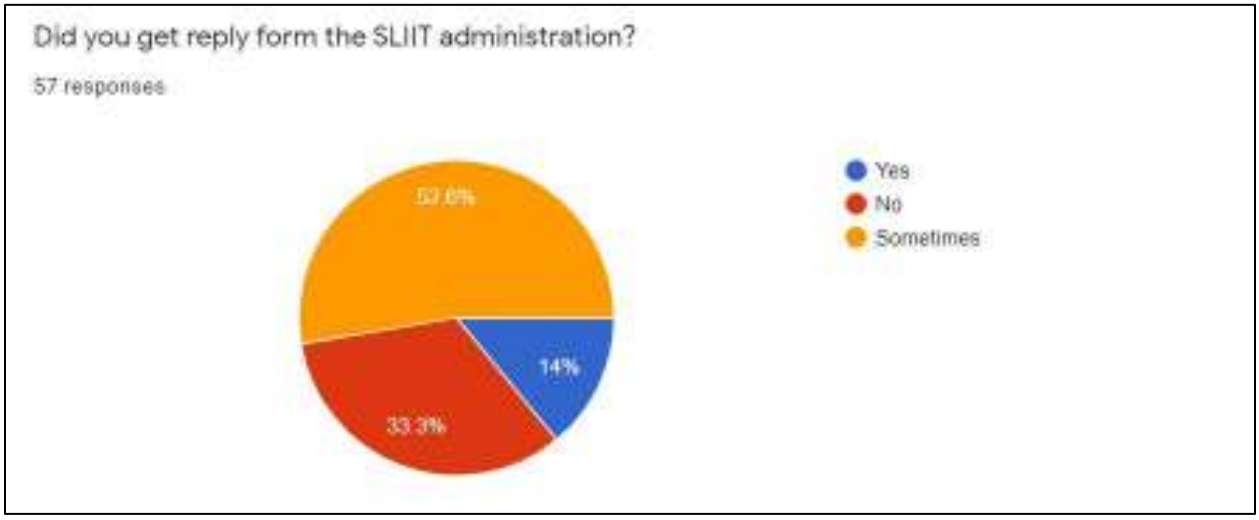


Figure 1: Summary of responses to know did students receive reply form university administration

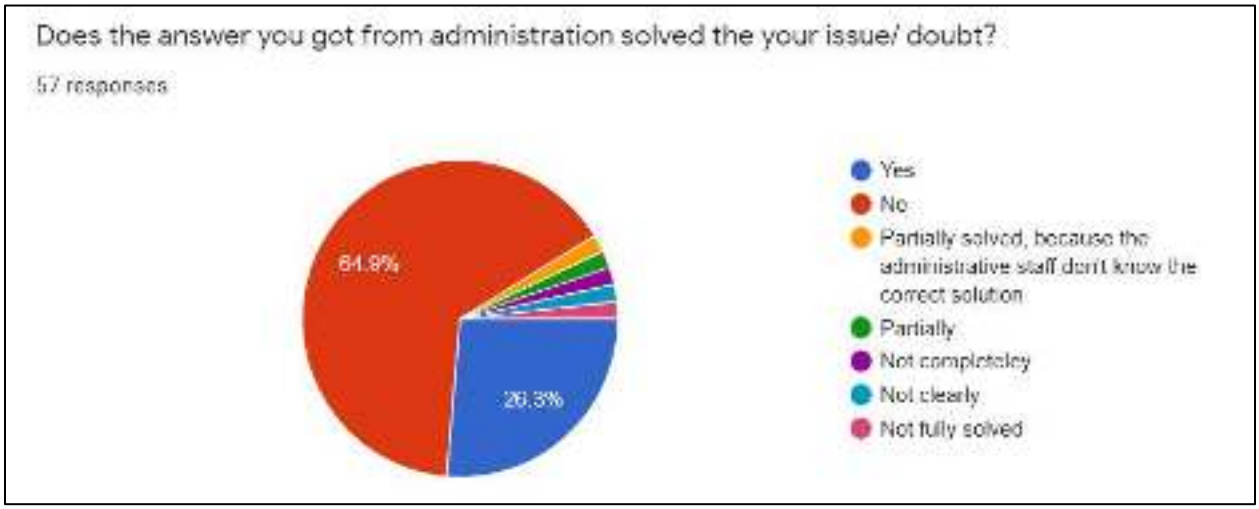


Figure 2: Summary of how the answer from university administration solved students' issues

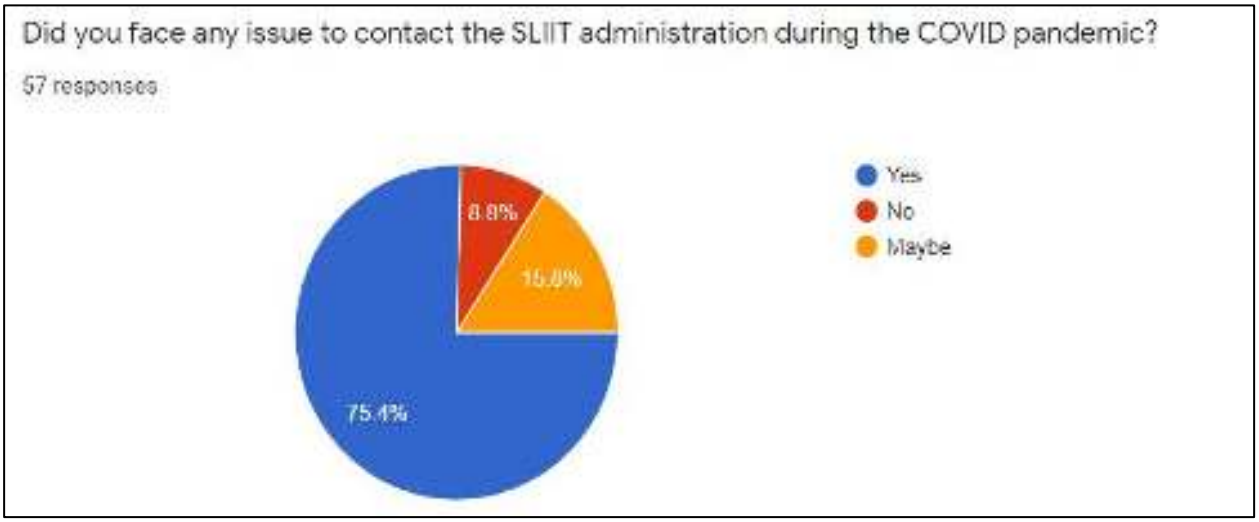


Figure 3: Summary of responses to know if students faced difficulties to contact administration during the COVID pandemic

Many kind of research have been conducted to create a better chatbot for university students. According to many authors 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 it challenging to answer all the questions asked by each student. Therefore, the 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 [5].

Usually, 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 [5]. To check the bot's accuracy many authors have asked the same queries in different ways by changing the working, sentence order and irrelevant orders [5], [6]. Bayu Setiaji and Ferry Wahu Wibowo developed a text based chatbot for Indonesian language [8]. 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 [7].

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 chatbots collected around 1500 questions and responsive answers form an educational organization [9]. 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 [5].

As per the above-mentioned readings it is clear that the idea of chatbot 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 do not have enough accuracy and enough functionality. Also, there are no researches prevailing that have multiple language support for users to communicate with the chatbot.

1.2 Research Gap

Having questions to university administration is common in every universities. Students at a university, lecturers and parents also have many questions that they would like to clarify with the university administration. A university administration is tried to contact by phone, email, and in person. But the administration cannot answer all the questions on time and accurate. Also, there are a lot of miscommunications happening because of the language issue. In that respect there is a need for a system to answer all administrative related questions in the language user prefers. For that purpose, Natural Language Processing techniques are used as an approach to create a chatbot with multiple language support.

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], [9]. 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 [5]. And another research also designed to solve administrative issues using chat-bot [7].

But these researches are 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 [8]. 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.

Table 1: Comparison of former researches

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	✓	✓	✓	✓

The proposed system Escort is designed to overcome the above limitations other researches have. Escort is 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.

2. 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 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.

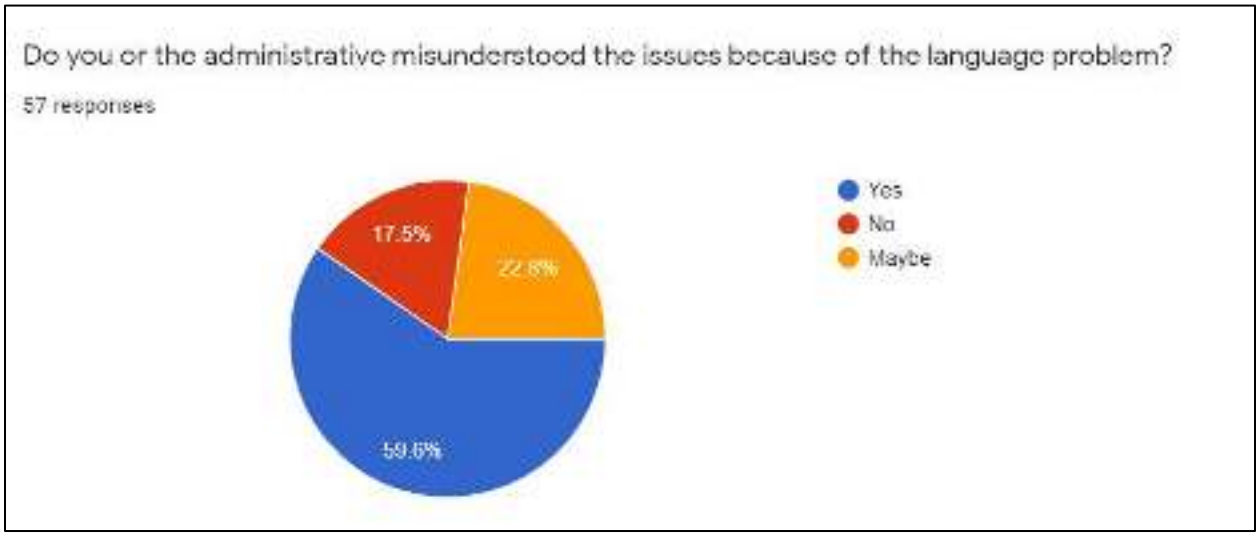


Figure 4: Summary of response to identify language issue

According to Figure 4, 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.

