ABSTRACT

In the modern Era of technology, Chatbots is the next big thing in the era of conversational services. It is designed to be the ultimate virtual assistant, entertainment purpose, helping one to complete tasks ranging from answering questions, getting driving directions, turning up the thermostat in smart home, to playing one's favourite tunes etc.

Whenever faced with unchartered territories humans have a tendency to get anxious and it is but natural and even more so in the case of students who are about to join a new course. They are always filled with endless queries and mostly all of them are similar across people. Rather than flooding a single person with all those and to answer them over and over again, it is apt to segregate usual common queries and their solutions at a single place which is easily accessible by all.

In this project the detailed design of a chatbot is given which provides an efficient and accurate answer for any query based on the data set of FAQs. This chatbot can be used by Sree Chitra Thirunal College Of Engineering to answer frequently asked questions to curious students in an interactive fashion.

CONTENTS

LIST OF ABBREVIATIONS

ML Machine Learning

NLP Natural Language Processing

HCI Human Computer Interaction

AI Artificial Intelligence

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

A chatbot system is defined as a "software program that simulate conversation with human users, using text, voice or images or a combination of spoken and visual heuristics"(Shawar Atwell, 2007). Several terms are used to describe chatbots such as machine conversation systems, conversational agents, virtual agents and dialogue systems.Common for all is that they utilize a computer program that can perform tasks automatically or with minimal human intervention.

One of the important goals in the field of Human Computer Interaction (HCI) is the outline of normal and instinctive connection modalities. Specifically, numerous endeavors have been committed to the improvement of frameworks to communicate with the client in a characteristic language. Computer based chatbots are getting to be distinctly famous as an intuitive and successful open framework between human and machines. Chatbot is a manufactured substance that is intended to reproduce a clever discussion with human accomplices through their regular language. Currently, chatbots are utilized by a great many web clients to intercede access to information or learning bases and furthermore to do nonspecific discussions.

The first chapter gives a basic introduction about chatbots and their working. The problem chosen and the objectives the project is supposed to solve the problem considered. The second chapter gives the literature survey for the project in which all the existing system models that are similar to the project model is explained. Then the chatbots used in Universities are explained. The system proposed for the project is explained in detail in the beginning of third chapter. The third chapter gives details about system analysis and design that was carried out. The dataset that was developed for the project and how it was built, the format chosen to save the information everything is explained in this chapter. The next section details about the interface that is used and implemented in the

project. After detailing the interface, the various technologies used to implement these are explained in detail. Then the architecture layout of the chatbot is given along with a diagrammatic representation. The next section gives detailed explanation about the algorithm used in the project and its working. The last section gives information about the methodology used for developing the project and the reasons for selecting the stated methodology. The last two chapters deals with the final results and conclusion for the project. Finally, the references for the project are given.

1.1 AI, machine learning and chatbots

It is important to separate between the terms artificial intelligence(AI), chatbots and machine learning. A chatbot is basically an interface designed as a conversation, and is what meets the eye whereas AI and ML stays behind the stage curtain. Machine learning is a term used for the algorithms that gives systems the ability to learn from data, and is a subset of AI. Artificial intelligence, with a sci-fi ring to its name, is harder to define. In short, it is the branch of computer science that deals with the simulation of intelligent behaviour in computers. What exactly intelligent behaviour from a computer is, and what people perceive as intelligent behaviour from a computer, that is another question that this project will stir into.

1.2 How Do Chatbots Work?

Most people won't build their chatbots from scratch as there are plenty frameworks and services that can help. However, in order to grasp how bots work, we have to go over some developer talk.

1.2.1 Backend

Chatbots can be built in basically any programming language that allows you to make a web API. For most people this will be either Node.js or PHP, but there are many bot libraries written for Java and Python as well. The backend receives messages, thinks of a response, and returns it to the user. It performs a very important task in the case of a chatbot as the main query processing and training part of the chatbot is carried out by the backend. This section is to be implemented with extreme care and an appropriate programming language is to be chosen so as to be able to add features by the developer.

1.2.2 Frontend

This can be one of the popular messengers apps (Facebook Messenger, Slack, Telegram), or a simple chat interface like the Realtime Chat With Node.js. You are not limited to a single platform - the same bot can be implemented in more than one place. It is the part of the chatbot that makes the bot usable to users and connects it to real world entities. The implementation of an appropriate front end for the chatbot depends on the type of chatbot that is being implemented and the platform where the bot is to be published. If the chatbot is a web application it would be implemented on a website or a webpage but if it is a software type then it would be implemented on a platform operating system.

