International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 01 | Jan - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

Development and Implementation of Chatbot Using Machine Learning

Chakshu
Computer Science & Engineering
Presidency University,
Bangalore, India

Choleshwar Kumar
Computer Science & Engineering
Presidency University,
Bangalore, India

Amarjeet Kumar
Computer Science & Engineering
Presidency University,
Bangalore, India

Komma Bhanu Prakash Reddy
Computer Science & Engineering
Presidency University,
Bangalore, India

Samrat Sarkar
Computer Science & Engineering
Presidency University,
Bangalore, India

Dr. Iqbal Gani Dar
Assistant Professor of
Computer Science & Engineering
Presidency University,
Bangalore, India

Abstract- **The Machine Learning based computer program designed to stimulate human like conversation. It provides automated support and services. Using Natural Language Processing and Machine Learning algorithm. It provides 24/7 accessibility and personalization. This chatbot provide customer service, user interaction. If the user requires any information which is not present in database, then its automatically train the data as per the user required. It takes less time to solve the problem, query and any other questions. In this chatbot we collected the user information and train the data. It interacts to customer through text or voice. It can respond to user queries instantly, reducing wait time. It enhanced user experience like personalized support, Multilingual support or improved engagement. Its analyze user behavior and provides user preference and main point. It encrypts user data, ensuring that user information is protected. It's also save the cost of the user such as Lower training cost.**

Keywords: Machine Learning, Natural Language Processing, Virtual Assistant, Data sets, Text-based, Voice based.

I. INTRODUCTION

A human-computer discourse framework, the frame of chatbots that work through normal dialect preparing. It utilizing either content or discourse, is a mechanized framework, which ordinarily runs 24/7, and its essential utilize is driving transformation; it is implied to handle millions of demands per hour.

Chatbots are a group of intelligent, conversational computer programs that are activated by input in a distinctive vernacular. They are able to understand input, execute tasks, and receive commands. Indeed, even though chatbots have been around for a while, they are becoming more advanced due to open-source development platforms, increased preparation control, along with data accessibility. These elements have sparked the widespread adoption of chatbots across various industries and domains. We encounter chatbots in a variety of contexts, including social media groupings or client gain, shipper websites, and profit-keeping money administrations, among others.

For the most part, chatbots are made to achieve specific assignments. Client care chatbots are made as it were for clients who look for benefit, and conversational chatbots are made to banter with clients. Be that as it may,

it is conceivable to prepare on expansive datasets and file human-level interaction, but the organizations ought to thoroughly test and check their chatbot sometime recently discharging it into generation.

II. LITERATURE REVIEW

Here, the factors are based on which feelings are classified using artificial intelligence techniques. Using convolutional neural networks, deep learning, and repeated neural networks (RNNs), the researchers develop models for classifying emotions from a portion of identified data. When using Natural Language Generation (NLG) and Characteristic Dialect Handling (NLP) to guide client discussions, etymological interaction is crucial. Here, voice and content acknowledgment are used in a multi-modal strategy. They have gathered corpuses to help them learn the semantic information of words and speak to them as vectors. They have also gathered information about lexical equivalents of words. [1]

Table 1. Chatbot using Machine Learning [8], [9]

Author	Year	Title	Objective	Methodology	Key Findings	Limitations
Zhang et al.	2020	Development of Intelligent Chatbots Using NLP	To explore the integration of NLP techniques in chatbot systems.	Implementations Of chatbot using deep learning and NLP.	NLP-based chatbots provide context-aware and fluent conversations.	High computational cost for training large datasets.
Gupta & Verma	2019	Chatbot Framework for E-commerce platform.	To evaluate chatbot effectiveness in improving user engagement in e-commerce	Experimental study with user feedback	Chatbots enhance customer satisfaction and reduce response time	Limited evaluation across diverse industries
Lee et al.	2021	Sentiment Analysis in Chatbot Interactions	To analyze how sentiment analysis improves chatbot-	Sentiment analysis model integrated with chatbot	Sentiment-aware chatbots adapt responses to user emotions	Difficulty in understanding mixed or ambiguous sentiments
Kumar et al.	2022	Machine Learning Models for Multilingual Chatbots	To assess ML models for multilingual chatbot development	Comparative study of ML models	Transformers outperform RNNs in accuracy and language versatility	Limited focus on low-resource languages

Smith & Brown	2020	Ethical Challenges in ML-based Chatbots	To examine ethical concerns of chatbots using sensitive user data	Review of case studies and regulatory guidelines	Privacy and bias are significant issues in chatbot deployment	Lack of concrete solutions to mitigate ethical risks

