

## RETRIEVAL BASED CHATBOT SYSTEM

Tanay Sharma<sup>\*1</sup>, Tarunpratap Singh Rathore<sup>\*2</sup>, Narendrapal Singh Rathore<sup>\*3</sup>

<sup>\*1,2</sup>Student, Department Of Computer Science And Engineering, Acropolis Institute Of Technology And Research, Indore, MP, India.

<sup>\*3</sup>Associate Professor, Department Of Computer Science And Engineering, Acropolis Institute Of Technology And Research, Indore, MP, India.

### ABSTRACT

The Health Bot is a retrieval based chatbot which is directed specifically towards people who are suffering from the novel coronavirus or whose friends/families are suffering from the same. Programming of this chatbot is done in python along with using concepts of Machine Learning and Deep Learning. The chatbots are broadly classified into two types: Retrieval Based and Generative Based. The Health Bot is a chatbot based on Retrieval Based A retrieval-based chatbot uses predefined input patterns and responses. It then uses some type of heuristic approach to select the appropriate response. It is widely used in the industry to make goal-oriented chatbots where we can customize the tone and flow of the chatbot to drive our customers with the best experience. The goal is to reach out and help as many people as possible who are suffering during this pandemic by providing them with resources and connections in time. The end result is to provide a connecting bridge between people and resources to make utmost utilization of it.

**Keywords:** Machine Learning, Deep Learning, Retrieval Based, Generative Based , Python.

### I. INTRODUCTION

A chatbot is a program that can be operated on computer devices which allows humans to interact with technology using a variety of input methods such as voice, text, gesture and touch. From past decade chatbots are being used in customer service environments but recently people are understanding the vast capabilities of chatbots and how easy can our lives be with the help of chatbot hence, they are now being used in a variety of other roles within enterprises to improve the customer experience and business efficiencies as well as in health and medical science. Chatbots have a variety of different names along with different levels of intelligence. A basic chatbot might give you a feeling of a front-end solution for answering standard FAQs. Chatbots built using some of the frameworks may offer slightly more advanced features like slot booking or simple transactional capability, such as taking food orders etc. All the AI chatbots are the only ones which are literally capable of meeting customers expectation as they can provide controlled user experience. At present moment almost all the enterprises are looking to deploy a personalized conversational AI chatbots for their company. A Chatbot can be of different types. Some of the methods/systems of building a chatbot are as follows:

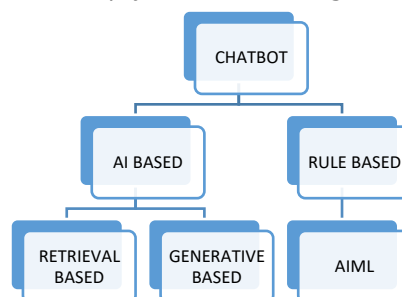


Figure 1: Overview of chatbot

### II. METHODOLOGY

The system we have proposed is Retrieval based chatbot system. This suits perfectly for our idea as the amount of queries in health and medical industry is limited and can be guessed easily. Hence, for such queries we can give exact instantaneous response to the users. The working principle of retrieval based chatbot is based on trees and graphs. The basic methodology of retrieval based chatbot is to provide best possible response to the questions asked by the user. The responses are manually created by the developer which are based on

knowledge of pre-existing data. Bot is trained in such a way that there will be a set of frequently asked questions for which this bot will provide a suitable answer. Retrieval based chatbots are the most common kind of AI based chatbot used nowadays. As the responses are pre-defined hence, it allows bot developers to control user-experience, they can make it more friendly depending upon the satisfaction of a certain group of customer using it. It is best for customer support, lead generation, feedback, order placing etc.