1.2.3 Integration

After developing the frontend and backend, they both have to be integrated and a connection has to be maintained between the two so that the user input that is being sent from the front end is accepted at backend and the output selected from the backend is send to the front end. If there is any problem in the input and output, it should be able to pass through the integration environment used and the system must not crash. There are multiple integration environment that can be used and depending on the language used to write the backend we can use any one.

1.3 Need of project

Whenever faced with unchartered territories humans have a tendency to get anxious and it is but natural- and even more so in the case of students who are about to join a new course. They are always filled with endless queries and mostly all of them are similar across people. Rather than flooding a single person with all those and to answer them over and over again, it is apt to segregate usual common queries and their solutions at a single place which is easily accessible by all.

Problem Definition

Presently, the college has no provision for any such system and therefore our suggestion to create one to simplify the redundant nature of work and utilize manpower efficiently. The need for a virtual assistant is imminent in this scenario.

1.4 Objective

The project is made with the following objectives in mind :

- To provide responses to query regarding college activities. This is one of the most important function the chatbot is expected to perform. The students will have queries regarding college activities which are not easily available. If the student wishes to get some information, the only resource he or she has is the college website and reading through the entire content is a tedious process.

- Solving queries regarding college admission. The chatbot also has the objective to solve the questions of the user regarding the admission procedures and details about the various departments and how to obtain admission to any particular department.

- Solutions to doubts regarding college facilities. Any student who wishes to join a college would like to get to know about the various facilities provided by the college such as the lab facilities and whether uniforms are allowed or it could be even information about hostel facilities in and around the college. The students can get the information related to labs on the college website but details regarding uniforms is not made available.

- Solutions to queries regarding college staff. It is possible that it is not just students that are trying to access the wesite chatbot, there is also a possibility that people who want to know about the staff qualification at the college. This information has to be provided to the user by the chatbot.

CHAPTER 2

LITERATURE SURVEY

"Chatbot and conversational agents: A bibliometric analysis"[7] tells about the increasing demand of chatbots. "Chatbots are replacing some of the jobs that are traditionally performed by human workers, such as online customer service agents and educators. From the initial stage of rule-based chatbots to the era of rapid development in artificial intelligence (AI), the performance of chatbots keeps improving. Chatbots can nowadays “chat” like a human being and they can learn from experience." it quotes.

"An e-business chatbot using AIML and LSA"[1] explores the possibilities of a chatbot being used in the business world.The paper has stated the difficulties of the accuracy with which an online chatbot would converse by copy pasting pre written answers.It uses AIML and LSA to create the chatbot unlike our chatbot.

"Programming challenges of chatbot: Current and future prospective" [3] speaks about all the challenges that are encountered while building a chatbot. Chatbot building is a tedious task. It speaks of the scalability, stability and flexibility issues.

"Chatbot using a knowledge in database human-to-machine conversation modelling" [2] gives the insight to chatbot building. It explains the sentence similarity algorithm and concepts of pattern matching.

"Chatbot for university related FAQs"[8] is the paper that has done work similar to the present project. Manipal university has come up with a chatbot for their university. The bot is a FAQ bot and only college related FAQs can be asked. They have used AIML to build the bot.

"Evaluation of Modern Tools for an OMSCS Advisor Chatbot"[9] At Georgia State, “Pounce” the chatbot has been helping students navigate the application process. At technical university of Berlin, Germany, a chatbot named "Alex" helped students in test groups and schedule classes more efficiently.

Instead of searching through online timetables, students type normal questions to Alex, such as when classes are, who will be teaching them and what exams they have to take. The bot replies in natural-sounding sentences and can ask follow-up questions to get to the bottom of what a student really wants to know.

Georgia Institute of Technology went a step further by introducing a virtual teaching assistant named Jill Watson — which is powered by IBM Watson — in an online course about artificial intelligence. Jill Watson was first used in a spring 2016 course to answer questions in a class chatroom, and students weren’t aware they were interacting with AI until the last day of class.

Ashok Goel, the professor of computer science and cognitive science behind Jill Watson, tells that he was surprised at the chatbot’s effectiveness in boosting student engagement. Since the initial roll out, Georgia Tech reports that the virtual TAs have continued in the course, with some students even creating their own personalized chatbots.

