

Self-Learning Conversational AI Chatbot Using Natural Language Processing.

Mrs. P. Elakkiya¹, M. Varun², P. S. Mohamed Riswan Maliq³, S. Pragadeesh⁴

Assistant Professor, Department of Computer Science and Engineering¹

Students, Department of Computer Science and Engineering^{2,3,4}

Anjalai Ammal Mahalingam Engineering College, Thiruvavur, India

elakkiya@aamec.edu.in and mohamedriswanmaliq@gmail.com

Abstract: ChatBot has the capability to recognize the specific domain of any query that is posted. Cosine Similarity algorithm is applied here to find the right answer to the user query. The bot itself is intelligent enough to identify the frequently asked unanswered question and notify the same for admin feedback. Once the admin provides a response for it, the bot is enhanced intelligent to look for new set of questions and answers and respond the same to the users in the future. This operation can be configurable in such a way that the system can decide the threshold limit for the unanswered question.

Keywords: Natural Language Processing, Cosine Similarity Algorithm, Artificial Intelligence, User Interface, Machine Learning

I. INTRODUCTION

A chatbot is a piece of software that assists in the natural development of a conversation with the user Artificial intelligence has become increasingly complicated as information technology and communication have advanced. Human acts such as taking a picture are used by AI systems. making a decision at a specific time, executing day-to-day chores, responding to people swiftly, and solving problems making a decision at a specific time, executing day-to-day chores, responding to people swiftly, and solving problems with the internet It's a highly effective way to handle an benefit from everything that's just outside your door. The chatbots are good enough to trick users into thinking they're talking to a human, but they have a limited knowledge base at runtime and no way of keeping track of all the discussions.

Chatbots employ machine learning to assist AI in understanding user queries/doubts and providing an appropriate response to the user. For conversing or engaging with the user, they are created utilising the Artificial Intelligence Markup Language. Answering engines are another name for chatbots. Because the knowledge has already been programmed in advance, this application works in a very straightforward manner.

Mining are some of the approaches employed in the application chatbot's knowledge, which has been gathered from a variety of sources.. The chatbot compares the user's supplied sentence to an existing pattern in the knowledge base. Each pattern is compared against the chatbot that receives questions from users, tries to understand the question, and provides appropriate answers. It does this by converting an English sentence into a machine-friendly query, then going through relevant data to find the necessary information, and finally returning the answer in a natural language sentence. The bot itself is intelligent enough to identify the frequently asked unanswered question and notify the same for admin feedback. Once the admin provides a response for it, the bot is enhanced intelligent to look for new set of questions and answers and respond the same to the users in the future.

II. LITERATURE REVIEW

Many programmes consolidate a human appearance and attempt to replicate human communication, but in the vast majority of situations, the data used for bot conversation is stored in a database established by a human specialist. We may create several types of chatbots using AI; for example, in this work, we created a Conversational chatbot ForCollege Enquiry. Enquiry process, Fees structure, Course information, Eligibility criteria description, and Admission are only a few of the fields. This research shows how we might deal with identifying the most important facts in writings describing the life of an authentic figure in order to create a conversation operator that could be used in middle school CSCL scenarios.

A chatbot, according to Benton and Radziwill (2017), is an online medium for engaging with humans when they are actually conversing with computer software that is brought to life by natural language input. Others define it as a computer software that uses artificial intelligence to emulate human conversation. According to Schlarl (2004), a chatbot is software that allows textual dialogue using natural language. Users find it difficult to understand that the chatbot is not a real person, which emphasises the crucial need of a vast knowledge base, which is the existing set of rules a chatbot possesses. (Scharl, 2004). Chatbots will soon become one of the most effective ways for businesses to communicate with individual customers and swiftly resolve their issues. Moore (Moore 2017). Furthermore, the recent advancements in Artificial Intelligence and major developments in messaging services have been heavily credited to the recent Chatbots have piqued people's interest (Guzman&Pathania,2016) Chatbots exist in task-specific apps and duplicating a human dialogue to be educational, conversational, or based on esteem.