A voice recognition chatbot is developed in this paper. Third-party expert systems are used to help manage the questions that the bot is asked but is unable to understand. Based on client performers and content, the webbots are created as web pals. They focused on the made strides structure here in case the application isn't really text-based but rather voice-based prepared. In this case, voice acknowledgment involves two steps: capturing and processing an input flag. acknowledgment of information from the server's response and data preparation. The server used in this instance is built using Cleanser as a dark box technique. The use of master frameworks enables the advancement of infinite and autonomous ideas. [2]

A conversation between people and machines will be facilitated by this chatbot. In order to identify the sentence and decide whether to respond to the address, the framework will keep the information database here. Bigram will be used to determine the input sentence's similarity score. The chatbot data is stored in an RDBMS. [3]

The chatbot was implemented using design comparison, which recognizes sentence structure and spares reaction design. Here, the developer demonstrates how to use the chatbot. Database, computer program, programming language, working framework, and the processes of input and output are stored. Here, content is used to take the input, trim is used to remove any more accentuation, and arbitrary is used to choose a response from the database. The purpose of the chatbot is entertainment. [4]

The author wrote about Google Collaborator in 2020, which includes the subsequent Google Present period. This was created in 2012 and provided responses according to the locations and preferences of users. Google Right Hand has a more user-friendly client interface and might offer deeper counterfeit insights. Because Google Right Hand is explicitly linked to a user's Google accounts, its main drawbacks are identity theft and security incursion. [5]

Distinctive reports displayed in web-the substance is inspected by labeling dataset using n - gram based moo dimensional demonstration and TF-IDF framework that makes the S, U, V within the final portion, 3 frameworks increase - cosine closeness result. [6]

III. PROPOSED SYSTEM

The chatbots are conversational virtual collaborators that computerize intuitively with the clients. Chatbots are fueled by manufactured insights utilizing Machine Learning strategies to get it common dialect. The thought process of the paper is to assist the client in looking at the inquiry. At first, when the client visits the site to begin with registers themselves and afterward can inquire the bot their queries. The framework employments a master system to reply the inquiries reply isn't display within the database. Here the domain experts moreover ought to give different points of interest. The information of the chatbot put away within the database within the frame of pattern-template. Here SQL is utilized for dealing with the database. The diagram of and created arrangement is clarified step-wise as takes after:

- Planning of Advising Process-Specific Questions to begin with, the precise arrangement of any questions related to a directing handle sets the arrangement swirling. Any arrangement is embraced after cautious think about of the differential needs and desires of users-such as a plans client, and a current client. [7]
- Taking care of Numerous Shapes of the Same Inquiry through Semantic Investigation: The chatbots effectiveness can be improved encourage through semantic investigation by dealing with numerous shapes of the same inquiry. Usually conceivable with the utilize of characteristic dialect handling strategies in which show how the same address is being communicated. The result is reliable and precise reactions to clients independent of the inquiry expressing, making it user-friendly and proficient.
- The created arrangement makes beyond any doubt that the chatbot can handle questions from straightforward to complex. Be it the address almost affirmation due dates or complex questions with respect to scholastic approaches, the chatbot is outlined to get it and react suitably.
- Optimization of execution and client encounter of the chatbot arrangement is backed by the integration of different innovations. NLP, Machine Learning calculations, or indeed profound learning models may be coordinates to permit the chatbot to superior decipher the expectation of a client and way better get it the setting of a communication. The arrangement grasps a comparative examination of diverse advances to choose the foremost appropriate ones.

IV. SYSTEM ARCHITECTURE

Figure 1 shows that Chatbot Development flow chart. In this flow chart User input the query and it will pass through different steps to generate output. From beginnings when the user giving a query as input like Text. After that the text will moved forwarded to Convert text as Lowercase letter. Next step is tokenization process in which all the text should be divided in sub-words and send to next step. In this step the program should check if the sentence or word are meaningful then if generate output for the next step. The next step is stemming it is the process of checking the prefix, suffix or infix.