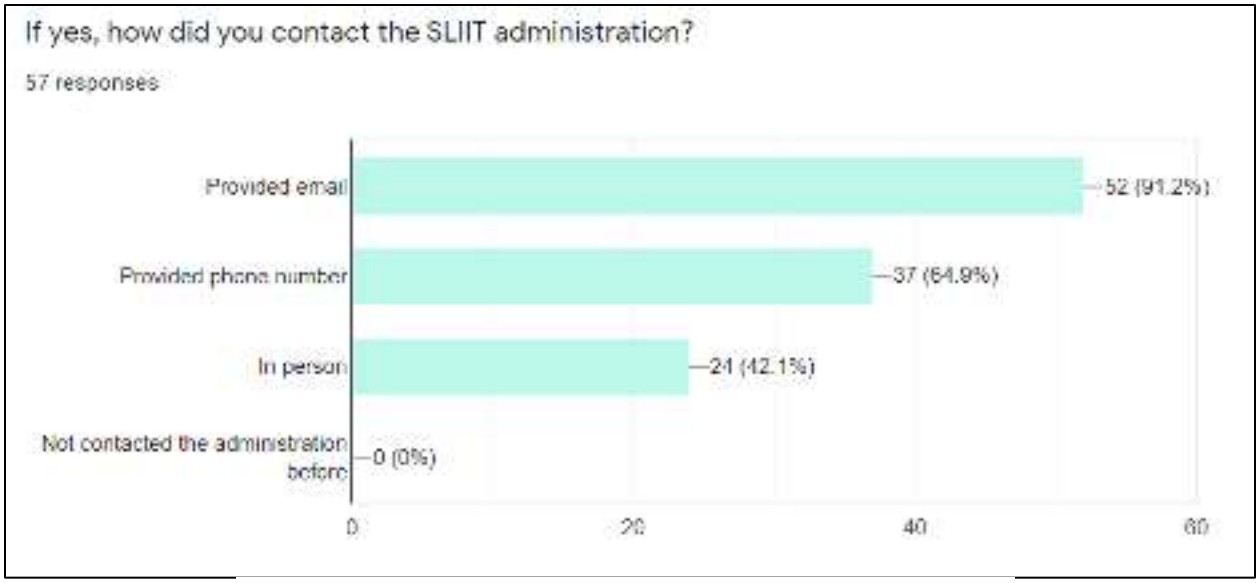


Figure 5: Summary of response to know communication platform

Also, using a chatbot will reduce the waiting time for asking simple doubts [5]. During the survey, students also expressed that having a chat-bot will help them to reduce time and get accurate response (Figure 6).

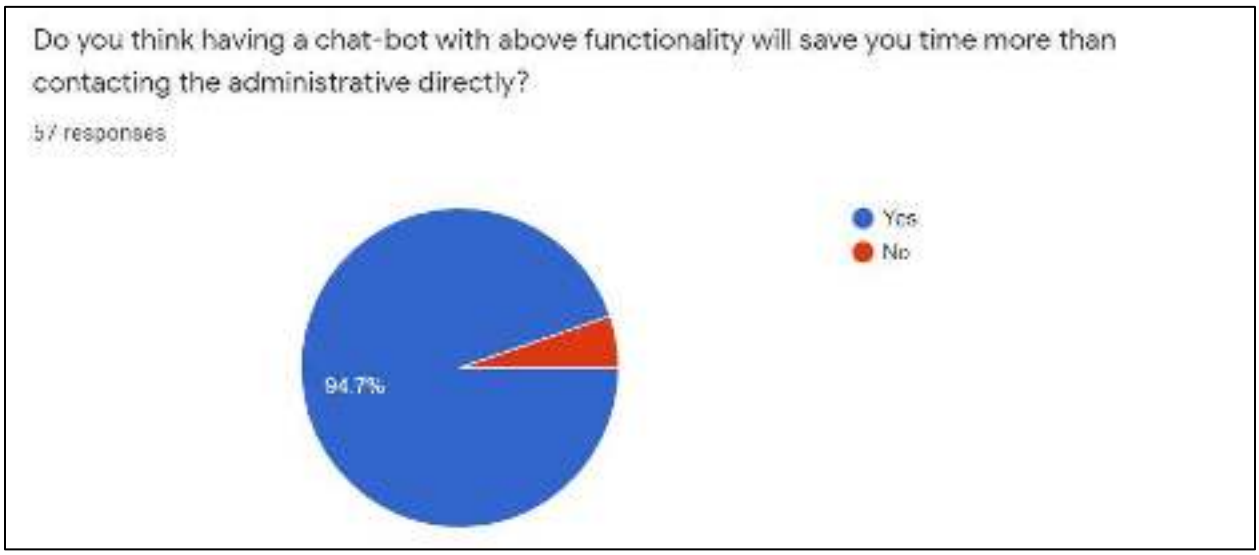


Figure 6: Summary of response to know if chat-bot is useful among students

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.

3 OBJECTIVES

3.1 Main Objectives

The main objective of implementing Escort administrative chatbot 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 questions/ 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 Thanglish will help them to choose the language with which they are comfortable. The methodology of administrative issue related chat-bot is explained further in the document.

3.2 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 Thanglish)
4. Implement a website which can be accessed by all technical devices.

4 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, and Thanglish) and start asking questions related to administration.

The system will identify the input and prepare the text data by commencing preprocessing, [11] which will use various steps like removing punctuations, tokenization and removing URLs according to our dataset.

4.1 Data Collection

The primary goal of collecting requirement is to identify the administration related questions students are having. The questions for the data set are collected for SLIIT students via google questionnaire. The questionnaire had 9 sections, where students can add their questions according to the category. Later the questionnaire was shared among common users, so they can ask their common questions about SLIIT. The total questions received was around seven hundred and the data set has around three thousand questions with answers. The answers for each question are written with the help of SLIIT administration hundred and the data set has around three thousand questions with answers.

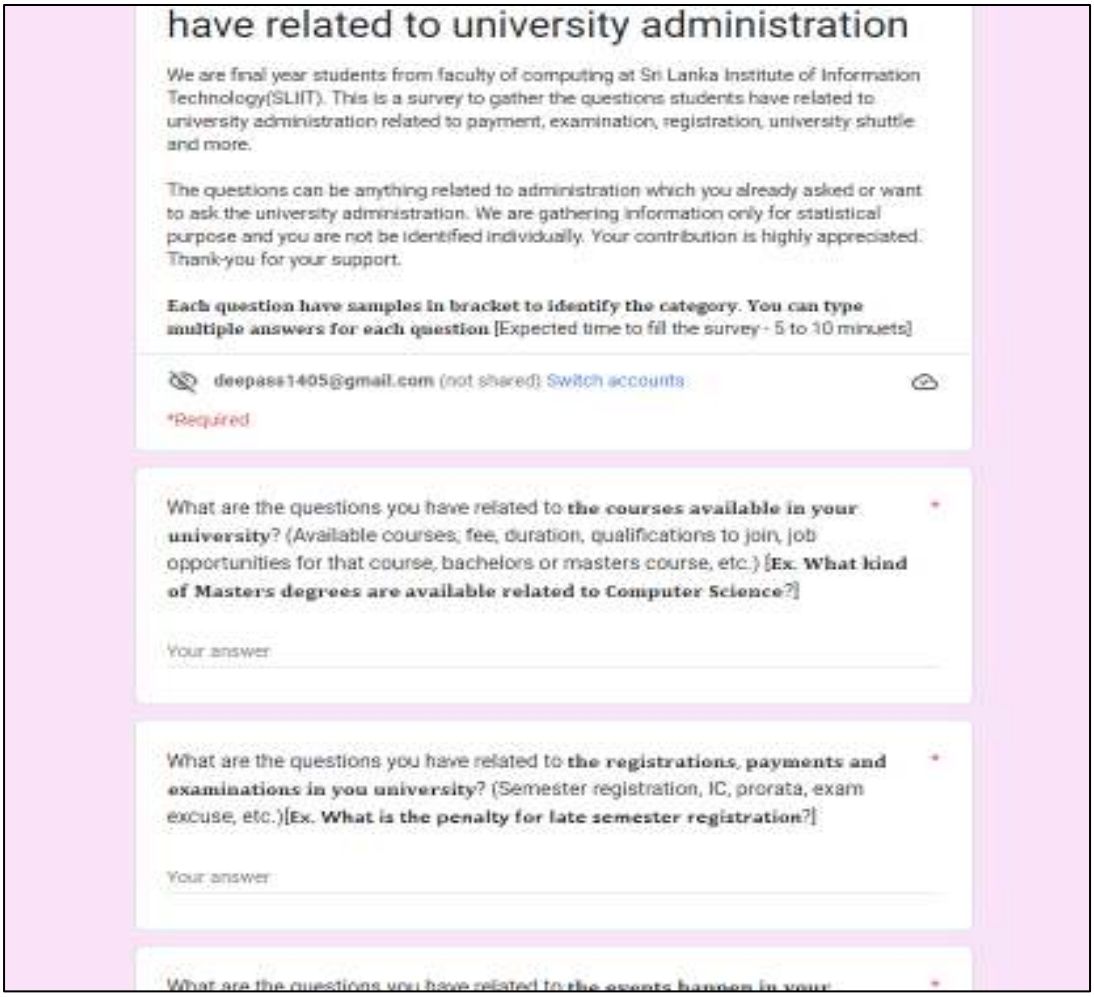


Figure 8: The google form of data collection

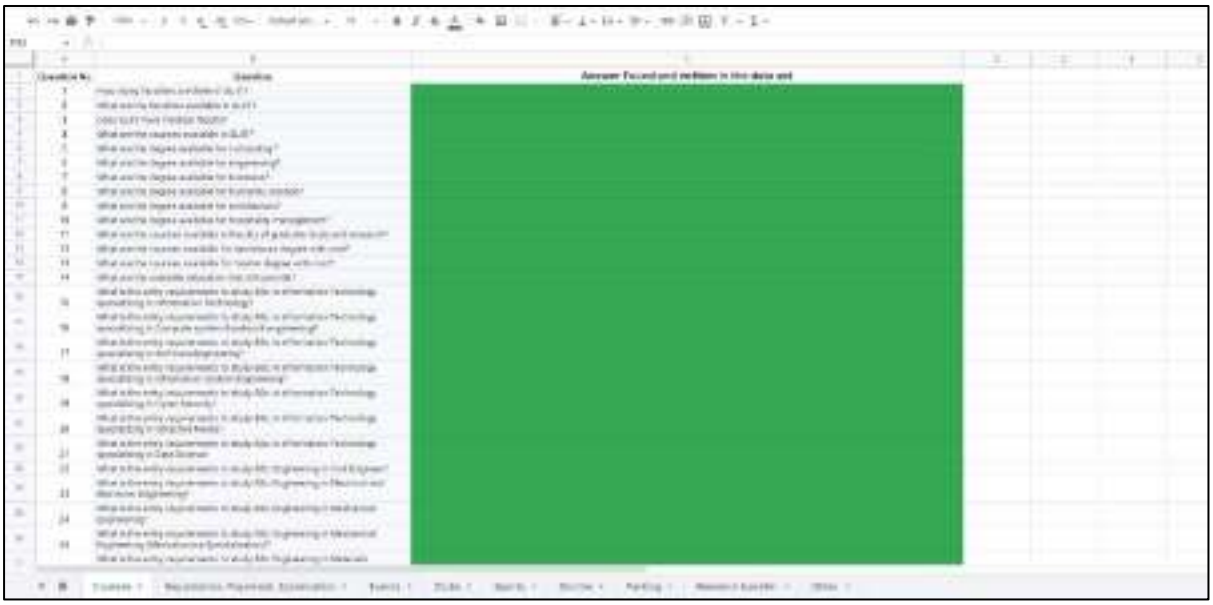


Figure 7: Categories and dataset of each category

4.2 Natural Language Processing techniques and Model Build

A. Text preprocessing

In NLP, text preprocessing is the first step of implementing a model [12].To achieve an accurate chatbot the techniques tokenization, lower casing, stop words removal and lemmatization are being used. Tokenization is a text preprocessing technique which split the sentence into smaller units. Sentence tokenization is used to split the sentences into words. Then lemmatization is used to switch each words into its base root mode. Then the words are converted to lowercase and the stop words have been removed.

```
for intent in intents['intents']:
    for pattern in intent['patterns']:
        stop_words = set(stopwords.words('english'))
        word = nltk.word_tokenize(pattern)
        words.extend(word)
        documents.append((word, intent['tag']))

        if intent['tag'] not in classes:
            classes.append(intent['tag'])

words = [wordLemmatizer.lemmatize(word.lower())
         for word in words if word not in ignore_patterns and stop_words]
```