The Bot Who Started a Thousand... Other bots include:

ALICE, one of the very first bots to go online – and one that has kept up exceptionally well despite being built and launched more than 20 years ago – would be missing from any list of pioneering Chatbots. Dr. Richard Wallace created and launched ALICE – which stands for Artificial Linguistic Internet Computer Entity, an acronym that could have come straight out of an episode of The XFiles – in the early days of the Internet in 1995. (As you can see in the image above, the International Research Journal of Engineering and Technology website's aesthetic has remained practically constant since then, serving as a powerful reminder of how far web design has progressed.) Despite the fact that ALICE is based on an outdated codebase, the bot provides a remarkably accurate conversational experience to its consumers. Of course, no bot is flawless, even one that is old enough to drink legally in the United States if it only had a physical form. ALICE, like many current bots, struggles with the complexities of some queries and responds with a combination of unwittingly postmodern replies and remarks that show ALICE has higher self-awareness, for which we should thank the agent. Despite its flaws, none of today's chatbots would be conceivable without Dr. Wallace's revolutionary work. Wallace's bot was also the model for the companion operating system in Spike Jonze's 2013 science-fiction romance film Her.

The Latent Dirichlet Allocation (LDA) technique was utilised in the operation of managing queries in a real-time inquiry at the University of Salerno for a group of Computer Science students. This resulted in a satisfying result. Rasika analysed existing chatbots such as Facebook Chat, Natasha from Hike, and Wechat, and used recurrent neural networks (RNN), pattern matching, Natural Language Processing (NLP), and data mining methods to try to design a superior performing system. The proposed work involves the approach of Artificial Intelligence and Cosine similarity algorithm to give appropriate result to users. The model gives accuracy result to users based on "THRESHOLD LIMIT". How much e optimise Trained dataset we increase the threshold limit. We set the Threshold limit from 0.1 to 1.0 based on accuracy.

III. METHODOLOGY

The training procedure is necessary because it improves the DataSet, which improves the responses that are crucial for subsequent processing. You will be required to submit a list of statements for the training process, with the order of each remark determined by its placement in a given conversation. Data from conversation transcripts is contained in a conversation dataset. This information is used to train a Smart Reply model that suggests text responses to human agents interacting with customers. Finally, we establish the frequency thresholds for each question.

A Conversational User Interface, or CUI, is a text-based interface that allows people and computers to communicate using natural language. Based on a large library of conversational patterns, language analysing software helps bots detect and interpret human interactions. CUI can help users interact better on the platform, through mobile apps, and even over the phone. Its main advantage is that it is simple to use, allowing users to ask for exactly what they want without having to memorise specific keywords or phrases. It has the effect of a one-on-one chat. A bot, on the other hand, can have thousands of conversations at once. Furthermore, the language processing technology that powers chatbots and voice interfaces is capable of learning and evolving with its users.

Classification module consists of several phases including Similarity check and Admin feedback of unanswered questions. In this case, the user's query is initially passed on to the NLP. The term that fits the dataset was then extracted

by NLP. Using the Cosine Similarity Algorithm, the term is matched with dataset questions, and the user is given the proper response. If Keyword does not meet a predetermined frequency criteria for any of the questions in the dataset. The bot provides the user with a default response. The query is then forwarded to the administrator for response. If the administrator authorises the query, it will be answered. Finally, the Bot's Trained Dataset is updated.

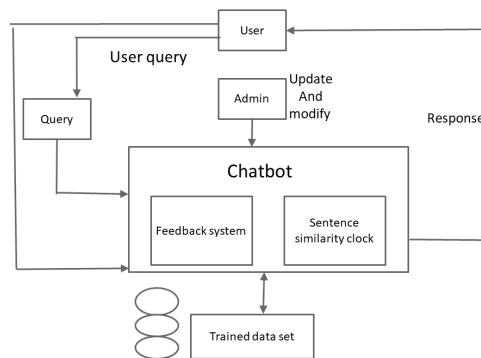
IV. ALOGRITHM

Cosine similarity algorithm is one of the popular algorithm of Machine learning .Cosine similarity measures the similarity between two vectors of an inner product space. It determines whether two vectors are pointing in the same general direction by measuring the cosine of the angle between them. In text analysis, it's frequently used to determine document similarity. Thousands of characteristics can be used to characterise a document, each of which records the frequency of a specific word (such as a keyword) or phrase in the document. As a result, each document is an object that is represented by a term-frequency vector.

