

## SANSKRUTI M. JAJOO

Module : Human – AI Interaction (Coursework 1 Report)  
Project Name : Educational Chatbot - Your Peer and Your Mentor  
Student Id : 20370257

### Core Idea of the project:

It is often observed in the education ecosystem that some students are hesitant to discuss concepts and doubts with their teachers. On the other hand, students find it easier and more comfortable to study with their fellow students. This situation has its obvious disadvantage that the fellow students are not the subject matter experts and also lack experience. Therefore, the progress of the students also gets hampered. So, here comes the need of an intelligent agent, who can play both the roles.

This forms the core idea of this project. The automated, intelligent tutoring system, which provides a dedicated learning environment for students to study in and then analyses their reactions and patterns of engagement with artificially intelligent learning content, is the most valuable use of AI. Similarly, AI chatbots can assist students in learning through a sequence of messages based on a study material.

Major questions that need to be answered from this project are:

1. How well a chatbot can help a student to improve learning progress in fact-based subjects?
2. How chatbot lowers the burden of faculty?

### Background and Relevant ideas:

Quiz generation via Over generating Transformations and Ranking outlined in the paper [4] can be used for the future advancements of **Max**. The paper [3] proposes the question answering and quiz generation from the content uploaded by user. It uses Inverse Document Frequency scoring to get similarity between the questions where **Max** is developed on term frequency as well.

### Description:

The idea of chatbot giving the feeling of human, came up with the common cool name of the chatbot as "**Max**". "**Max**" is made of a happy encouraging tone and with a sense of humor to act as a friend.

In this project the "**Max**" can perform the following educational functions:

1. **Discuss** a random topic in a conversational form.
2. Show the **score** of the answers given by the user for the asked question.
3. **Book a test** by asking the details from user.
4. Encourage user on getting less marks or on predicting the emotion as sad with **quotes**.
5. Help to relax the user with a **joke** or playing game "**Guess Who?**".
6. **Small Chat** with user like a colleague.
7. **Clear doubts** of student by answering the questions.

### Design:

#### Tools and Technology used:

1. Python: A highly productive general purpose computer programming language.
2. Jupyter Notebook: A server client application where we can perform different actions on the programming code and is useful for analyzing the data.
3. Python libraries: Collection of methods and functions in python programming language reducing the user effort and extra code lines.