Figure 9: Implementation of text preprocessing techniques

B. Modal Build

The model is built using the above text preprocessing techniques. The model is developed by using Natural Language Toolkit (NLTK) and Keras library for artificial neural networks. The data set is created as a json file which have the university administrative related questions and responsive answers for them. The data set is called and put into tokenization, ignore patterns, lemmatization, and lowercasing techniques. Once the text preprocessing part is completed the words and their responsive labels are stored in PKL files, which is a file created by pickle that enables object to be serialized to files on disk and deserialized back into the program at run time. The model type being used is Keras sequential which allows to build a model layer by layer. ‘Dense’ layer type is used to create

the layers and for activation function ReLU (Rectified Linear Activation) is being used. For the last layer ‘Softmax’ activation is used. The model is trained for two hundred epoch, which means the entire dataset is passed forward and backward through the neural network two hundred times. Once the training is completed the model can be used to ask questions and get response [13].

```
model = Sequential()
model.add(Dense(128, input_shape=(len(train_x[0]),), activation='relu'))
model.add(Dropout(0.3))
model.add(Dense(64, activation='relu'))
model.add(Dense(32, activation='relu'))
model.add(Dense(16, activation='relu'))
model.add(Dense(len(train_y[0]), activation='softmax'))

model.compile(loss='categorical_crossentropy',
              optimizer='rmsprop', metrics=['accuracy'])
history = model.fit(np.array(train_x), np.array(train_y),
                    epochs=200, batch_size=5, verbose=1, validation_split=0.33, validation_data=(x_val, y_val))
model.save('model/english_administrative_chatbot.h5', history)
```

Figure 10: Implementation of layers and activations

4.3 System Architecture

The system architecture is shown in the Figure 11

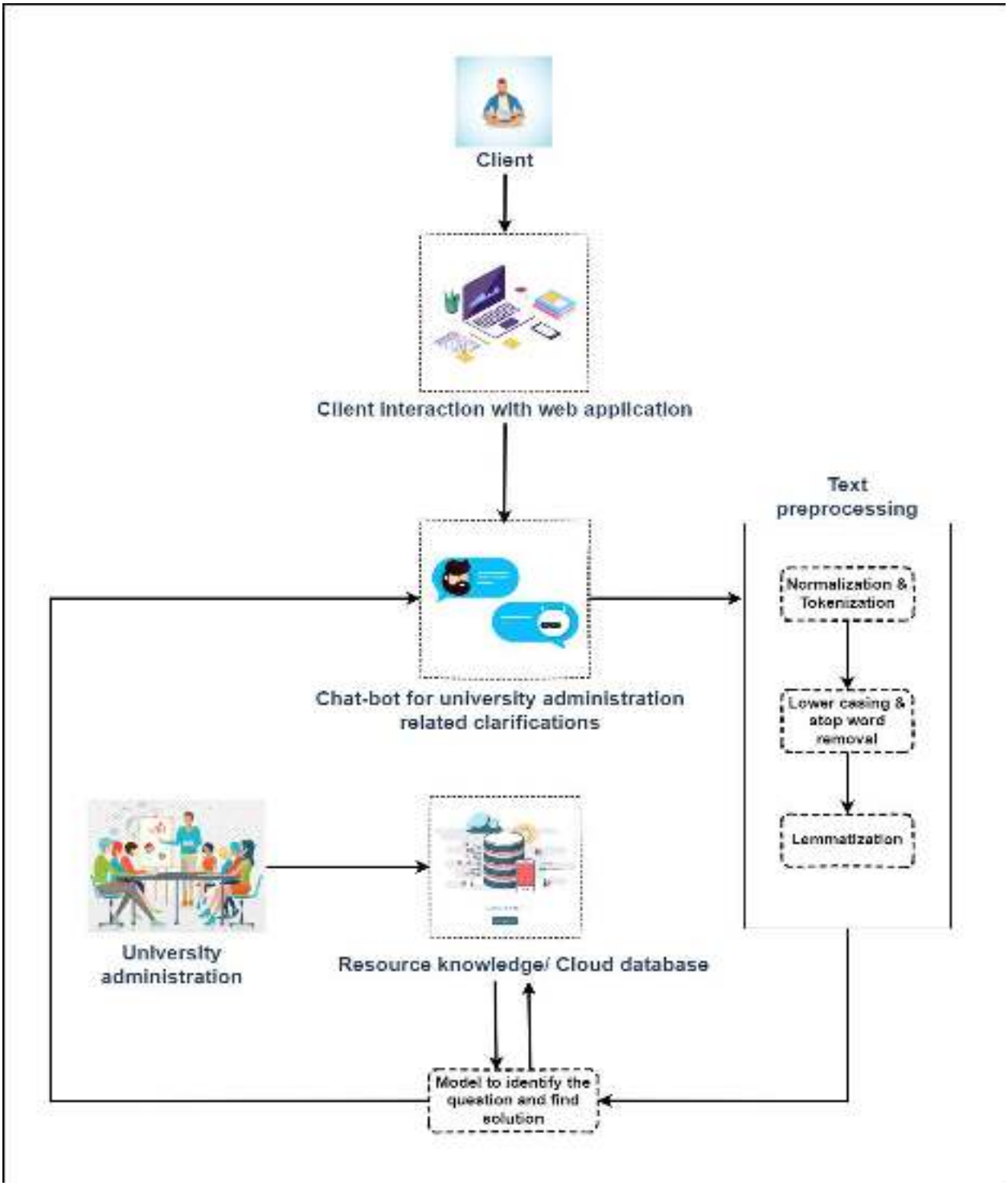


Figure 11: High Level Architectural Diagram

4.4 Commercial aspects of the product

Escort, which is provides a perfect solution for university administration related questions, which many students, lecturers, and parents have. Many people do not have enough time to wait to get answers for their questions and the administration also do not have much time to answer each person's question. Also, the language barrier is leading the communication into difficulties which also lead to work information sharing.

We have implemented a successful and accurate administrative chatbot system, which helps to save a lot of time of both users and university administrators. The system is implemented with multiple language option, which will help to overcome the communication issues. As the system will operate 24/7 users do not have to waste their time on waiting for an answer.

5 IMPLEMENTATION & TESTING

5.1 Implementation

All three language models (English, Tamil, and Thanglish) were implemented in this stage. Each model was trained using more than 2500 data set. As the first procedure, the chatbot model was implemented and trained in python language. NLTK is used to implement the text preprocessing part and Keras sequential method was used to create and train the model. Once the model was created an API as implemented to communicate with frontend.

Frontend was developed using Java Script (React JS framework). Once the user enters the web site, they can choose their preferred language and start asking their questions. All the interfaces were implemented with user-friendliness and responsiveness which user can access the website via any smart device (computer, laptop, mobile phone, and smart television).

The following are some of the interfaces which were implemented for administrative chatbot.

A. Administrative home page
B. English chatbot
C. Tamil chatbot
D. Thanglish chatbot

A. Administrative home page

Administration home page will help students to know what the system does, and it allow users to choose the language they preferred.

Figure 12: Administration chatbot home page

B. English Chatbot

English chatbot page is to ask questions only in English. It has the instructions about the questions users can ask. A chat option is provided in the page where users can easily type their questions and the get response immediately.

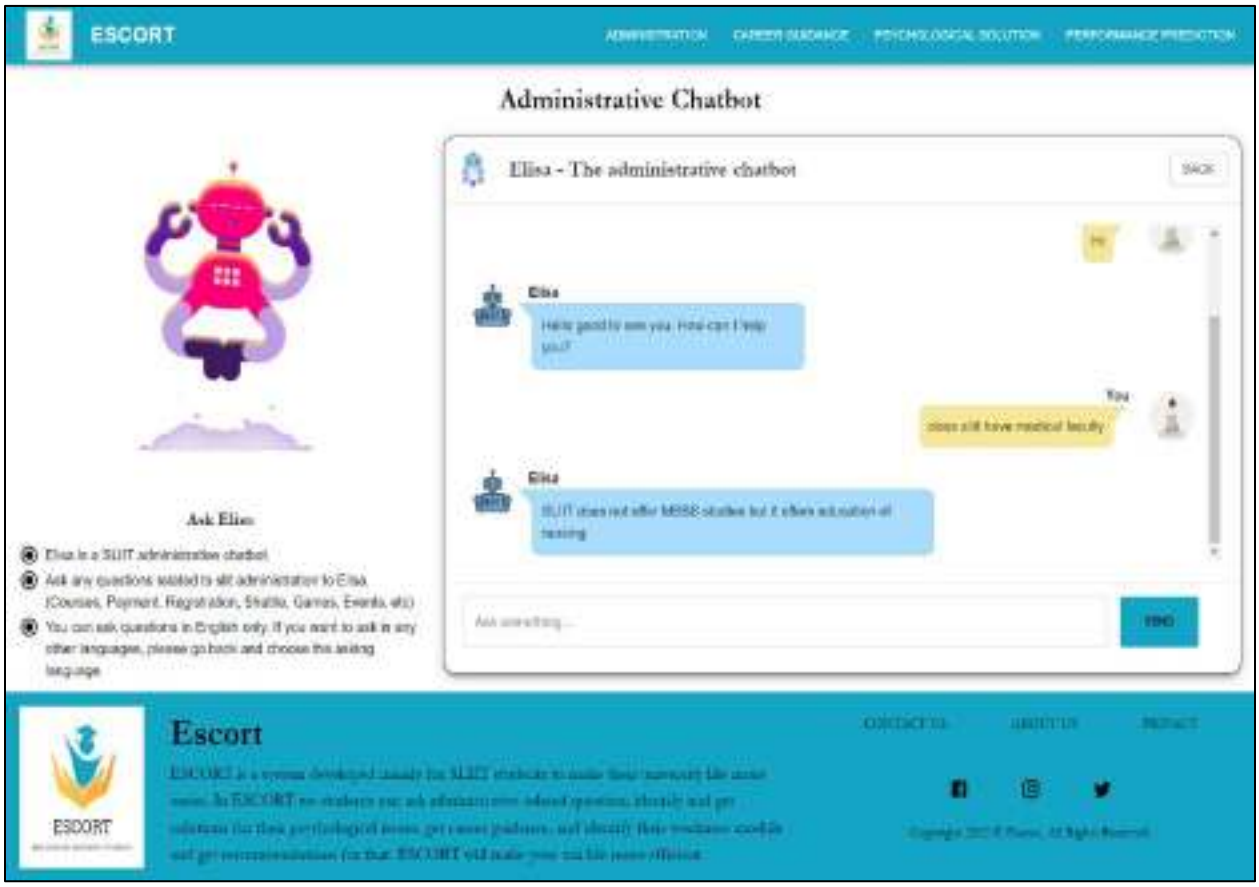


Figure 13: English chatbot

C. Tamil Chatbot

Tamil chatbot page is to ask questions only in Tamil. It has the instructions about the questions users can ask. A chat option is provided in the page where users can easily type their questions and the get response immediately