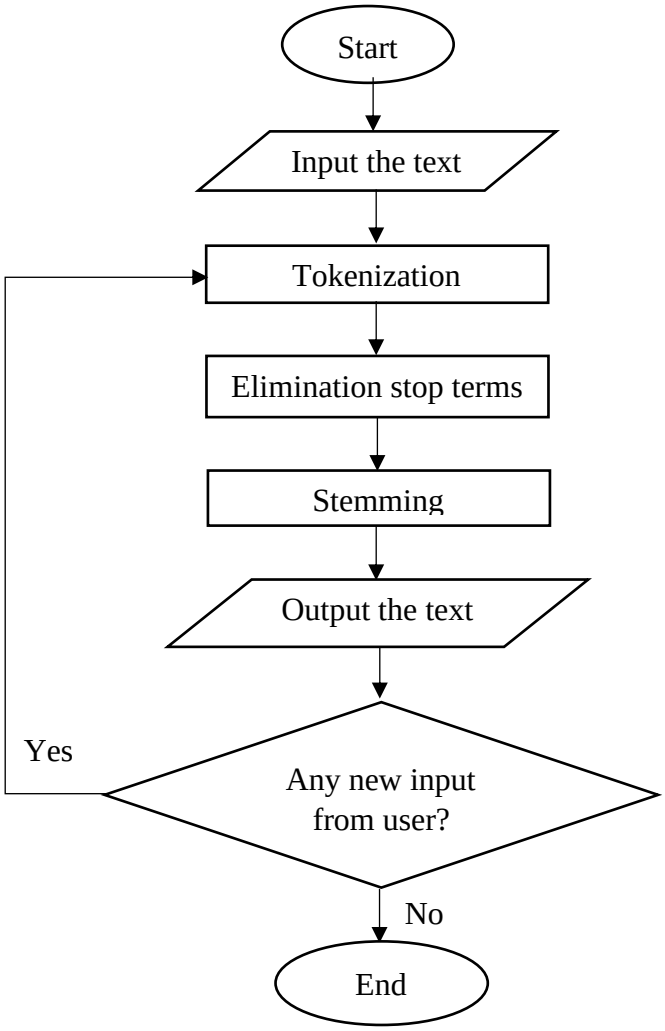

Figure 1. Chatbot Development

This check the grammatical uses in sentence is correct or not. If the grammatical is not in proper form, then it shows error. After all step's passed the program produce output and again ask Any New Input or End the task. It will help users to ask the query and solve the problem. It is very fast to find the solution, if the solution is not in the database it automatically updates the data. It shows the data in easiest way like tabular form, numbering, using pointers.
The steps for Chat development are:

- Lowercase: The text input which is collected from user and convert it to lower case letter. This avoid number value.

- Tokenization: This is the process in which main sentence is substituted to sub sentence. All the words should be counted as Tokens.

- Eliminating Stop terms: The process check if the sentence is meaningful or not. If the sentence is not meaningful then the sentence should be void.

- Stemming: This is the final step to check the sentence grammatical are in proper way. It checks the prefix, suffix, and infix. If it is not in proper form, then it shows error.

Using the formula below, the word frequency is used to determine how frequently a phrase has appeared in a given sentence.

$$tf = tfi \text{----------(1)}$$

The weight of rare terms across all reports in the document is calculated using IDF. The IDF score of the words that appear frequently in the document is high. It is provided by the following condition:

$$idf = logN/df \text{----------(2)}$$

The weight of the term or word in the document is determined by combining the tf and idf. The weight of each term in the document is calculated by multiplying the tf and idf values.

$$Wi = tfi * logN/df \text{----------(3)}$$

Algorithm

input = Ask from chatbot
sentTokens ← Sentence tokenized
tf idf ← T F − IDF (data victories)
vals ← Cosine
query (tf idf)
reqT F IDF
if reqT F IDF > 0
  return
else
  return “Ask again”
end if

V. RESULT ANALYSIS

The chatbot models are tested on a test dataset consisting of 140 queries from a single university. The sklearn metrics library is used to compute confusion matrices and accuracy.

Table 2. Matrix of Chatbots [3]

Bot	True	False
True	9	5
False	4	119
Roger	True	False

True	5	8
False	3	150
Victus	True	False
True	10	1
False	2	140
Radio	True	False

The neural network is not constructed in this model. Rather, each sentence is transformed into a vector using TF-IDF Vectorization, and the similarity between each sentence and the query is determined using Cosine Similarity. This model only selects the most similar sentence; it must comprehend the content of the query. The Victus confusion matrix is shown in Table 2.