CHAPTER 3

SYSTEM DESIGN

3.1 Proposed System

Chatbots assist people daily with everything from ordering pizza to dealing with customer service issues. So, it's no surprise that higher education institutions are embracing them to interact with their No. 1 customer: students. Whether it's navigating the admissions process or scheduling classes, universities have embraced artificial intelligence to streamline student interactions and offer timely support.

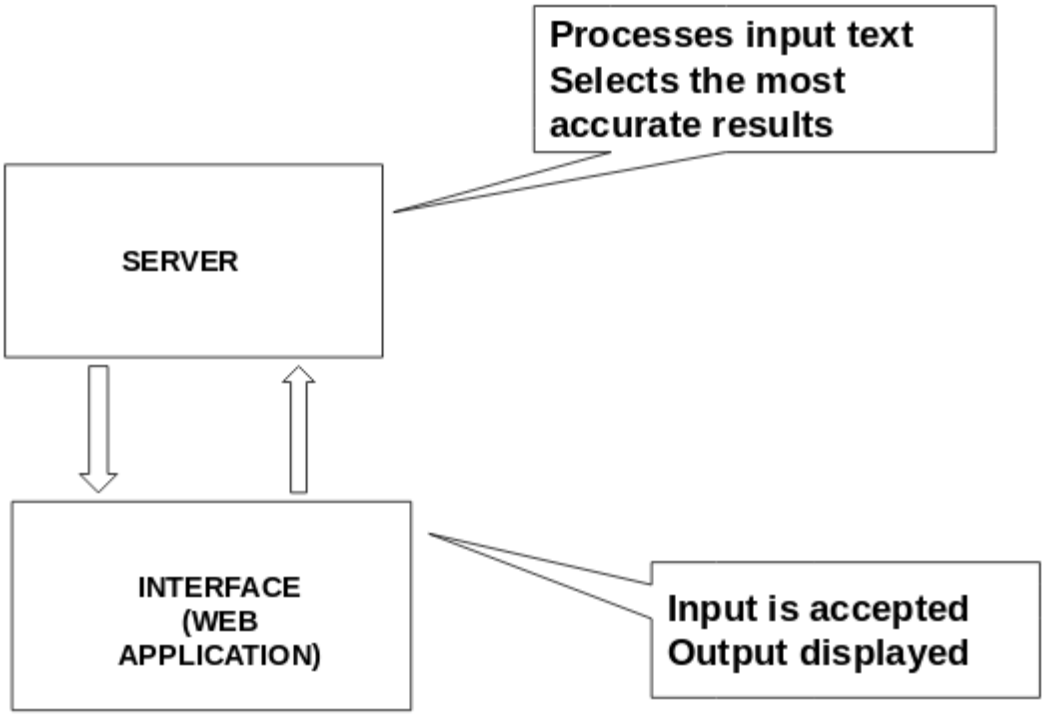

Figure 3.1: Model of a Chatbot

.

Figure 3.1 gives the standard model of a chatbot that is in use. The main back-end server is connected to the chat application interface that helps user to connect to the bot. Instead of blindly searching the internet for information on colleges, students could be asking chatbots their questions. In this work we have developed a interactive chatbot for college related queries, and the work flow of proposed framework is as given below. User discussion as a rule begins with welcome or general questions. User inquiries are encoded and send to the server that sends it to the engine back-end code that is working. It uses NLTK libraries and various functions to calculate the probability of each word in the query with its presence in the dataset and then gives the response with the most probability match. This operation is divided into three parts:

1. User post the query on chatbot
2. Processing is done on the users query to match the required response
3. Pattern matching is performed between user entered query and knowledge (pattern).

Finally pattern based answer is presented to the user to answer their query. This project presents the chatbot for educational sector, where user (a student or parents) can ask query regarding college admission, about college information and other things related to academics. As discussed earlier, user can post their query on chatbot and response is generated based on algorithm followed.

3.2 Architecture

The architecture model of a chatbot is decided based on the core purpose of development. There are two types of possible responses of chatbot: it can either generate a response from scratch as per machine learning models or use some heuristic to select an appropriate response from a library of predefined responses.