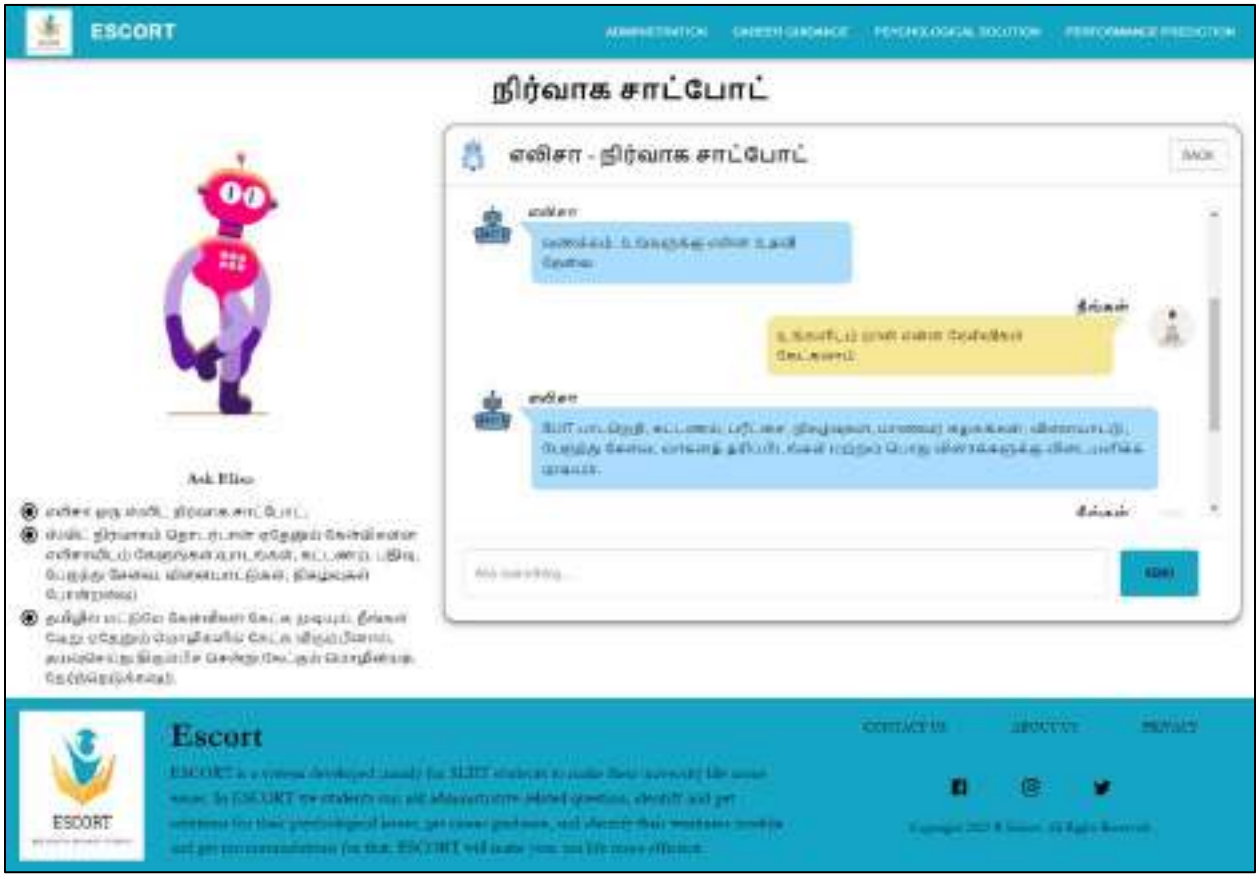


Figure 14: Tamil chatbot

D. Thanglish Chatbot

Thanglish chatbot page is to ask questions only in Thanglish. It has the instructions about the questions users can ask. A chat option is provided in the page where users can easily type their questions and the get response immediately

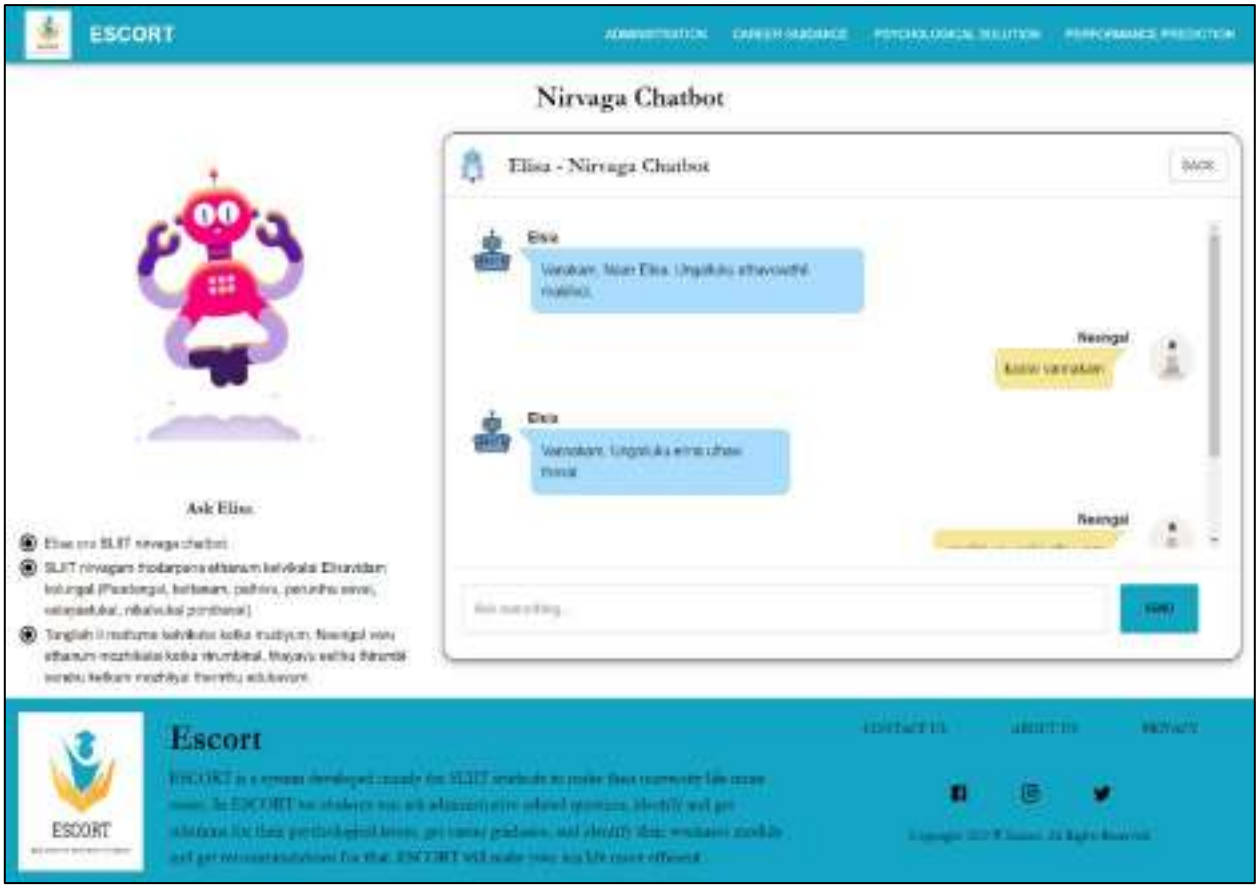


Figure 15: Thanglish chatbot

E. Mobile responsive view

The system is mobile responsive, so users can easily access via any smart devices.

Figure 16: Mobile responsive

5.2 Testing

Testing the system while integrating and before the release is must for every software product. There can be some minor to major issues in the system and it can be easily identified during the testing phrase. Initially the requirement specification testing has been performed to ensure that the implementation is in line with the requirement.

It is import for a system to be responsive and accessible via most of the common smart devices. To check the responsive of the system 'Browser Stack' application has been used. More than that the following testing have been performed on the system.

A. Unit testing

Unit test is a method where testing the smallest piece of code that can be logically isolated in a system. Normally unit test is implemented by the software developer who developed the system. This will help to detect the early flaws in code. In the administrative chatbot system unit test played a main part which helped to identify the minor bugs.

B. Module testing

Used to check the individual components of the system which helps to early detection of errors. Module testing helps to break the system in to smaller parts and test, which reduces the complexity of testing the whole system.

C. Integration testing

Integration testing is the phase in software testing in which individual software modules are combined and tested as a group. It occurs after unit testing and before system testing.

D. System testing

System testing is the testing of a complete and fully integrated software product. System testing is actually a series of different tests whose sole purpose is to exercise the full computer-based system.

E. Performance testing

This test helps to predict how dependable and efficient the system is. This non-functional software technique will be used to determine how the speed, responsive, scalability and stability the system when it is holds up under a given workload. Administrative chatbot system was performed this testing and the test was success. Even in a high network traffic the system performed well.