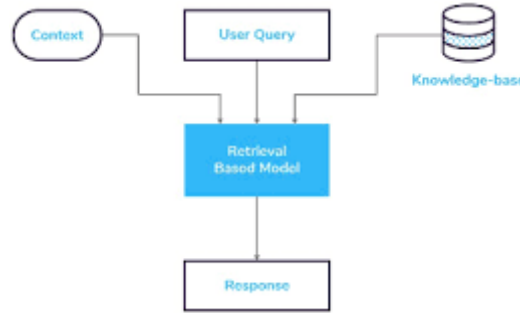


Figure 2: Retrieval Based Chatbot

Figure 2 shows the basic architecture of the Retrieval based chatbot system. Here, a user utters an input in the form of a query, the system then starts searching for a response depending upon the intent of the query. A matching is done between the query fed by the user and the set of queries that are already stored in the repository. The chatbot is trained on the dataset which contains categories(intents), pattern and responses. We use a special recurrent network(LSTM) to classify which category the user’s message belongs to and then we will give a random response from the list of responses.

### III. MODELING AND ANALYSIS

The basic aim of building a Deep Learning model using the input questions to drive the conversation in a more meaningful way is considered while building the system model. Our proposed system works on a manually developed Health related querying dataset for COVID-19 affected people. The Healthbot query produces relevant responses based on the pre-defined repository. We have used a special recurrent network which is Long Short Term Memory to classify which category the user’s message belongs. The LSTM belongs to Keras library. Along with Keras, NLTK, tensorflow were also used.

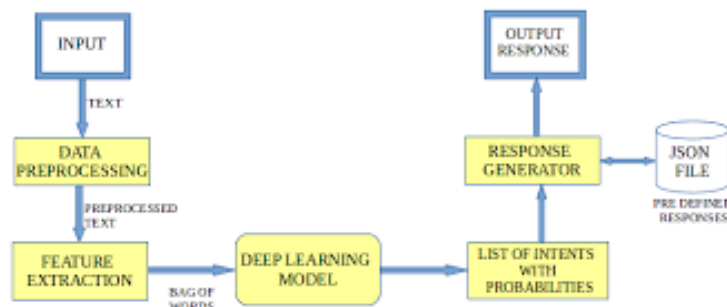


Figure 3: System Overview

Figure 3 shows the basic architecture of system behind the working of the chatbot. First, the user gives an input in the form of a query, then the processing of data takes place on that text which is done with the help of python libraries like panda and numpy. After data processing the text is divided into a list of token of words which is then given to our deep learning model. Through the deep learning model these tokens are being classified into different list of intents. There is a list of intents which is predefined by the developer, the deep learning model assigns that query of user to matching intent. With those intents there’s a repository of responses which is connected to each intent. From those lists of responses the bot sends a response at random. This is how a retrieval based chatbot works.

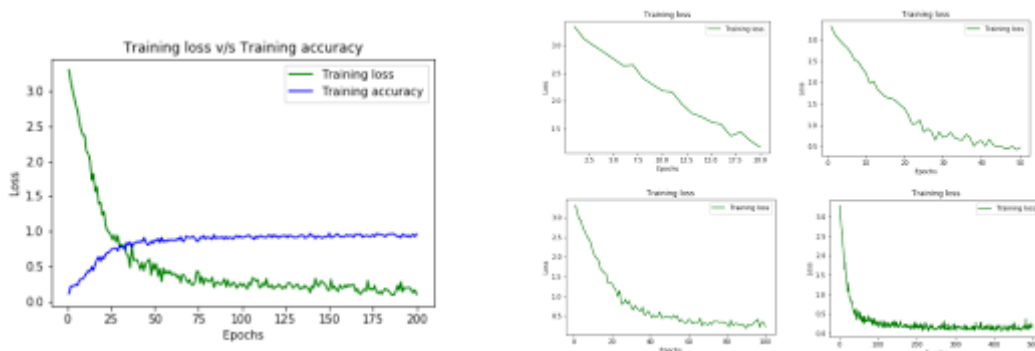
### IV. RESULTS AND DISCUSSION

In the process of building a chatbot there has to be a lot of hit and trials and dealing with errors before you reach the end result. But once you get it, it is all worth it. We tried multiple machine learning and deep learning models before we reached to our end result. Once the most accurate model is selected you can just try it out by

asking queries and getting desired responses. As it is a retrieval based chatbot and it gives pre-defined responses therefore, it makes developer’s work easy while testing. Also, if you want to increase your output you just need to keep updating your repository. Updating your repository frequently is also an easy way of increasing your projects’ efficiency. Our project is working fine and giving all the desired outputs.