Though pattern-based heuristics deliver good results, the problem is that it requires all the patterns to be programmed manually. This is a tedious task, especially if the chatbot has to distinguish hundreds of intents for different scenarios. Intent classification is completely based on machine learning technology that allows for the training of bots. With a training set of thousands of examples that are most likely to be faced by the chatbot, it can be trained to pick up patterns of data and learn from it. Scikit-learn is

a popular machine learning library that helps in executing machine learning algorithms. Developers even have the option to use one of cloud APIs among api.ai, wit.ai, and Microsoft LUIS. Recently bought by Facebook, wit.ai was the first machine learning API for chatbots.

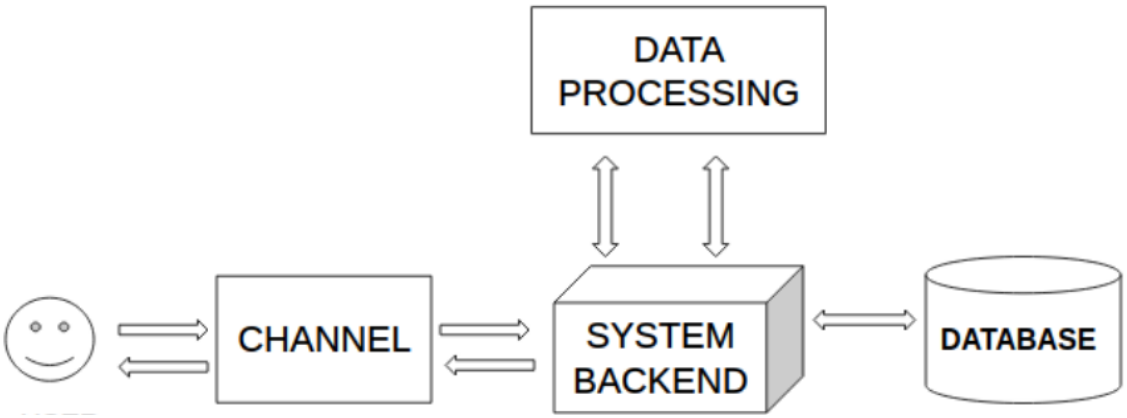

Figure 3.2: Architecture

.

The chatbot architecture shown in Figure 3.2 describes how the various components of the chatbot can be integrated together to form the final system. It consists of a channel which is used to interact with the user. The channel could be any software that will run on a system or if the chatbot is implemented as a web application, the website will act as the channel. The channel will be connected to the system backend which is the main part of the entire system that will be used to process the queries that the user sends to the system. This query is passed on to the data processing block of the system which will implement the algorithm used in the system to get the desired results. The keyword matching algorithm used here will attempt to identify keywords in a sentence. In the case that one or more keywords are found in the user's input text then an answer will be retrieved. The database or the dataset used here contains the required information that will be parsed by the chatbot to obtain the desired results.

Neural Networks

Neural networks are a set of algorithms that are designed to recognize patterns. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, be it images, sound, text or time series, must be translated. Neural

networks help to cluster and classify. It is a clustering and classification layer on top of the data. It helps to group unlabeled data according to similarities among the example inputs, and they classify data when they have a labeled dataset to train on. It extracts features that are fed to other algorithms for clustering and classification and thus can be used as components of larger machine-learning applications. Important things to be kept in mind are :

- Neural networks require clear and informative data (and mostly big data) to train.
- It is prudent to use Neural Networks for complex problems.
- Hardware requirements are essential for running a deep neural network model.

It's easy to classify TensorFlow as a neural network library, but it's not just that. It is designed to be a powerful neural network library. But, it has the power to do much more than that. Every library has its own "implementation details" which follows its coding paradigm.

3.3 Algorithms

3.3.1 Sentence Similarity Measurement

Semantic similarity is giving score for semantic relation between two sentences or strings. So, if there are two sentences or strings, from measuring it can be determined the similar of two sentences or strings. The higher score of the sentence semantic similarity, the more similar meaning of two sentences. The score of the sentence semantic similarity is from 0 until 1. The equation of sentence similarity represented by equation:

$$\frac{\varsigma\left(S_1 \epsilon S_2\right) \cup \varsigma\left(S_2 \epsilon S_1\right)}{\varsigma\left(S_1\right) \cup \varsigma\left(S_2\right)} \tag{3.1}$$

The symbol of is count of the sentence or string that is symbolized as s. The used algorithm to compute the semantic similarity between two sentences is given below

1. Each sentence is divided into a list of tokens.
2. Implement bigram on the divided sentence
3. Apply the following equation to the tokens

$$\frac{\varsigma\left(S_1 \epsilon S_2\right) \cup \varsigma\left(S_2 \epsilon S_1\right)}{\varsigma\left(S_1\right) \cup \varsigma\left(S_2\right)} \tag{3.2}$$