F. Functionality testing

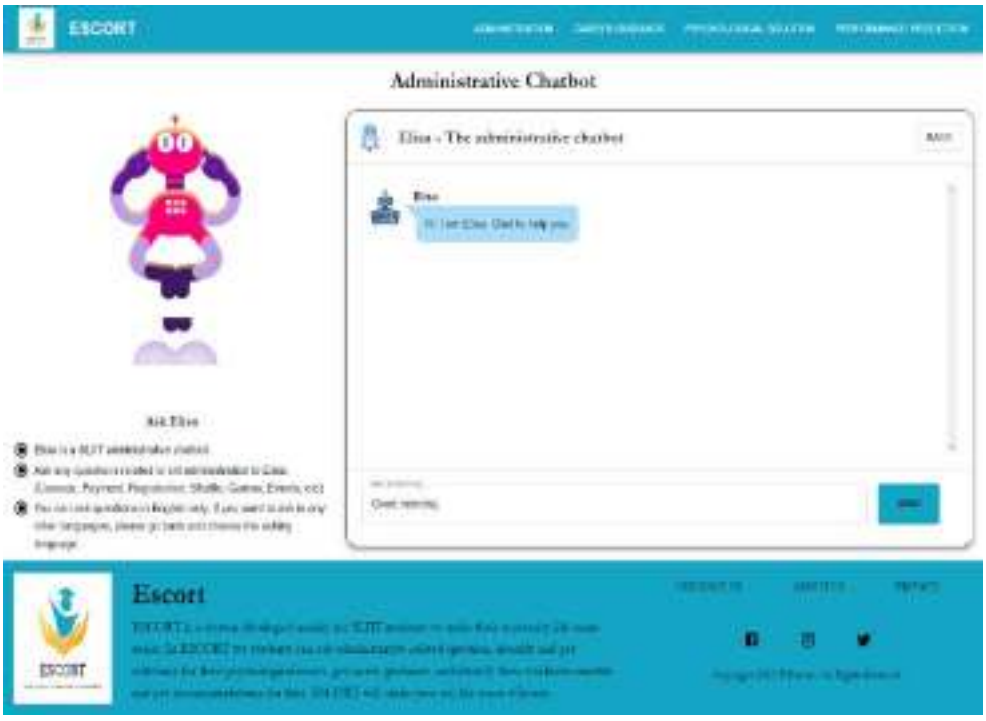
To ensure that the specifications defined in the beginning are fulfilled is tested by functionality testing.

G. Acceptance testing – Alpha testing

Users of the software work with the development team to test the software at the developer's site. To conduct this testing some of the SLIIT students were asked to use the system for their need.

5.2.1 Test cases

Table 2: Test case 1

Test case	Test case 001
Description	User interact with the English chatbot system for the first time.
Summary	User try to interact with the English chatbot for the first time by saying 'Good morning.' The system replies with the correct response.
Pre-condition	Choose 'English' button to navigate to English chatbot
Post-condition	User got the correct response
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose 'English' button 4. Type the question 5. Click 'Send' Button
Test input	Good morning 
Expected result	Hello good to see you. How can I help you? / Hi. How can I help you?

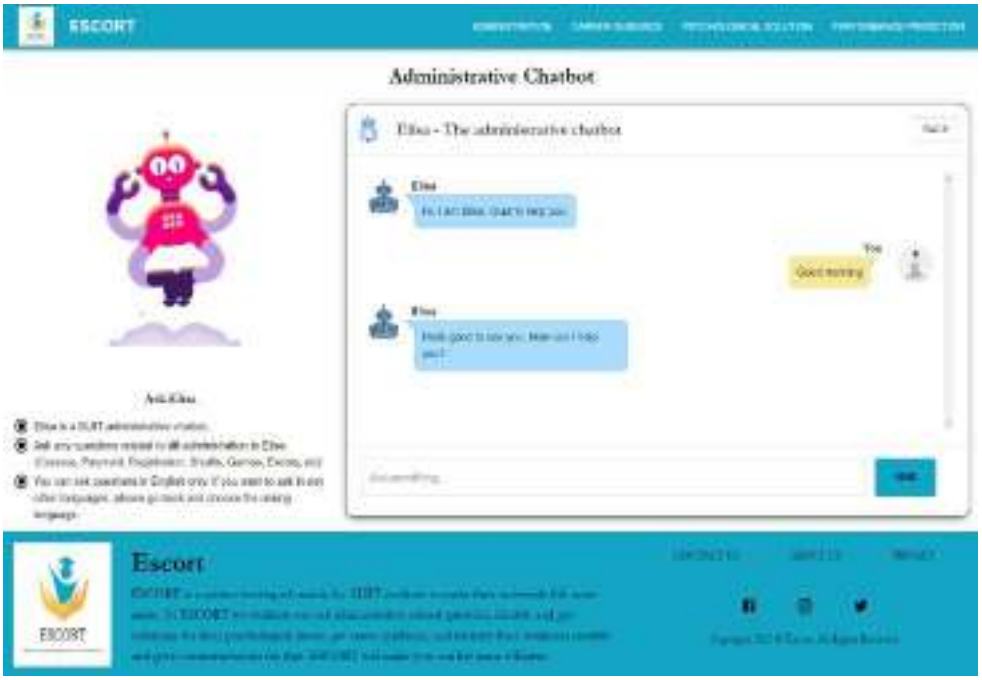
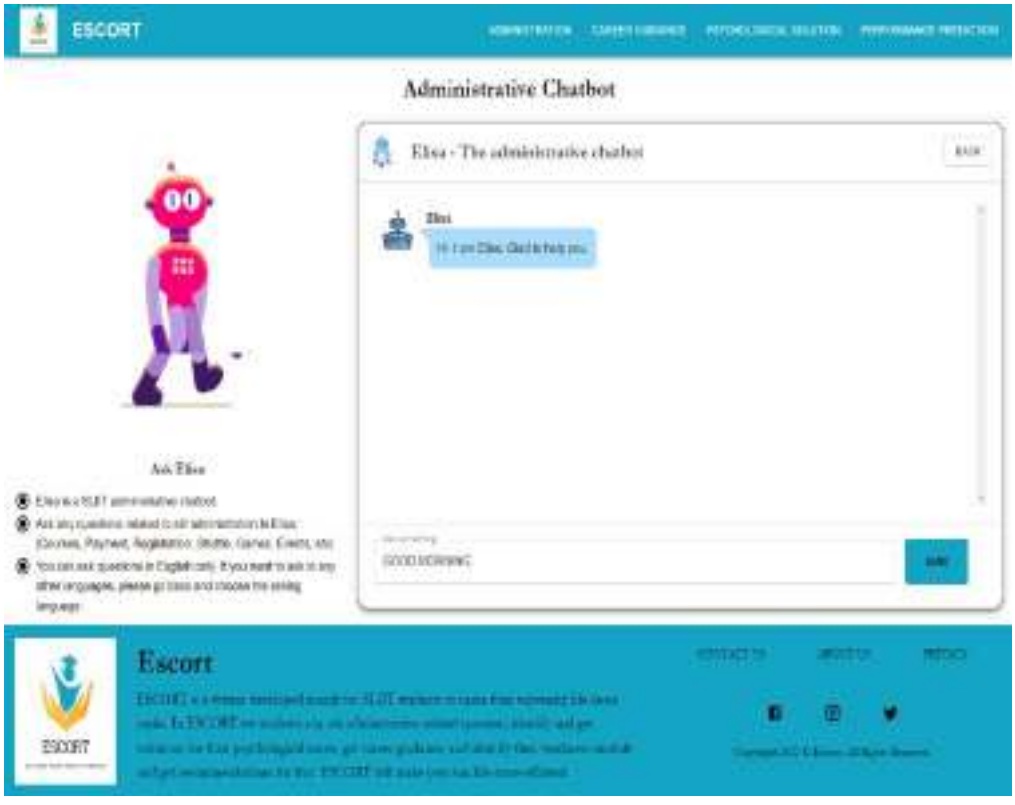
Actual result	Hello good to see you. How can I help you? 
Test result	Pass

Table 3: Test case 2

Test case	Test case 002
Description	User interact with the English chatbot system for the first-time using capital letters only.
Summary	User try to interact with the English chatbot for the first time by saying 'GOOD MORNING.' The system replies with the correct response.
Pre-condition	Choose 'English' button to navigate to English chatbot
Post-condition	User got the correct response
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose 'English' button 4. Type the question 5. Click 'Send' Button
Test input	GOOD MORNING 

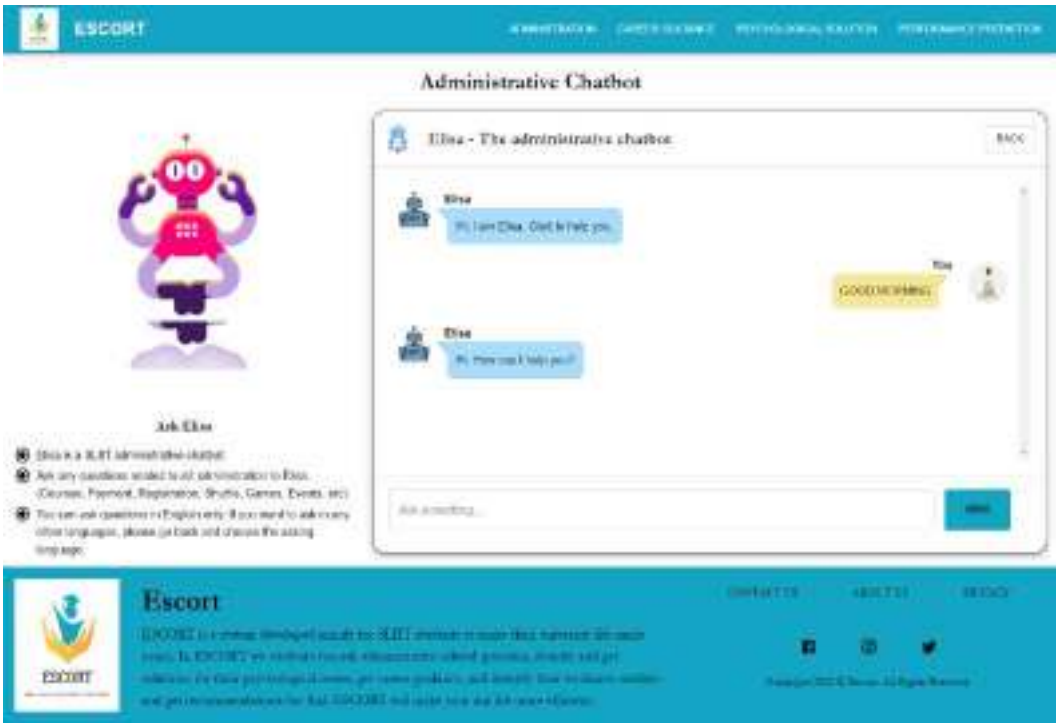
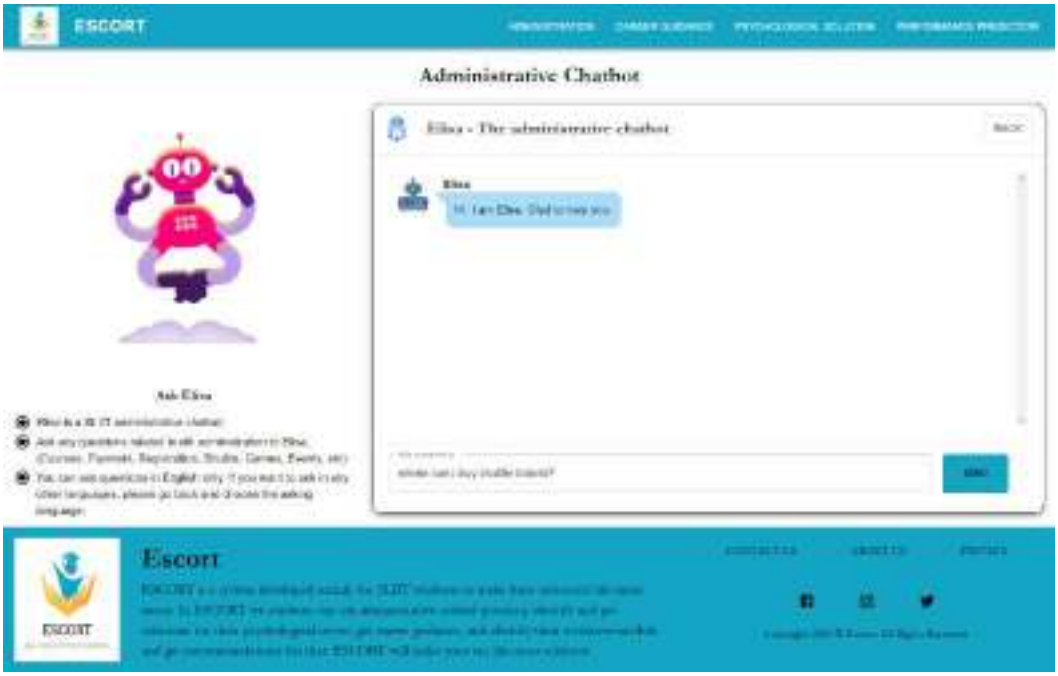
Expected result	Hello good to see you. How can I help you? / Hi. How can I help you?
Actual result	Hi. How can I help you? 
Test result	Pass

