

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 2, February 2019

A Smart Chatbot for Education System

Divyani Sejkar¹, Sheetal Jagdale², Ritu Punamiya³, Pushpanjali Ghadge⁴, Nivedita Bhirud⁵

U.G. Student, Department of Computer Engineering, Vishwakarma Institute of Information Technology,
Kondhwa, Pune, India¹

U.G. Student, Department of Computer Engineering, Vishwakarma Institute of Information Technology,
Kondhwa, Pune, India²

U.G. Student, Department of Computer Engineering, Vishwakarma Institute of Information Technology,
Kondhwa, Pune, India³

U.G. Student, Department of Computer Engineering, Vishwakarma Institute of Information Technology,
Kondhwa, Pune, India⁴

Associate Professor, Department of Computer Engineering, Vishwakarma Institute of Information Technology,
Kondhwa, Pune, India⁵

ABSTRACT: Chatbots are programs that communicate through text with human using Artificial Intelligence. It is used in many areas like entertainment, Business, education. Chatbot has become more popular in business right now as they reduce service cost and handle multiple users at a time. The system developed has been made to identify the question and make response to answer the question. Chatbots are completely text based user interface. The designed Chatbot will provide efficient answer for any question using Artificial Intelligence and Latent Dirichlet Allocation. The user can ask question related to college activities through the chatbot from anywhere. The system analyses the question and answers to the user. The user just has to register himself to the system. In chatbot Natural Language Processing technologies are used for parsing, tokenizing, stemming and filtering the content of questions.

KEYWORDS: Latent Dirichlet Allocation [LDA], Artificial Intelligence [AI], Natural Language Processing [NLP].

I. INTRODUCTION

Chatbots are the kind of program that are generally used to simulate user conversation. During the admission process or even after admitting into a college students have various doubts regarding college activities. It is not feasible for the students to visit college help desk to answer its doubts. Chatbot for education system will be developed using Artificial Intelligence techniques such as Natural Language Processing (NLP), Latent Dirichlet Allocation algorithm that will communicate with the students on the real time basis by clearing the rings of doubts of students. The system will be web application which will provide the appropriate responses to the user's queries. The responses will be generated using Artificial Intelligence algorithms. This will avoid the need of user to go physically to the college for inquiry. The user will first have to register himself into the system before starting the chat. User can ask questions related to various college activities such as locations of different departments, queries related to student, section, exam section, account section, Placement activities and other cultural activities. The system will reply to user with the help of Graphical User Interface (GUI). It will help the students to be updated about the college activities.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 2, February 2019

II.RELATED WORK

2.1 Paper Name: ELIZA—a computer program for the study of natural language communication between man and machine Authors: Joseph Weizenbaum.

Description: ELIZA is basically a program which allows conversation with humans in natural language. Based upon the decomposition rules the input sentence is analysed which is triggered by the keywords in the sentences. With the help of the decomposition rules which are associated with the reassembly rules responses are generated. It recognizes the keywords in the sentence and using these keywords it produces a response from the predefined responses. During the actual working of ELIZA, the words which are established for every scan of input, is applied by the hashing algorithm which points to some list structure which could contain that word as a keyword. However it faced with some technical problems like recognizing the most important keyword in the sentence and also the problem to choose the appropriate transformation rule.

2.2 Paper Name: The Anatomy of A.L.I.C.E. Authors: Richard S. Wallace.

Description-ALICE is basically the program that uses some pattern matching rules to analyse user queries. Artificial Intelligence Mark-up Language is used to store the data. Alice produces the responses based on a rule which matches the input patterns to the output text. The Chabot engine and language knowledge model are different in the ALICE architecture. In this AIML consists of AIML objects which are data objects which consist of topics and categories kind of units. Categories are nothing but the rules which are used for matching the input text and then converting it into a pattern and the Alice response or the output is considered as the template. The pattern matching technique is used to recognize and search the longest pattern.To perform the word to word match the AIML interpreter is used which searches for the best matching pattern. The output is then built by the interpreter when the match is detected.

III.EXISTING SYSTEM

In the existing system the students need to always visit the institute to clear their doubts. This traditional method of students visiting the college help desk is not effective as it is not always possible for the students to reach the college and also it is not very effective for the college help desk to have one to one conversation with all the students to answer their queries. Also the traditional method does not allow the students to communicate with the faculty everytime.Sometimes not all the information is available on the institute website which creates a major drawback for the students to get their doubts cleared.

3.1 Disadvantages of Existing System

1. Time consuming process.
2. Not very flexible as students can't reach the help desk every time.
3. Not very efficient.

IV.PROPOSED SYSTEM

In our system we basically provide a web interface to the students to enter their queries. The commands or the queries entered by the user are in the natural language. The system then processes the input entered by the user in natural form. It performs tokenization where the input text entered is broken down into various tokens. The various stop words like "a","an","the" etc are removed. It then performs Stemming by converting the words into their base forms or root forms which helps in better classification of text. After the pre-processing of the data it tries to recognize the intent and the entity and finds the topics using LDA Algorithm. It is basically a topic modelling algorithm which is used to identify the topics in the sentence. It works on the principle of computing the probability of words in the topic and the probability of topic in the sentence. Based on the probability it reassigns the topics to the words. After mapping the