4. Obtain the similarity score for both the sentences.

An example of algorithm can be given as below,

s 1 = “burung”

s 2 = “burrungg”

so s1 and s2 can be divided into bigram set as

s 1 = “bu”, “ur”, “ru”, “un”, “ng” 5

s 2 = “bu”, “ur”, “rr”, “ru”, “un”, “ng”, “gg” 7

s 1 s 2 = “bu”, “ur”, “ru”, “un”, “ng” 5

s 2 s 1 = “bu”, “ur”, “ru”, “un”, “ng” 5

The similarity score in this case is obtained as follow,

$$\frac{5 \bigcup 5}{5 \bigcup 7} \approx \frac{10}{12} \approx 0.83333 \tag{3.3}$$

3.4 Dataset

The task of collecting data was long and tedious one since we were planning on hard coding the questions and answers to provide a very personal experience to our users. The aim was not to provide just any question and answer bot but was to provide a friend and make it as real as possible. The following was done for data collection:

1. A survey was carried out in the college to know the various questions the students have had before joining the college. All the years and all the departments were made a part of it. Some students from M.tech were also made a part of it.

2. Study of a similar project – Manipal University had created a similar project; they created a question and anwer bot for their college. We went through their paper work but didn’t have a hands on experience on their bot because their site doesn’t have one as of now.

3. Google forms were issued students were asked to blow it up with their questions. The form was made available for only SCT students since only looking for SCT related queries were required. It was easy to fetch the general doubts students have before joining an institute and that was not the focus here so it was made just a part of the project, it was not the project. The form gave the students an option to give answers also. If they had a response for any of the query, they could do so.

4. All the queries and their respective responses were held in a single file.

5. JSON format was used to store the data.

6. Tags and contexts were used to differentiate queries related to different subjects.

3.5 Technology Stack

3.5.1 Front end

HTML,CSS and Javascript

HyperText Markup Language (HTML) is a markup language[1] for creating a web-page. Cascading Style Sheets (CSS) is used to style HTML elements while JavaScript is used to manipulate HTML elements and CSS styles. The college website is developed using Bootstrap.

3.5.2 Back end

Python3

Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. Version 3.6.5 was used to carry out the project.

NLTK Toolkit

The Natural Language Toolkit (NLTK) is a platform used for building Python programs that work with human language data for applying in statistical natural language processing (NLP).

Pandas and Numpy

Pandas is one of the data centric python packages that makes importing and analyzing data much easier. Pandas version 0.23.4 was used for the project. Numpy is that

library for computing in python.

3.5.3 Web integration

Flask

Flask is a micro web framework written in Python. It is classified as a micro-framework because it does not require particular tools or libraries. The version used in implementing the project is 1.0.2

3.6 Chat Interface

Chat interface is the medium used by the chatbot to connect with the user. The chatbot is to be embedded in the official website of the college which is www.sctce.ac.in

HTML and CSS were used in the interface. The bot was embedded such that as soon as the home page opens, the bot button would be visible to the user. The user can access it if need be. The user can interact to the bot only through texts. As mentioned earlier, answering FAQ questions such as Examination results, timetable etc have not been the project's priority, they are just a part of the project. Hence they've been replied with links to the sites which would give them their answers. More personal questions have been given answers. The questions that the bot doesn't know answers to, will take the user to search on Google and politely letting the user that bot Is unable to answer the query. For asking questions, the user has to input their query into the text box provided and click on the submit button. The chat interface shows the user's question and the respective response. On asking the next question, the question asked previously and its response is no longer available.

3.7 Methodology

3.7.1 Incremental Model

The incremental model follows a series of processes, which are used during development. Usually the stages will require the gathering of requirements and their analysis. The design of the system is the next stage, followed by coding the actual system. Then evaluation, testing and debugging, if necessary, is the next step. Finally the system will

either be accepted or not and therefore maintained or rejected respectively. It is vital to move to the next process of the incremental model if the previous step has been completed.