Table 4: Test case 3

Test case	Test case 003
Description	User interact with the English chatbot system and ask a random question in lower case sentence.
Summary	User try to interact with the English chatbot by asking random question in lower case. The system replies with the correct response.
Pre-condition	Choose 'English' button to navigate to English chatbot
Post-condition	User got the correct response
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose 'English' button 4. Type the question 5. Click 'Send' Button
Test input	where can i buy shuttle tickets? 
Expected result	You can buy the shuttle tickets in the SLIIT administration office in the main building of Malabe SLIIT.

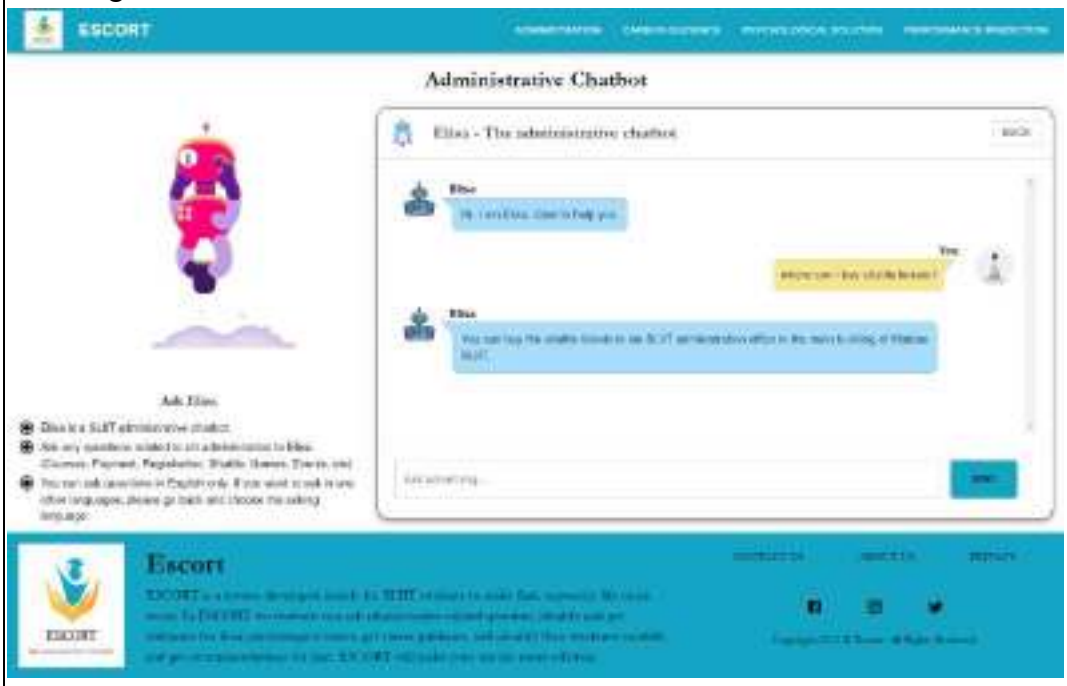
Actual result	You can buy the shuttle tickets in the SLIIT administration office in the main building of Malabe SLIIT. 
Test result	Pass

Table 5: Test case 4

Test case	Test case 004
Description	User interact with the Tamil chatbot system for the first time.
Summary	User try to interact with the Tamil chatbot for the first time by saying ‘எனக்கு உதவி தேவை.’ The system replies with the correct response.
Pre-condition	Choose ‘தமிழ்’ button to navigate to Tamil chatbot
Post-condition	User got the correct response
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose ‘தமிழ்’ button 4. Type the question 5. Click ‘Send’ Button
Test input	எனக்கு உதவி தேவை 
Expected result	வணக்கம். உங்களுக்கு என்ன உதவி தேவை / உங்களை சந்திப்பதில் மகிழ்ச்சி. உங்களுக்கு என்ன உதவி தேவை

Actual result	உங்களை சந்திப்பதில் மகிழ்ச்சி. உங்களுக்கு என்ன உதவி தேவை 
Test result	Pass

Table 6: Test case 5

Test case	Test case 005
Description	User interact with the Tamil chatbot system and ask random question.
Summary	User try to interact with the Tamil chatbot and asking random question. The system replies with the correct response.
Pre-condition	Choose 'தமிழ்' button to navigate to Tamil chatbot
Post-condition	User got the correct response
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose 'தமிழ்' button 4. Type the question 5. Click 'Send' Button
Test input	ஸ்லிட் க்கு மருத்துவ பீடம் உள்ளதா? 
Expected result	ஸ்லிட் MBBS படிப்பை வழங்கவில்லை. ஆனால் அது நர்சிங் கல்வியை வழங்குகிறது.

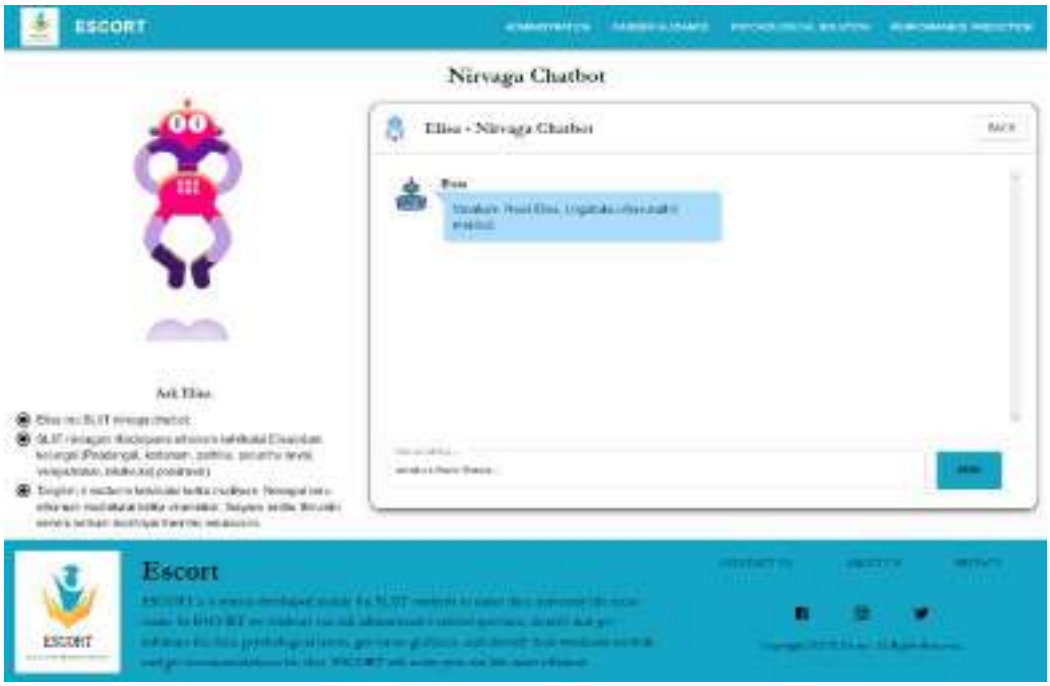
Actual result	ஸ்லிட் MBBS படிப்பை வழங்கவில்லை. ஆனால் அது நர்சிங் கல்வியை வழங்குகிறது. 
Test result	Pass

Table 7: Test case 6

Test case	Test case 006
Description	User interact with the Tamil chatbot system and ask same question in different way.
Summary	User try to interact with the Tamil chatbot and asking the same question in a different way. The system replies with the correct same response.
Pre-condition	Choose 'தமிழ்' button to navigate to Tamil chatbot
Post-condition	User got the correct response
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose 'தமிழ்' button 4. Type the question 5. Click 'Send' Button
Test input	நான் ஸ்லிட்டில் டாக்டருக்கு படிக்கலாமா? 
Expected result	ஸ்லிட் MBBS படிப்பை வழங்கவில்லை. ஆனால் அது நர்சிங் கல்வியை வழங்குகிறது.

Actual result	ஸ்லிட் MBBS படிப்பை வழங்கவில்லை. ஆனால் அது நர்சிங் கல்வியை வழங்குகிறது. 
Test result	Pass

Table 8: Test case 7

Test case	Test case 007
Description	User interact with the Thanglish chatbot system for the first time.
Summary	User try to interact with the Thanglish chatbot for the first time by saying 'enaku uthavi thevai.' The system replies with the correct response.
Pre-condition	Choose 'Thanglish' button to navigate to Tamil chatbot
Post-condition	User got the correct response
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose 'Thanglish' button 4. Type the question 5. Click 'Send' Button
Test input	enaku uthavi thevai 
Expected result	Vannakam. Ungalluku enna uthavi thevai / Ungallai santhipathil makeltchi. Ungalluku enna uthavi thevai