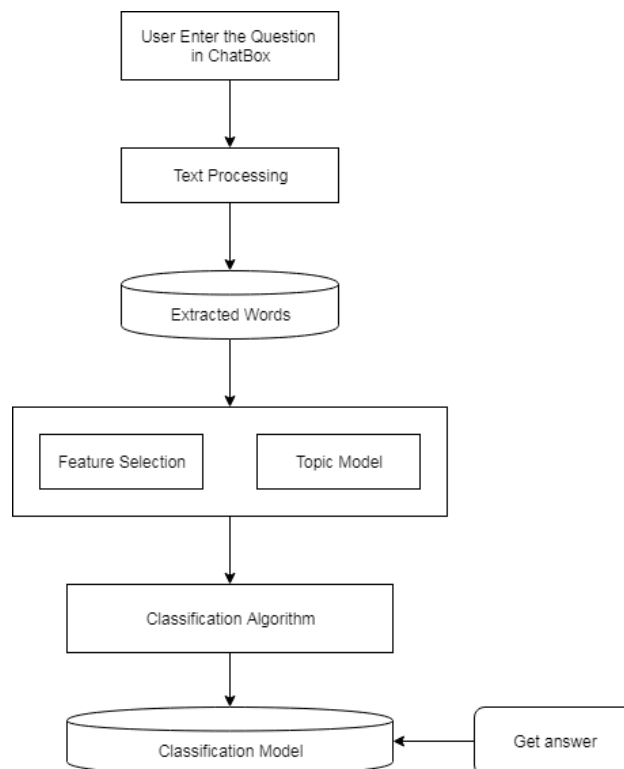
International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 2, February 2019

sentences1 to the topics it fetches the answer from the knowledge base having the highest probability and provides it as the response to the users.



1. User login and enter the question:

- The user has to firstly register himself into system. After registration user can login into system using valid credentials and can start entering the question related to the college activities.

2. Generic Question Construction:

- After the user login in the system, user starts querying through the interface provided. On receiving these queries the system executes the LDA algorithm to identify the topics. Based on the topic identification it provides the appropriate response to the user.

3. Responding Chatbot:

- After receiving the queries it performs the natural language processing.
- It performs the tokenization of the input text by breaking the sentence into various tokens.
- It then removes the stop words encountered in the system like “a”, “an”, “the” etc.
- It then reduces the text by converting the words into base words or root words which is stemming.
- Topic modelling is performed to extract the various topics from the sentence. Topics are nothing but words and various words are clustered to form topics.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 8, Issue 2, February 2019

4. Generic response Constructor:

- This module converts the response fetched from the knowledge base back into the natural language understandable by the users. The system on finding the appropriate response provides it to the user through the interface provided.

5. Questions directing to Authorized Users:

- In case if the response provided by the system is not fruitful, this unit allows the users to ask the questions back to some authorized users in the system which could provide the appropriate response

V. CONCLUSION

We proposed the smart chatbot for education system using AI mechanism and LDA algorithm for topic modelling. This system is basically a web application where user will start the chat by asking textual questions through the interface. The knowledge base is maintained which is updated everytime the user ask the question. The bot is initially trained to increase its efficiency. Hence it is more intelligent to answer the complex queries asked by the students. It removes the drawback of time consumption of the traditional method as it can handle simultaneous queries related to the students and can provide quick responses.

REFERENCES

- [1] Jincy Susan Thomas, "Chatbot using gated end-to-end Memory Networks" International Research Journal of Engineering and Technology, Vol.05, pp.2395-0072, Mar-2018.
- [2] N.T Thomas, "An e-business chatbot using AIML and LSA", International Conference on Advances in Computing, Communication and Informatics (ICACCI), Jaipur, 2016, pp.2740-2740-2742. DOI:10.1109/ICACCI.2016.7732476.
- [3] Aditya Deshpande, Alisha Shahane, Darshan Gadre, Mrunmayi Deshpande, Prof Dr. Prachi Joshi, "A Survey of various Chatbot Implementation Techniques", International Journal of Computer Engineering and Applications, ISSN 2321-3469.
- [4] Francesco Colace, Francesco Pascale, Antonio Ferraioli, Luca Garofalo, Saverio Lemma, Marco Lombardi, "A Conversational Workflow Model of Chatbot", DOI-10.18293/SEKE2017-208.
- [5] Luke K Fryer, Mary Ainley, Andrew Thompson, Aaron Gibson, "Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners", DOI: 10.1016/j.chb.2017.05.045.
- [6] Rishabh Upadhaya, "Chatbot platform as command & control channel in botnet (a proof of concept)", DOI: 10.13140/RG.2.2.12489.34401
- [7] Bayu Setiaji, Ferry Wahyu Wibowo, "Chatbot using a Knowledge in the database", IEEE Conference on Intelligent Systems, Modelling and Simulations DOI 10.1109/ISMS.2016.53.
- [8] Weizenbaum Joseph, "ELIZA-a computer program for the study of natural language communication between man and machine", New York: W.H. Freeman and company, pp.2,3,6,182,189. ISBN 0-7167-0464-1.
- [9] Prof.D.S.Thosar, Varsha Gothe, Priyanka Bhorkade, Varsha Sanap, "Review on Mood Detection using Image Processing and Chatbot using Artificial Intelligence", International Research Journal of Engineering and Technology, Vol.05, pp.2395-0072, Mar-2018.
- [10] P. Anupriya, S. Karpagavalli, "LDA based topic Modelling of Journal Abstracts", International Conference on Advanced Computing and Communication Systems [2015].
- [11] A M Rahman, Alma Islam, "Programming Challenges of Chatbot: Current and Future Prospective", IEEE Conference, DOI.10.1109/R10-HTC.2017.8288910.
- [12] Boris Galitsky, Dmitry Ivoysky, Sergei O. Kuznetsov, Fedor Strok, "Matching sets of parse trees for answering multi-sentence questions", RANLP, pages 285-293.
- [13] Ganguly D, Roy D, Mitra M, "Word Embedding based Generalized Language Model for Information Retrieval", Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '15. <http://doi.org/10.1145/2766462.2767780>.