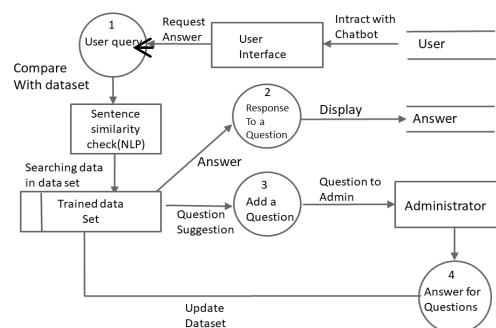
$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

Cosine similarity is a similarity metric that can be used to compare documents or, for example, to rank documents based on a vector of query words. Allow two vectors, x and y, to be compared. When we use the cosine measure as a similarity function, we get $\frac{x \cdot y}{\|x\| \|y\|}$, where $\|x\|$ is the Euclidean norm of vector, defined as It is the vector's length in terms of concept. Similarly, the Euclidean norm of vector y is $\|y\|$. The cosine of the angle between vectors x and y is computed by the measure. A cosine value of 0 indicates that the two vectors are orthogonal (at 90 degrees to each other) and do not match.

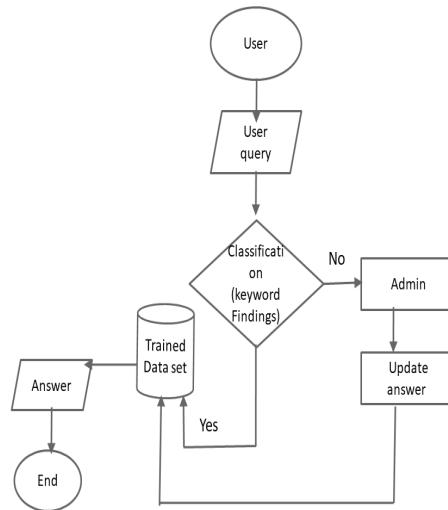
V. SYSTEM ARCHITECTURE



VI. DATAFLOW DIAGRAM



VII.FLOWCHART



VIII. FINAL PROTOTYPE

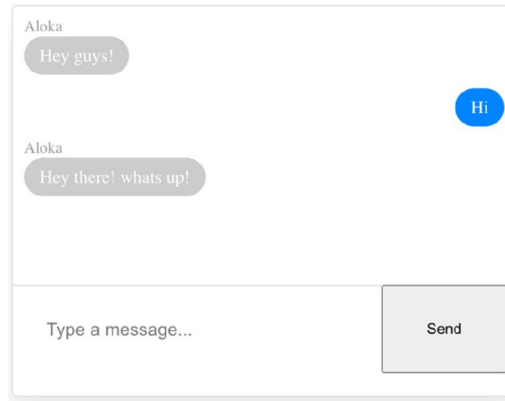
```

List<String> greetingList = new ArrayList<String>() {
    {
        add("hi");
        add("hello");
        add("hey");
    }
};
double thresholdValue = 0.5;
@RequestMapping(value = "/greeting", method = RequestMethod.GET)
public String getGreeting() {
    return "Hey there!!!";
}
@RequestMapping(value = "/getResponse", method = RequestMethod.POST)
public String getResponse(@RequestBody Map<String, String> requestBody) {
    String message = requestBody.get("message").toLowerCase();
    message = message.toLowerCase();
    if (greetingList.contains(message)) {
        return "Hey there! whats up!";
    }
    List<Map<String, String>> feedDataList = loadFeedData();
    double maxSimilarity = -1;
    String answer = "";
    if (feedDataList != null) {
        for (Map<String, String> data : feedDataList) {
            double current = similarity(message, data.get("question"));
            if (current > maxSimilarity) {
                maxSimilarity = current;
                answer = data.get("answer");
            }
        }
    }
    if (maxSimilarity >= thresholdValue) {
        return answer;
    }
    return "Sorry could not find anything";
}
@RequestMapping(value = "/getUnanswered", method = RequestMethod.GET)
public List<String> getUnanswered() {
    List<String> unanswered = null;
    try {
        File file = new ClassPathResource("unanswered.json").getFile();
        ObjectMapper objectMapper = new ObjectMapper();
        @SuppressWarnings("unchecked")
        Map<String, List<String>> map = objectMapper.readValue(file, Map.class);
        unanswered = map.get("unanswered");
    } catch (Exception ex) {
        ex.printStackTrace();
    }
    return unanswered;
}
  