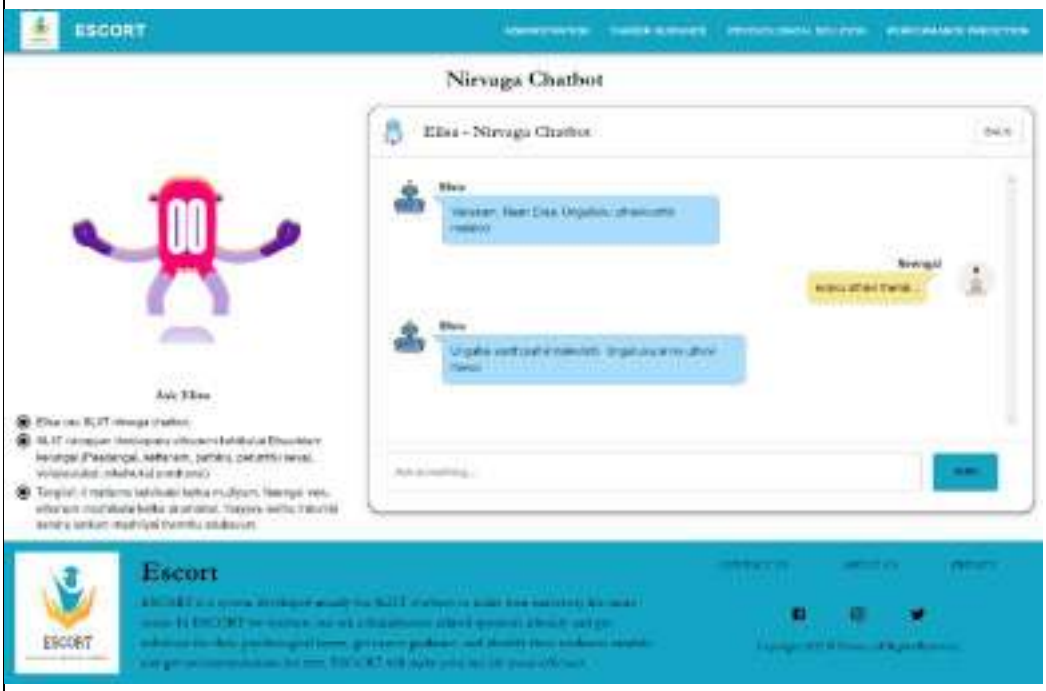
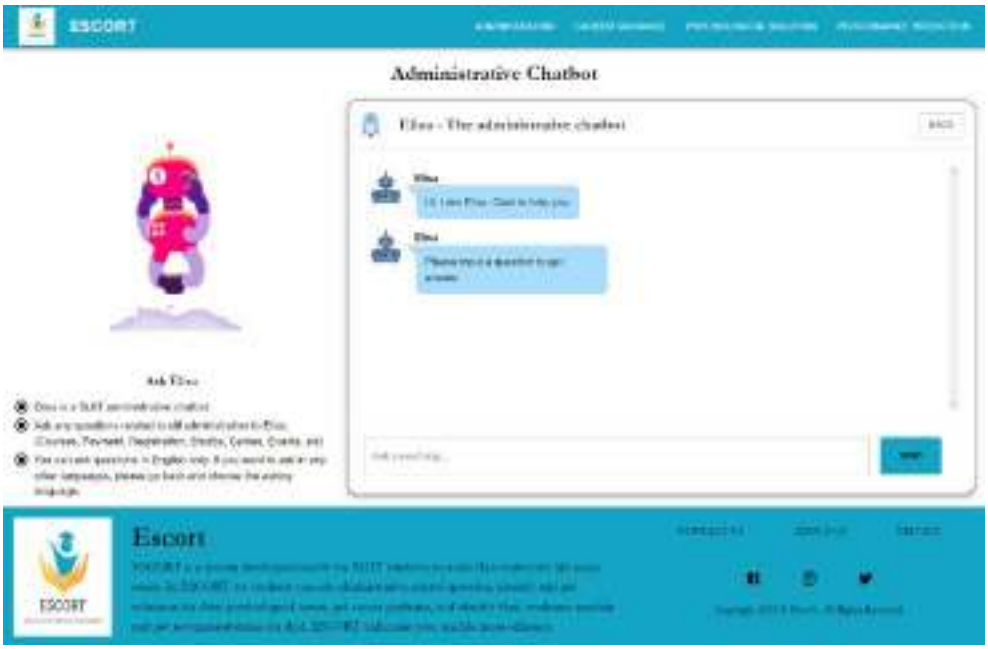
Actual result	Ungallai santhipathil makeltchi. Ungalluku enna uthavi thevai 
Test result	Pass

Table 9: Test case 8

Test case	Test case 008
Description	User try to pass empty value as a question
Summary	User try to pass empty value as question and the system reply by asking to enter a question
Pre-condition	Choose any language button to navigate to responsive language chatbot
Post-condition	User got the response as to enter a question
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose any language button 4. Click 'Send' Button without typing any question.
Test input	Empty
Expected result	Please input a question to get answer / தயவு செய்து வினா ஒன்றைக் கேட்கவும் / Thayavu seithu vinaa ondrai kertkavum
Actual result	Please input a question to get answer / தயவு செய்து வினா ஒன்றைக் கேட்கவும் / Thayavu seithu vinaa ondrai kertkavum 

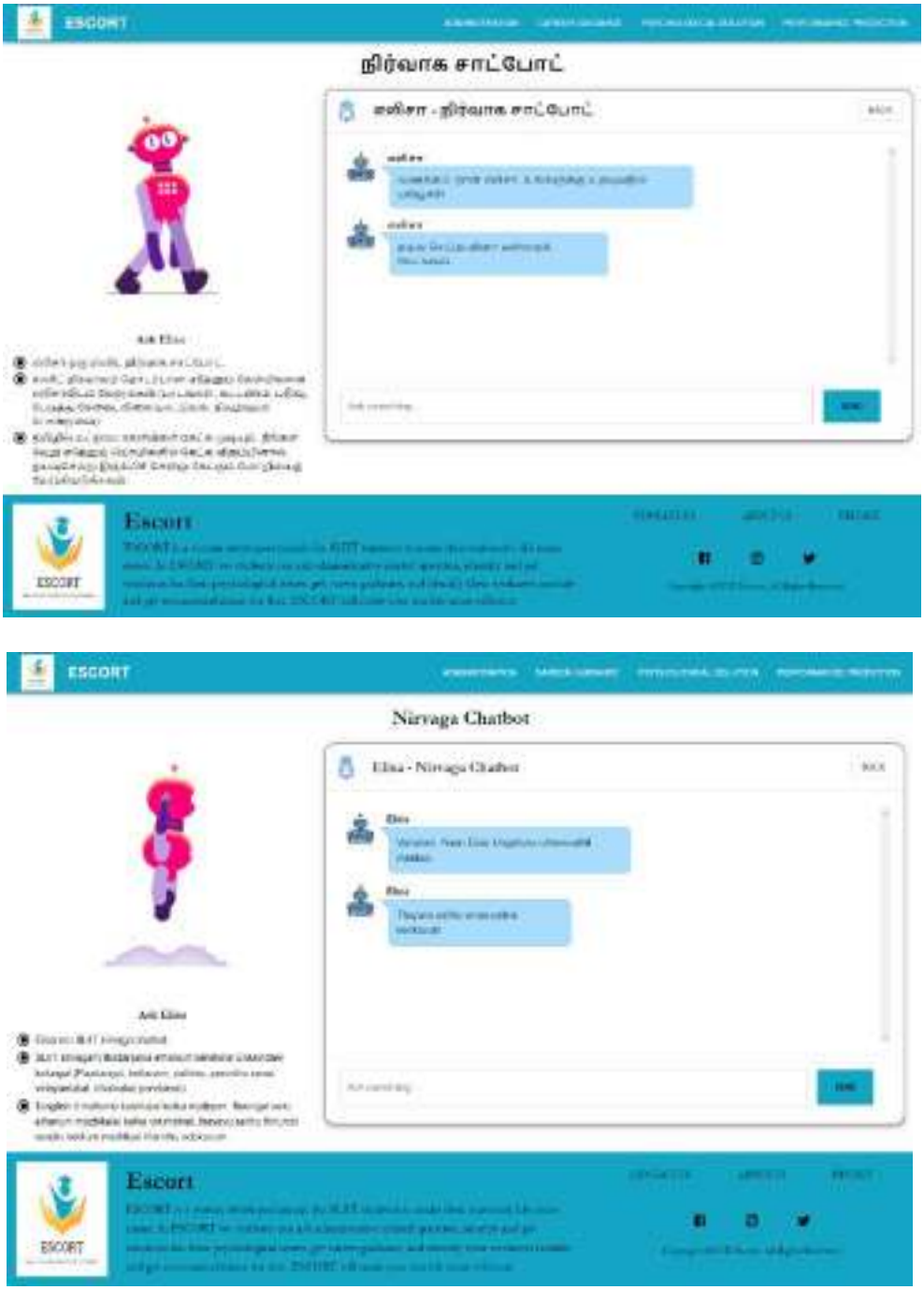
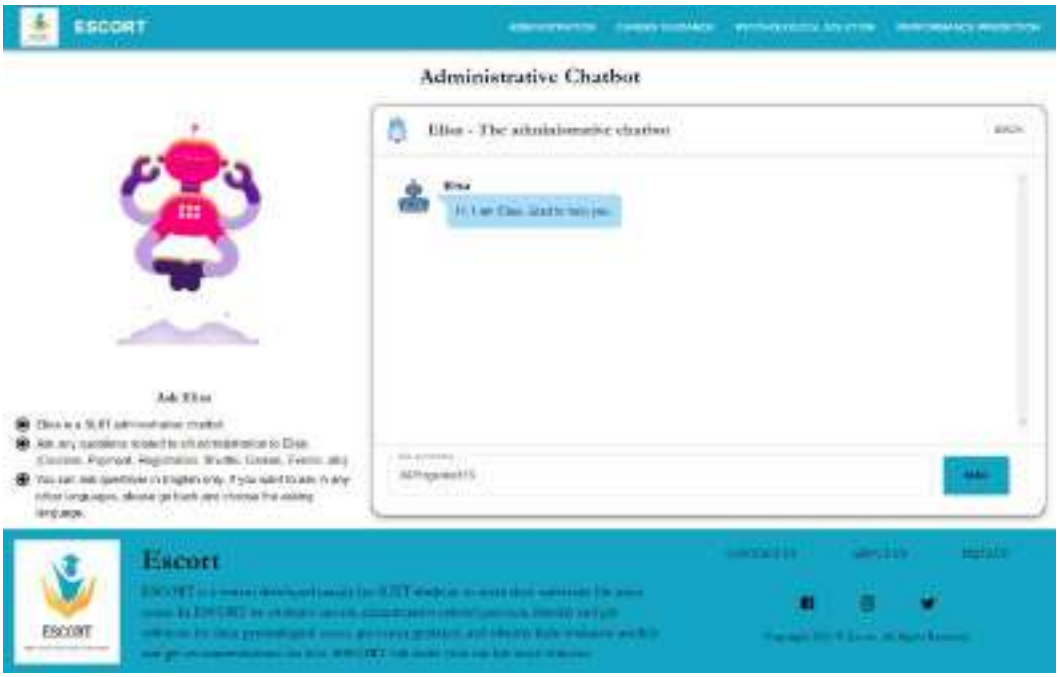
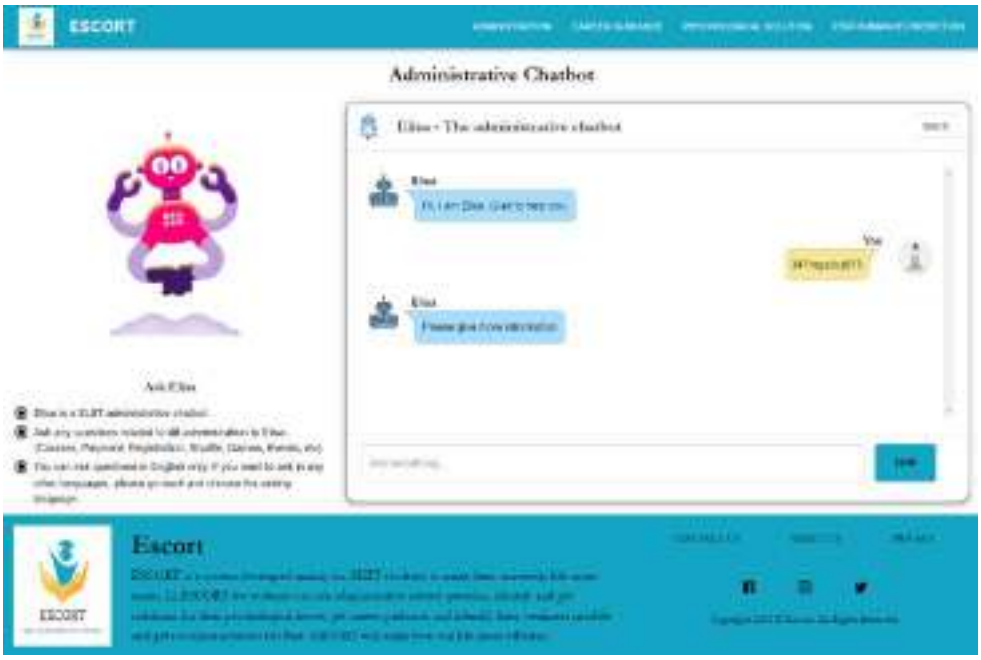
	