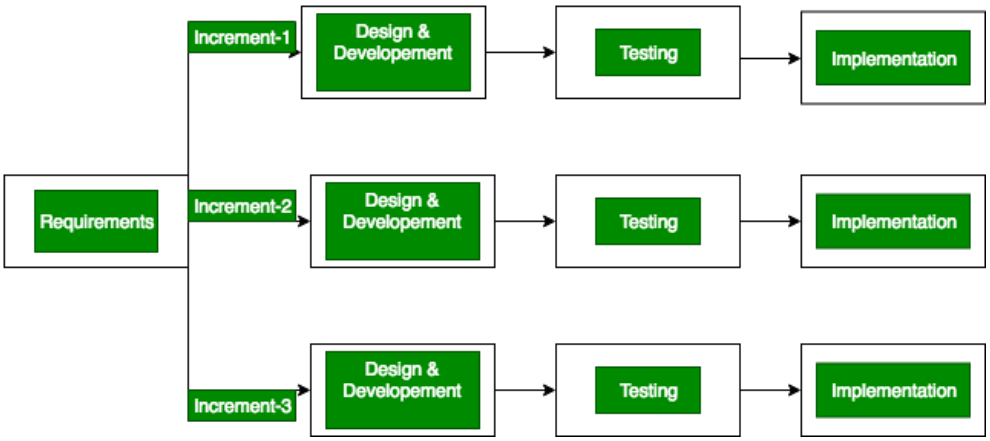

Figure 3.3: Overview of Methodology

.

Requirement gathering and analysis

This is the phase in which the need for the project is specified. All possible requirements of the system to be developed are captured in this phase and documented in a requirement specification document. The project is intended for students who wish to get details regarding the college without having to go through the entire college website. This need was translated into the software by preparing a chatbot accessible very easily from the college website which would provide replies to queries regarding the college.

System Design

The requirement specifications from first phase are studied in this phase and the system design is prepared. This system design helps in specifying hardware and system requirements and helps in defining the overall system architecture. The requirements were used to develop the basic architecture of the chatbot. It consists of a back end which would process the query using the dataset provided. A front end that would be used to connect with the user and a data set that gives the required information about the college which is provided to the user when he or she requests for any specific information.

Testing

All the units developed in the implementation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures. Indiviual testing of units was carried out module-wise. Initially the training of the chatbot was carried out. The connection establishment was checked first after which the query processing part was checked. The test plan was developed after carrying out the following procedures:

1. Analysis of system for testing purpose

   In this step, the inputs that are to be provided to the bot and the various outputs that are expected from the bot was formulated.

2. Scope of modules defined

   The various modules that comprise the whole model is defined to be in scope or out of scope. The modules which can be tested are defined as in scope. In the project, the server and client modules, the trainer module and the bot engine module are to be tested. There are no modules in the project that are not to be tested.

3. Testing type identified

   The strategy to be followed for testing was identified. The techniques to be used were decided to be Unit testing, Integration testing and System testing.

   - Unit testing - Each individual software module was tested. The trainer module was run initially to check whether model is being formed and trained by the system. The server module was run and tested. The client module was run and tested.
   - Integration testing - A group of software modules were collected together and tested. The client and server module was together tested for the connection established between the components. The bot engine was run to check whether it is able to fetch responses from the processed data.
   - System testing - The fully completed model was tested on the system after integrating with the website to check whether it is able to fulfil the requirements.

S.No	Input provided	Expected output	Obtained output	Accurate(Yes/No)
1.	hello	Hello, I am here to solve all your queries	Hello, I am her to solve all your queries	Yes
2.	what is your name	Hello,I am SCTCE's virtual assistant	Hello,I am SCTCE's virtual assistant	Yes
3.	What are the departments	The college has 6 departments-CS,EC,BT,Mech(Auto & Prod)	The college has 6 departments-CS,EC,BT,Mech(Auto & Prod)	Yes
4.	how can i reach CCF	Walk right to the workshop towards canteen building	Call the college at 9446567842	No
5.	what is rank of college in Trivandrum	It is one of the top 5 Engineering colleges in Trivandrum	It is one of the top 5 Engineering colleges in Trivandrum	Yes
6.	what are you	Hello,I am SCTCE's virtual assistant	Hello, I am her to solve all your queries	No
7.	does the college have accreditation	Yes, the college is accredited by National Board Of Accreditation for 4 branches.	Yes, the college is accredited by National Board Of Accreditation for 4 branches.	Yes
8.	where is CCF in college	CCF is above the canteen.	First floor- administrative office of the college.	No
9.	Does the campus have hostel facilities	No, there are no hostels provided by college.	No, there are no hostels provided by college.	Yes
10.	Are mobile phones allowed on campus	Mobile phones are not allowed in the campus.	Mobile phones are not allowed in the campus.	Yes

Table 3.1: Test plan for system testing

Implementation

With inputs from the system design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality. Initially, the program was developed to run in the terminal. A client and server program was created for establishing proper connection between the user and the backend. The query provided by the user is processed by another module. The training part of the chatbot was written in the form of another module.