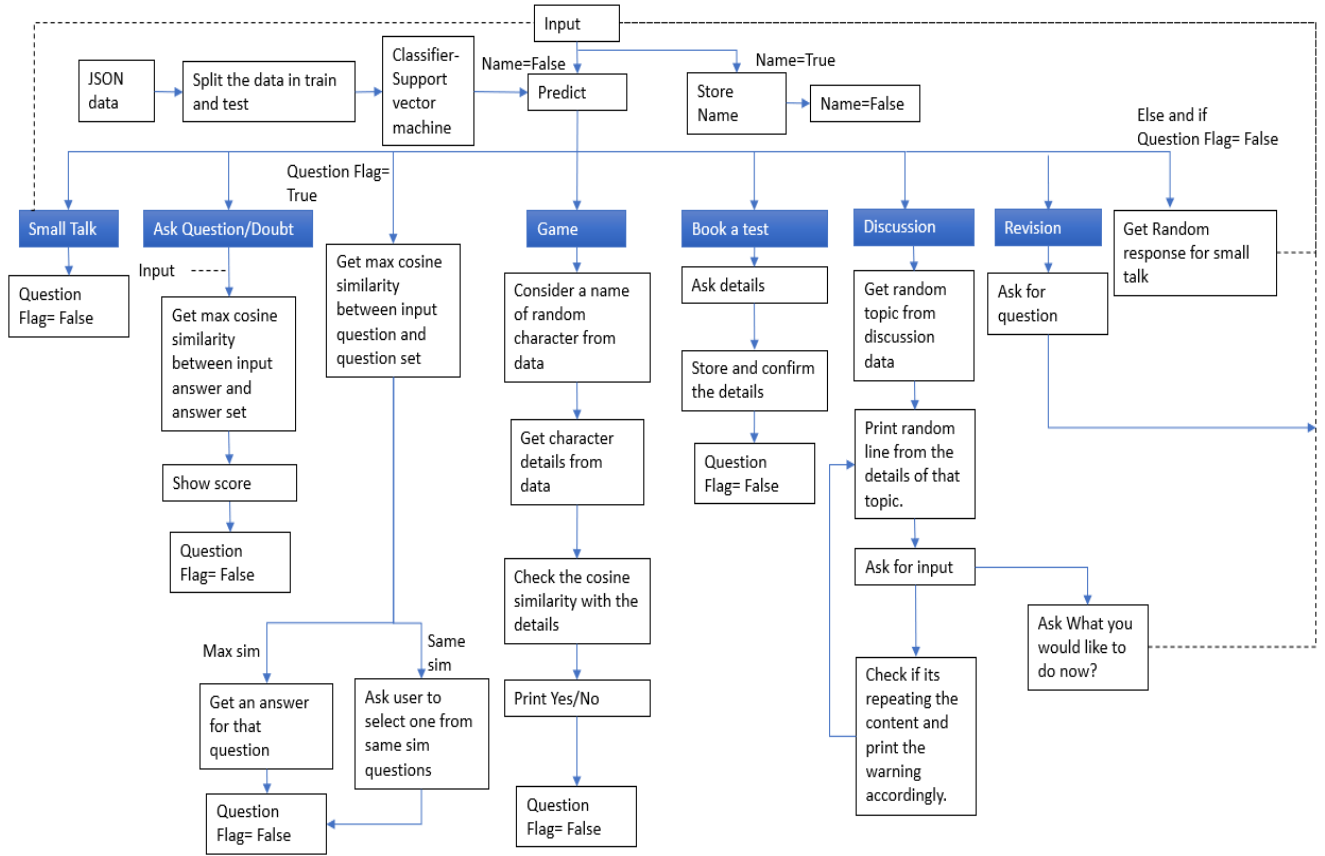


Figure 1: Design architecture of chatbot

**Implementation and challenges:**

	<b>Feature and Application</b>	<b>Techniques</b>	<b>Steps</b>	<b>Challenges</b>
1	<i>Intent matching</i>  Predicting the intent of the user and selecting the other feature applications based on prediction.	Support vector machine	1.Create a .json file with tags, patterns and responses. 2.Split the data of tags and patterns in intent.json file into train and test data using train_test_split (). 3.Apply Support vector machine on training data to create a model classifier. 4.Use predict function of classifier to predict the tag of user response.	1.Increase the accuracy of classifier. 2.Decide the train:test ratio.
2	<i>Identity management</i>  Try to detect name of the user from a user response to use it in conversation.	Parts of speech tagging	1.Use pos_tag of WordNetLemmatizer. 2.Get the Noun. 3.If there are no nouns then use the string as a name.	1.Tags changes with sentences. 2.User can give any random string.
3	<i>Small talk</i>  User can greet, discuss the emotions, ask for jokes.	Support vector machine	1.Predict the user response using intent.json patterns and Support vector machine as applied in Intent Matching feature. 2.Give the random response from the corresponding responses in json file.	1.Increase the accuracy of classifier. 2.Decide the train:test ratio.
4	<i>Transactions</i>  User can book a test. Max will ask more details.	Support vector machine	1.Predict the user response for booking the test using intent.json patterns and Support vector machine. 2.Get the answers for the time, no. of questions, marks for each question, date and topics to cover from user. 3.Store the main information using regular expression and confirm it from user.	

5	<p><i>Information retrieval &amp; question answering</i></p> <p>User can ask questions and get the answers from Max. If questions are asked by Max, then Max can check the accuracy of answers given by user.</p>	Cosine similarity	<ol style="list-style-type: none"> <li>1.Create vocabulary from tokenized stemmed words of the sentences.</li> <li>2.Create bag of words for the users question and the questions in the data.</li> <li>3.Calculate the cosine similarity between the two vectors</li> <li>4.Return the question with maximum similarity.</li> <li>5.Print the corresponding answer.</li> </ol>	Max similarity answer may not be the correct answer for that question.
6	<p><i>Discussion of any topic.</i></p> <p><b>Max</b> will say some random lines and then user will have to say some lines on the same topic.</p>	<p>Cosine similarity</p> <p>Jaccard similarity</p>	<ol style="list-style-type: none"> <li>1.Get the random topic from data.</li> <li>2.Print random line from topic details.</li> <li>3.Ask for user input.</li> <li>4.Check the cosine similarity with the previously used lines and the details for that topic.</li> <li>5. Give warning if needed on repeating the lines or giving unrelated line.</li> </ol>	
7	<p><i>Game playing</i></p> <p><b>Guess who?</b> is the game can be played with <b>Max</b>. In this user will have to guess the name of the character from the "Harry Potter" series. User can ask 'Yes or No' questions to Max.</p>	Cosine similarity on bag of words/vectors	<ol style="list-style-type: none"> <li>1.Create vocabulary from tokenized stemmed words of the sentences.</li> <li>2.Create bag of words for the users question and the concatenated details of the character in the data.</li> <li>3.Calculate the cosine similarity between the two vectors.</li> <li>4.Return the answer as Yes or No on the basis of similarity.</li> <li>5.Give the hint in the format of blanks.</li> </ol>	