Test result	Pass

Table 10: Test case 9

Test case	Test case 009
Description	User try to pass random value as a question
Summary	User try to pass random value as question and the system reply by asking to enter a question
Pre-condition	Choose ‘English’ language button to navigate to English language chatbot
Post-condition	User got the response as to enter a question
Test procedure	1. Go to Escort website 2. Choose Administrative from the top navigation bar 3. Choose ‘English’ language button 4. Click ‘Send’ Button without typing any question.
Test input	347hrgscku875 
Expected result	I am not sure I understand that. / Please give more information. / Sorry. I can't understand that.

Actual result	Please give more information 
Test result	Pass

6 RESEARCH & DISCUSSION

6.1 Results

The implemented system can successfully answer the questions user ask in their preferred language (English, Tamil, Thanglish) with the accuracy above 90%. After the user select a language and insert the question, the system will sand the answer as a reply.

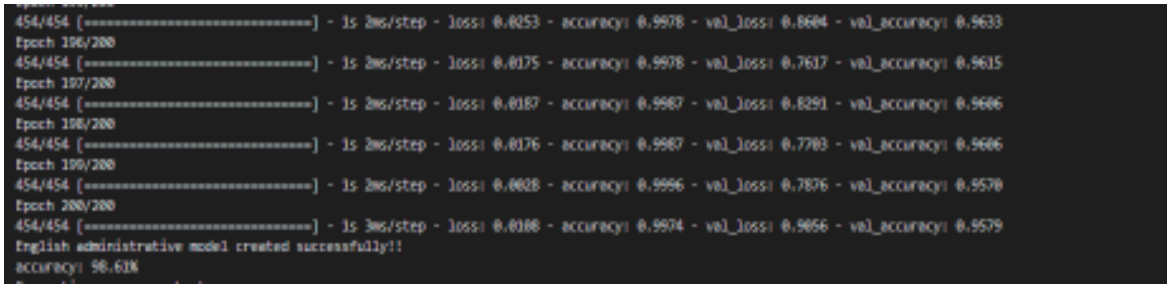


Figure 18: Accuracy of English model training

If a user, ask English language questions to the chatbot the following will be the answer.

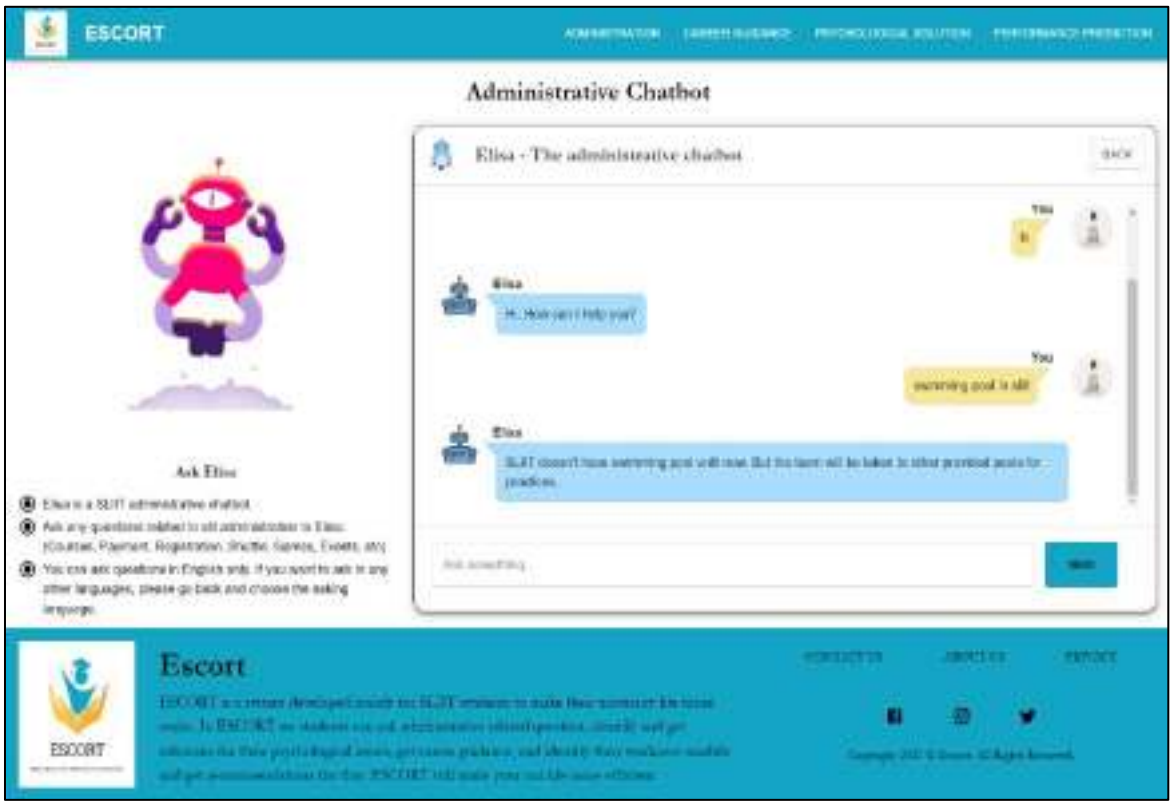


Figure 19: English chatbot

6.2 Research Findings

Considering the questions university students have related to university administration and the issues of getting proper answer for those, I decided to develop a chatbot system which is a peart of Escort. The goal of the system is to have multiple language option for users to ask and receive administrative related questions in nine categories and provide answers or solutions with high accuracy. From collecting the dataset to finish the implementation I faced some issues and some of them were resolved during the testing and for some I have used an alternative way.

6.2.1 Collecting dataset for the model

As mentioned in the methodology I have used google form to get the questions SLIIT students have related to SLIIT administration. As I need more dataset it was not success to collect the data set only from students. So, I have shared the questionnaire with some of the parents and other A/L finished students who would like to join SLIIT. As they also had a lot of questions, I was able to collect enough dataset for the model.

6.2.2 Technology requirement for system development

NLP is an advance technology which need more ram to run without issues. In the beginning I was using a laptop which had 8GB RAM. In the beginning I was able to train the model using that. But as the dataset increases the time for processing and training increased and because of that the API call took long time. After that I increases the RAM to 16GB and then the model worked without any issue and the model build time decreased more than 40%.

6.3 Discussion

Managing time for implementing the project and managing the documentations while studying is one of the major issues I have faced. Submitting the documents and other submission on time was a challenging part. To avoid last minute timing issues, I have maintained the teams task deliverables to manage my tasks.

I faced some issues while collecting the dataset. I have shared the questionnaire only with the SLIIT students and I was unable to collect enough dataset. To get more dataset I shared the questionnaire with some parents of SLIIT students and to some people who would like to join SLIIT in future. From that I was able to get enough dataset.

Once the dataset was collected, I started arranging the questions according to the category. The I contacted the SLIIT administration and got the answers for each question. Once the answers were collected, I wrote other forms for the same question (Ex: What is the fee for Software Engineering can be asked as Fee of Software Engineering). One all the questions have be added to other forms, I have translated the whole question and answer set to the Tamil language. In Tamil also I have identified other forms of asking the same question by arranging the words and change with synonymous a word. Then I translated to Tamil dataset to Thanglish.

Once the dataset was created, I started working on the model. First, I implemented the model for the English language, and I used NLTK and Keras sequence model to train the model. Once the model has trained the accuracy of the English model was around 82%. To increase the model, I added some more datasets and then the accuracy was 96%. Once I have completed the English model I started working on the Tamil and Thanglish models simultaneously. I got 89% and 83% accuracy for them.

After completing building all the models, I implemented the API in Python. Then I started working on the front end. For the front end, I have used React JS. I have developed the front end with a chatbot option and then I have connected the Python API of the chatbot. Then I asked some questions and checked the system in all three languages. Once all the implementation is completed, I have tested all of the functions. Finally, the administration chatbot component of Escort was designed and implemented as expected.

7. CONCLUSION

Most university students are facing difficulties to get solutions to administration-related issues and university administration is also having difficulties to solve the issues. The administration doesn't have enough time to answer all of the questions students have because of the high number of students count in university. Students are also facing communication language issues when they ask questions. Most of the students preferred to ask and get solutions in their native language. But English is being the primary language for most universities. Sometimes students are getting the wrong information because of this language issue.

This research aims to provide an online chatbot system, where users can choose their preferred language (English, Tamil, or Thanglish) and get responses in the same language. This system will help users to save more time and get a quick accurate answer to their administration-related questions. The system provides solutions/ answers with an accuracy rate above 90%.

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] A. Macaskill, 'The mental health of university students in the United Kingdom', Br. J. Guid. Couns., vol. 41, no. 4, pp. 426–441, Aug. 2013, doi: 10.1080/03069885.2012.743110.

[5] 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.

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

[7] 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.

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

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

[10] 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].

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

[12] "Text Preprocessing in Natural Language Processing", https://towardsdatascience.com/text-preprocessing-in-natural-language-processing-using-python-6113ff5decd8

[13] "Python Pickle File", https://fileinfo.com/extension/pkl#:~:text=A%20PKL%20file%20is%20a,stream%20that%20represents%20the%20objects

APPENDICES

Appendix A: Use case diagram