Once the functional and non-functional testing is done; the product is deployed in the customer environment. There are some issues which come up in the client environment. To fix those issues, patches are released. Also to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment. This is the last step carried out in the project development life cycle where the product will be deployed on the college website and will be maintained occasionally for bugs and reports.

Reasons for selecting Incremental Model

The advantages of incremental development are that it allows for departmentalization and control. A schedule can be set with deadlines for each stage of development and a product can proceed through the development process model phases one by one. Development moves from concept, through design, implementation, testing, installation, troubleshooting, and ends up at operation and maintenance. Each phase of development proceeds in strict order.

Some of the major advantages of the Incremental Model are as follows:

- Generates working software quickly and early during the software life cycle.
- More flexible and less costly to change scope and requirements.

- Easier to test and debug during a smaller iteration.
- Works well for smaller projects where requirements are very well understood.
- Easier to manage risk because risky pieces are identified and handled during its iteration.
- Easy to arrange and manage tasks.
- Process and results are well documented.

CHAPTER 4

RESULTS AND DISCUSSION

The results achieved by the project meets the requirements of the project. Chatbot is implemented to solve the academic related queries of the users. The chatbot uses parts of speech and bag of words to find replies for queries regarding Sree Chitra Thirunal College Of Engineering. The algorithm checks for the keywords that are commonly used and the important terms that will help the bot to develop the solution of the query that is given by the user.

4.1 Accuracy Estimation

There are two competing concerns: with less training data, the parameter estimates have greater variance. With less testing data, the performance statistic will have greater variance. The data should be divided such that neither variance is too high, which is more to do with the absolute number of instances in each category rather than the percentage. The neural network that was developed for the model consists of various metrices that are to be used to estimate the accuracy and loss of the model. The various factors that affect are number of neurons in the model, the number of epochs used to train the model, the batch size of input provided and the error threshold value that is set. Here, epoch defines the number of times the training data has to be passed through the network, batch size defines the number of inputs that is processed at a time by the neural network and error threshold defines the probability value below which model prediction error increases.

The model was trained using 8 neurons in hidden layer and 16 neurons in the hidden layer separately and the variations in the values of loss and accuracy were checked.

Number of epochs	loss	Accuracy
1000	2.51	0.650
2000	2.58	0.651
4000	3.25	0.517

Table 4.1: Eight neuron layer network for error threshold 0.02

Number of epochs	loss	Accuracy
1000	2.06	0.78
2000	3.52	0.50
4000	2.49	0.70

Table 4.2: Sixteen neuron layer network for error threshold 0.02

It was observed that on changing the epoch values for 8 neuron layer network, the loss increases and accuracy decreases as shown in Table 4.1. This proved that a eight neuron network is not suitable for carrying out the model training. Thus, we carry out the same tests on the sixteen neuron layer network to get the outputs as shown in Table 4.2. It was observed that when the number of epochs was made 1000 the loss obtained decreased with a considerable increase in accuracy.

Table 4.3 and 4.4 gives the values of loss and accuracy when the error threshold considered was 0.01 and the considered batch size of inputs was 200. In the case of 8 neuron network it is observed that accuracy obtained is less. As per Table 4.4, it is seen that the model underfits data for 2000 epoch values and overfits data for 4000 epoch value. After estimating accuracy and loss for various values the best metric combination was chosen as epoch value 1000 for the amount of dataset the model is to be trained on.

Number of epochs	loss	Accuracy
1000	3.06	0.650
2000	2.81	0.651
4000	3.11	0.517

Table 4.3: Eight neuron layer network for error threshold 0.01

Number of epochs	loss	Accuracy
1000	2.7	0.71
2000	3.04	0.51
4000	2.5	0.68

Table 4.4: Sixteen neuron layer network for error threshold 0.01

Amongst the total instances, a 80-20 split validation was carried out and it was seen that for every test set of queries provided each time as input there were correct responses provided 70% of the times. While the deep learning neural network was trained, it was seen that during each training there was a loss of 2.06 and an accuracy of 0.78.