**Evaluation:***User Testing:*

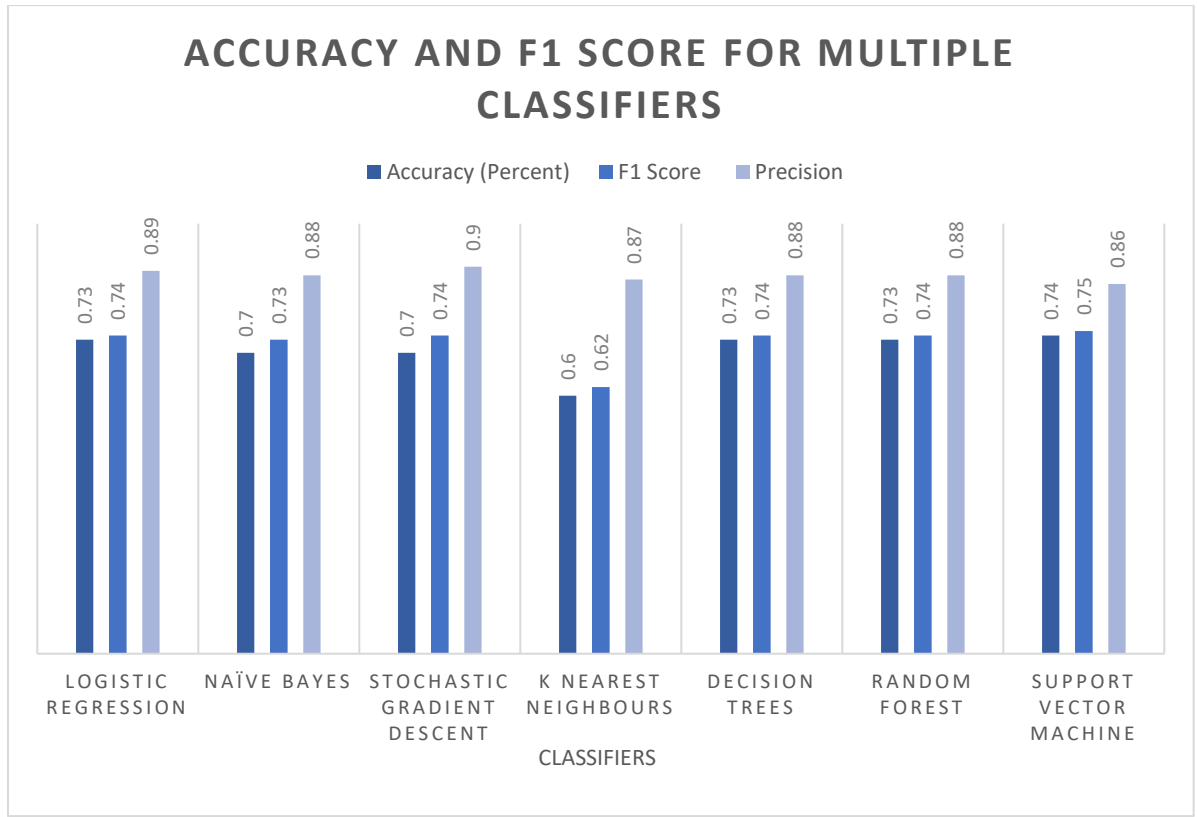
Requirements	Actions	Proposed Improvement	What went wrong?	Reason
1. Testing my understanding on some topics.	"I start the chatbot say Hello to it. Then I asked him to ask me few questions. I answered those questions and at the end I checked my score report."	"Give more detailed report"		
2. Revise few answers before the test. Then book a test.	"I started the chatbot. I asked few questions to revise. After getting answer I asked the chatbot to book a test. I gave the details asked by him. Then I closed the chat by saying bye."	"Add negative markings"	Got wrong answer for 'Name any cast of harry potter movie.'	Cosine similarity of Harry Potter Movie was greater for question 'how much are the harry potter movies worth'
3. Test my answers for few random questions. Then play a game after test to feel relax.	"I asked <b>Max</b> to ask me some questions. I answered it and then at the end I got my report. I got less marks so <b>Max</b> motivated me. After that I asked him to start the game. I closed the chat after playing game."		Gave wrong answer in game for the details not present in the data.	Lack of data
4. A small chat and then discuss any topic related to studies like usually do with colleagues."	"I gave my name. Then had some small funny talk like I used to do with my friends. Then we discussed one random topic he asked me to. Then I closed the chat"	"Discuss the topic which chosen by user"		Lack of data results in not fulfilling the request for discussing the topic selected by user.