```

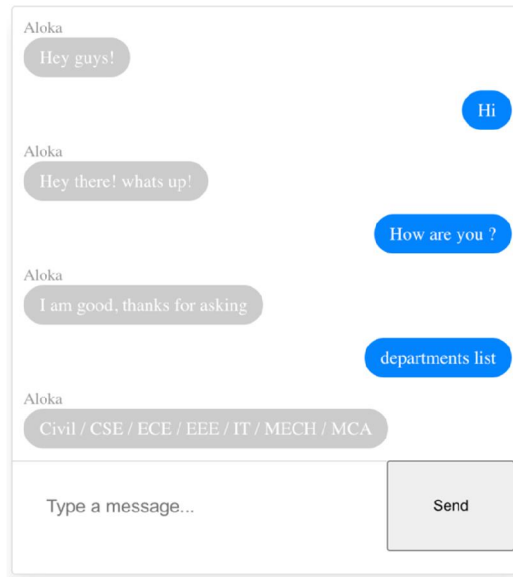
IX. RESULT

The proposed system was put to the test and proved to be effective and feasible. It saves manpower, time, and paper work for college administration. It also saves students the time and effort of driving all the way to campus for research. We designed a chatbot in this article that would communicate with users and deliver all college-related information. A chatbot connects the student/parent and the college administration. The admin will update any questions that the chatbot does not answer.

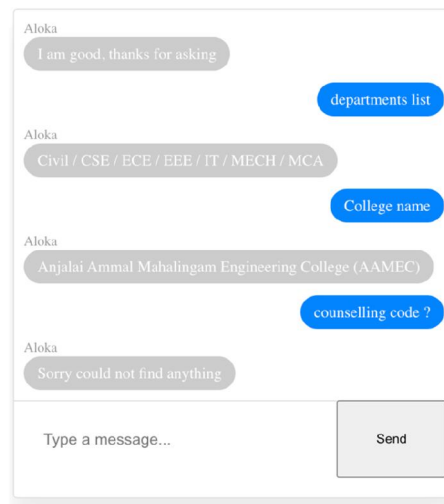
Greetings:



Successful answers:



Unsuccessful answer:



X. CONCLUSION

The major goal of this chatbot was to create an algorithm that could recognise user questions or queries and respond appropriately. To create a database in which all relevant data is saved and matched with inquiries when they arise. We successfully designed a chatbot that allows users to ask questions about the application process, course specifics, eligibility requirements, and admission. The chatbot analyses the question and responds appropriately. accuracy and has important practical application value.

REFERENCES

- [1] A. P. Chaves and M. A. Gerosa, "How should my chatbot interact? A survey on social characteristics in human-chatbot interaction design," *Inc. Hum.-Comput. Interact.*, vol. 37, no. 8, pp. 729–758, May 2021, doi:10.1080/10447318.2020.1841438.
- [2] M. Chung, E. Ko, H. Joung, and S. J. Kim, "Chatbot e-service and customer satisfaction regarding luxury brands," *J. Bus. Res.*, vol. 117, pp. 587–595, Sep. 2020
- [3] Ms. Ch.Lavanya Susanna and R. Pratyusha, "COLLEGE ENQUIRY CHATBOT" in *International Research Journal of Engineering and Technology (IRJET)* on 3rd March 2020.
- [4] S. Pérez-Soler, E. Guerra, and J. de Lara, "Model-driven chatbot development," in *Conceptual Modeling*, G. Dobbie, U. Frank, G. Kappel, S. W. Liddle, and H. C. Mayr, Eds. Cham, Switzerland: Springer, 2020, pp. 207–222
- [5] M. M. H. Dihyat and J. Hough, "Can rule-based chatbots outperform neural models without pre-training in small data situations: A preliminary comparison of AIML and Seq2Seq," in *Proc. 25th Workshop Semantics Pragmatics Dialogue*, 2021, pp. 1–3
- [6] J. Jiang and N. Ahuja, "Response quality in human-chatbot collaborative systems," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.* New York, NY, USA: Association for Computing Machinery, Jul. 2020, pp. 1545–1548, doi: 10.1145/3397271.3401234.
- [7] Nawaz, N., & Gomes, A. M. (2019). Artificial intelligence chatbots are new recruiters. *International Journal of Advanced Computer Science and Applications*, 10(9), 1-5.