CHAPTER 5

CONCLUSION AND FUTURE WORK

The main objectives of the project is to develop a chatbot that will be used to identify answers for user submitted questions. To develop a database where all the related data will be stored and to develop a web interface. The web interface developed will be for simple users. A background research including an overview of the conversation procedure and any relevant chatbots available was done. A database was developed, which stores information about questions, answers, keywords and feedback messages. A usable system shall be designed, developed and deployed to the web server. An evaluation of data collected by potential students of the college was carried out. Also after receiving feedback from the first deployment, extra requirements shall be introduced and implemented. Chatbots are used less in the academic arena thus it needs to be incorporated more into websites of colleges so that information can be passed on to the users without them having to go though the entire websites. It helps the student to get information regarding college without going through the website content and thus making human work less. It reduces tedious work that usually a user would have to carry out if they need to get any information from the website of the college. As there is no manpower employed in case of chatbot usage, it makes it easy for the users to get the required results of the queries. Information regarding activities happening inside campus and other academic information provided. The details about the college have been provided in the dataset which is used to find the required information for the query that is asked by the user.

REFERENCES

[1] Thomas N T Amrita (2016) quot;An E-business Chatbot using AIML and LSAquot; Intl. Conference on Advances in Computing, Communications and Informatics.

[2] Bayu Setiaji, Ferry Wahyu Wibowo (2016) quot;Chatbot Using A Knowledge in Database Human-to-Machine Conversation Modelingquot; International Conference on Intelligent Systems, Modelling and Simulation.

[3] A M Rahman, Abdullah Al Mamun, Alma Islam (2016) quot;Programming challenges of Chatbot: Current and Future Prospectivequot; IEEE International Islamic University Chittagong.

[4] Kane, D. A. (2016) "The Role of Chatbots in Teaching and Learning".

[5] Luger, E., Sellen, A. (2016) "Like Having a Really Bad PA: The Gulf between User Expectation and Experience of Conversational Agents" CHI Conference on Human Factors in Computing Systems (pp. 5286-5297).

[6] Luis Fernando D'Haro and Rafael E. Banchs (2017) quot;Learning to predict the adequacy of answers in chat oriented human agent dialogsquot; IEEE Region 10 Conference (TENCON), Malaysia.

[7] C.B.Lee, H. N (2017) quot;Chatbots and Conversational Agents: A Bibliometric Analysisquot; by, IEEE International Conference on Industrial Engineering and Engineering Management.

[8] Bhavika R. Ranoliya Nidhi Raghuwanshi, Sanjay Singh Department of Information and Communication Technology Manipal University (2017) quot;Chatbot for University Related FAQsquot; International Conference on Advances in Computing, Communications and Informatics.

[9] Eric Grigori (2017) "Evaluation of Modern Tools for an OMSCS Advisor Chatbot" Georgia Institute of Technology Atlanta.

CHAPTER 6

APPENDIX

The following are some samples of queries that are bot replies to :

User: Hi

Chatbot: Hey, I am here to solve all your queries.

User: Does the college have accreditation?

Chatbot: Yes, the college is accredited by National Board Of Accreditation

User: Is SCT better than CET?

Chatbot: No doubt :P

User: What is cult a way?

Chatbot: Cult a Way is the annual technical and cultural festival of SCT. It is a four- or five-day event and almost all colleges in the state participate. The events includes technical workshops and cultural events like dance programs, treasure hunt, quizzes, fashion shows, etc.

User: Why should I join SCT?

Chatbot: It is one of the top 5 Engineering colleges in Trivandrum and the placement opportunities provided are good.

User: Does the college have any socities?

Chatbot: Yes. The college has a lot of clubs.

User: Does the campus have hostel facilities?

Chatbot: No, the campus does not have hostel facilities.

User: What is the menu of the canteen?

Chatbot: You can check with the canteen staff for any special dishes made today. On a regular basis meals are available for lunch.

User: Does the college provide duty leave?

Chatbot: Yes.

User: Where is the attendance wing?

Chatbot: First floor- administrative office of the college.

User: When will the attendance wing update the site?

Chatbot: Let's hope soon. XD

User: Do the teachers interact with students about their personal problems?

Chatbot: Yes. The teaching staff is very friendly and helping.

User: Can non Kerala students get admission?

Chatbot: Yes.

User: *Unkown question to the bot*

Chatbot: Sorry I am unable to respond to it.

User: Bye

Chatbot: Thanks for visiting.