System Usability Score out of 10 from users:

Factor	1	2	3	4
<b>Conversation Flow</b>				
Understanding of question	9	7	8	9
Instant Response	9	6	9	10
Relevant Answers	7	7	9	9
Chatbot Engagement to continue the chat	8	9	9	8
Intelligence to understand intent	9	8	9	10
Onboarding	8	9	8	9
Persona	10	9	10	10
<b>Business specific</b>				
Confusion Handling	7	8	9	7
Accuracy	7	6	9	7
Speed	10	8	9	10
Troubleshooting	7	9	9	8
Navigation from one feature to other	9	9	9	9
<b>Average</b>	<b>8.33</b>	<b>7.81</b>	<b>9</b>	<b>8.8</b>

**Experimentation and results:**

## 1. Comparison of results from multiple classifier



## 2. Similarity

User question	Jaccard	Euclidean	Cosine	Manhattan
Data Question				
what is the location of great basin?	0.5	1	0.66	0.33
where is the great basin located on a us map?				
who own the Biltmore estates in nc?	0.6	1	0.77	0.33
what is the name of the family who own the Biltmore estates in nc?				
who was the congressman who was caught with an escort in ny?	0.75	1	0.87	0.5
who was the congressman who was caught with an escort in ny?				



**Discussion:**

It can be concluded that cosine similarity can be used to calculate the similarity between user's query and questions to return the related answers. An analysis of the graph illustrates that support vector machine classifier can be used to predict the intent of the users' response. Higher values of accuracy and F1 score supports the fact that most of the time it is effective.

Reflection:

Reasons for failure of **Max** in achieving goal accurately:

1. Uncertainty of user conversation
2. Lack of chatbot strategy.
3. Insufficient training data results in intent classification failures.

Deploying this chatbot in real world has both positive and negative impacts:

Impact:

Positive:

1. Student will be able to clear the doubts instantly from Max in more depth on the basis of FAQs.
2. Provide immediate feedback on errors.
3. Lower the burden of faculty as there is no longer a need to explain same thing again and again.

Negative:

1. Threat to the teaching assistant job as it changes the way educator works with the student.
2. Human error done by database editor can make negative impact on understanding of subject to student.

Bias and Fairness:

1. Discussion of any topic can be biased because of the content writer.
2. **Max** can only communicate in English language.

**Future work:**

1. The aim of clearing the doubts of students can be achieved by recording a history of FAQs and the explained correct answers for the same.
2. The questions and answers can be generated from the text content of the topic.
3. The test could be generated to certain level of difficulty with the help of difficulty indicator stored for questions in the database.
4. The chatbot will go through the process of analysing each student's level of knowledge and then provide them with the next parts of the lecture based on their progress.
5. In the discussion, chatbot should be able to compare the similar meaning sentences.

### **Conclusion:**

The aim of this project is to develop a chatbot helping students to improve the understanding of the fact-based subjects. This can be achieved by some educational functionalities performed by the chatbot like discussion on topics and question answering. Instant feedback of errors helps the user to deepen their understanding. The intent matching to predict the requirements of user is performed by Support vector Machine with accuracy score of 74% and F1 score of 0.75. This helps the chatbot to act as a friend and make user comfortable. With some advancement, this chatbot can be deployed in real world to benefit educational sector.

### **Reference and citations:**

- [1] Abbasi S and Kazi H. 2014. Measuring effectiveness of learning chatbot systems on Student's learning outcome and memory retention Asian Journal of Applied Science and Engineering, 3, 57-66.
- [2] [Chatbot-using-Natural-Language-Processing/chatapp.ipynb at master · SamiulGitHubUser/Chatbot-using-Natural-Language-Processing · GitHub](#)
- [3] Sreelakshmi A.S , Abhinaya S.B., Aishwarya Nair and Jaya Nirmala S. 2019.A Question Answering and Quiz Generation Chatbot for Education.
- [4] Heilman, M., Smith, N. A. (2010, June). Good question! Statistical ranking for question generation. In Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics (pp. 609-617). Association for Computational Linguistics.