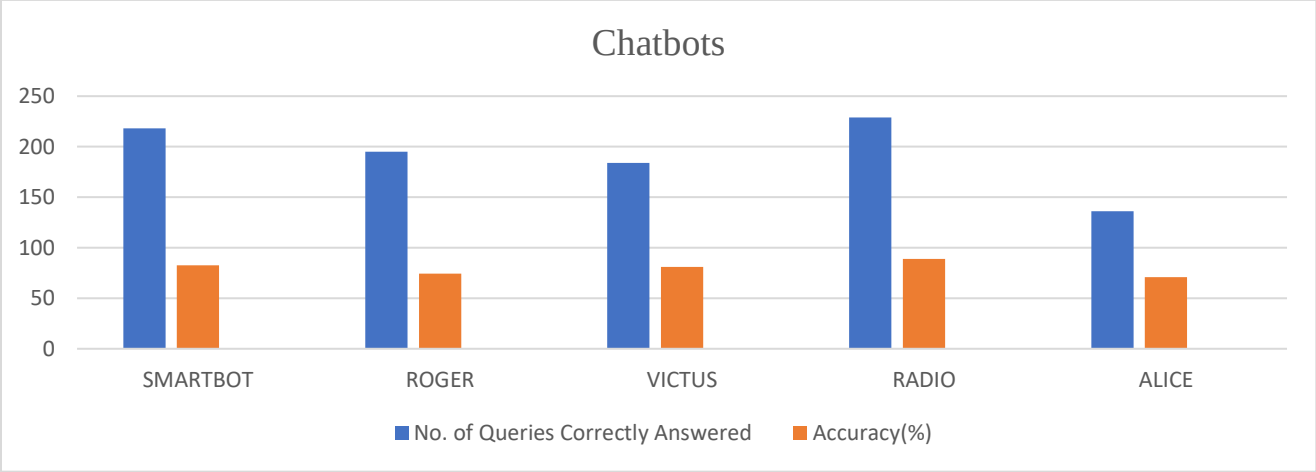

Figure 2. Report of chatbots

The neural network of this model was constructed using sequential modeling, which was intended to avoid the overfitting issue and enhance the model's functionality. The Smartbot confusion matrix is shown in Table 2.

VI. CONCLUSION AND FUTURE SCOPE

The aim of the paper is to create a framework that demonstrates high-quality responses in the shortest amount of time. As a result, the user is relieved of the stress of having an expert respond to their question directly. This program saves time while seeking a solution from a specialist or doctor for medical issues. Here, we created the program to extract the user query's keyword utilizing the TF-IDF and N-gram. To choose the appropriate response for the inquiry, each keyword is weighed down. The Web interface is designed to allow users to enter their queries. Upgrades to the application's efficacy and security make it better. Future chatbots will learn user preferences, habits, and behavior to provide more tailored experiences.

REFERENCES

[1] K. Oh, D. Lee, B. Ko and H. Choi, "A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation," 2017 18th IEEE International Conference on Mobile Data Management (MDM), Daejeon, 2017, pp. 371-375. doi: 10.1109/MDM.2017.64

[2] Du Preez, S.J. & Lall, Manoj & Sinha, S. (2009). An intelligent webbased voice chat bot. 386 - 391.10.1109/EURCON.2009.5167660

[3] Bayu Setiaji, Ferry Wahyu Wibowo, "Chatbot Using a Knowledge in Database: Human-to- Machine Conversation Modeling", Intelligent Systems Modelling and Simulation (ISMS) 2016 7th International Conference on, pp. 72-77, 2016.

[4] Dahiya, Menal. (2017). A Tool of Conversation: Chatbot. INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING. 5. 158-161.2017.

[5] E. Adamopoulou and L. Moussiades, “Chatbots: History, technology, and applications,” Machine Learning with Applications, vol. 2, p. 100006, 2020.

[6] Farheen Naaz, Farheen Siddiqui, modified n-gram based model for identifying and filtering near-duplicate documents detection, International Journal of Advanced Computational Engineering and Networking, ISSN: 2320- 2106, Volume-5, Issue-10, Oct.-2020

[7] Michelle Denise Leonhardt, Liane Tarouco, Rosa Maria Vicari, Elder Rizzon Santos, Michele dos Santos da Silva. \"Using Chatbots for Network Management Training through Problem-based Oriented Education\", Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007), 2007

[8] Zhang, X., Li, Y., & Wang, Q. (2020). Development of intelligent chatbots using NLP. *Journal of Computer Science, 19*(6), 1124-1140.

[9] Gupta, S., & Verma, P. (2019). Chatbot frameworks for e-commerce platforms. *E-commerce Research Journal*, *8*(3), 201-218.