**Table 1.** System Accuracy

S.no.	Value of n	Accuracy %
1	10	76
2	15	88
3	20	85



**Figure 4:** Graphical analysis

### V. CONCLUSION

In today’s world chatbots are one of the most useful and easy to access technology. It has various applications in vivid fields such as in health and medical field for providing basic information and for booking appointments, in business industry for customer support and feedback, in restaurants for placing for placing orders and paying bills, etc. Here, we have created a retrieval based chatbot named as The Health Bot which solves all your queries regarding novel coronavirus and also guides you about all the resources available around you. It first takes input from the user then after applying deep learning model gives you a suitable response from the predefined repository. In order, to keep this chatbot efficient you need to constantly keep updating the repository and add more data in it. This model is made using Keras and deep learning tools, hence we can make it more efficient by using more such libraries and making it more user friendly. We can also convert into a hybrid chatbot by integrating retrieval based chatbot with generative based chatbot working and properties with the help of a large amount of data.

### ACKNOWLEDGEMENTS

I would like to give credits to our professors Narendrapal sir and Kavita Namdev ma’am for continuously providing us help and motivation for completing this research paper.

### VI. REFERENCES

- [1] J. Singh, M. H. Joesph, and K. B. A. Jabbar, “Rule-based chabot for student enquiries,” in Journal of Physics: Conference Series, vol. 1228, p. 012060, IOPPublishing, 2019
- [2] O. Vinyals and Q. Le, “A neural conversational model”, arxiv preprintarXiv:1506.05869, 2015
- [3] Y. Wu, Z. Li, W. Wu, and M. Zhou, “Response selection with topic clues for retrieval-based chatbots”,Neurocomputing, vol. 316, pp. 251–261, 2018
- [4] B. Hancock, A. Bordes, P.-E. Mazare, and J. Weston, “Learning from dialogue after deployment: Feed yourself, chatbot!”, arxiv preprint arxiv:1901.05415,2019
- [5] P. Lison, “Probabilistic dialogue models with prior domain knowledge,” in Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pp. 179–188, Association for Computational Linguistics, 2012
- [6] O. Dusek and F. Jurccek, “A context-aware natural language generator for dialogue systems,”arXiv

- preprint arXiv:1608.07076, 2016.
- [7] J. D. Williams and G. Zweig, "End-to-end lstm-based dialog control optimized with supervised and reinforcement learning," arxiv preprint arXiv:1606.01269, 2016
- [8] R. Yan, "'chitty-chitty-chat bot": Deep learning for conversational ai.," in IJCAI, vol. 18, pp. 5520–5526, 2018.
- [9] A. Vazquez, D. Pinto, and D. Vilari no, "A computational model for automatic generation of domain-specific dialogues using machine learning," in Proceedings of the XVIII International Conference on Human Computer Interaction, pp. 1–2, 2017
- [10] B. Wu, B. Wang, and H. Xue, "Ranking responses oriented to conversational relevance in chat-bots," in Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 652–662, 2016
- [11] S. Konstantopoulos, "An embodied dialogue system with personality and emotions," in Proceedings of the 2010 Workshop on Companionable Dialogue Systems, pp. 31–36, Association for Computational Linguistics, 2010.
- [12] E. Varghese and M. R. Pillai, "A standalone generative conversational interface using deep learning," in 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 1915–1920, IEEE, 2018
- [13] B. Salehi, P. Cook, and T. Baldwin, "A word embedding approach to predicting the compositionality of multiword expressions," in Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 977–983, 2015
- [14] N. Shakhovska, O. Basystiuk, and K. Shakhovska, "Development of the speech-to-text chatbot interface based on google api," in CEUR Workshop Proceedings, vol. 2386, pp. 212–221, 2019
- [15] Surendran, Aiswarya. "Conversational AI - A Retrieval Based Chatbot." easychair.org. [https://easychair.org/publications/preprint\\_download/LnTd](https://easychair.org/publications/preprint_download/LnTd) (accessed June 28, 2021).