A Chatbot in Malayalam using Hybrid Approach

PRAVEEN PRASANNAN[1]   **STEPHY JOSEPH[2]**   **RAJEEV R R[3]**

Dept. of Language Technology, International Centre for Free and Open Source Software (ICFOSS)

praveenprasannan@icfoss.org   **stephy@icfoss.org**   **rajeev@icfoss.in**

Abstract

Chatbot is defined as one of the most advanced and promising expressions of interaction between humans and machines. They are sometimes called as digital assistants that can analyze human capabilities. There are so many chatbots already developed in English with supporting libraries and packages. But to customize these engines in other languages is a tedious process. Also there are many barriers to train these engines with other morphologically rich languages. A hybrid chatbot employs the concepts of both Artificial Intelligence (AI) and rule based bots, it can handle situations with both the approaches.

1. Introduction

Chatbot is a computer program that converses with humans like a human partner in natural language. Nowadays, chatbots are being used widely in various domains like Tourism, Health, Business, Customer support. The basic architecture of a chatbot is shown in Figure.1. When a user asks a query to a chatbot, the chatbot in return interacts in human natural language with the users. In general chatbots analyze and identify the user's request intent and extract the relevant entities. After analyzing, an appropriate response will be delivered to the user.
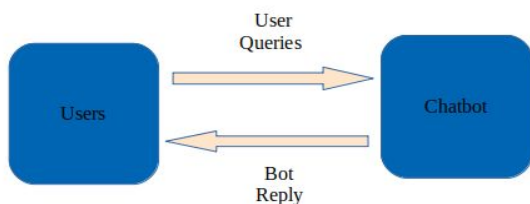

Figure.1

In this work we developed a chatbot in Malayalam in the Pandemic Situation Coronavirus. One of the biggest threats faced by the society during the Corona pandemic was Mis-Information, Dis-information and Mal-information. Government wanted to establish a single source of truth, where the public can rely for authentic information. To support the cause and to fulfill the need to support the general public due to the rapid spread of COVID-19 Pandemic during the months of February and March 2020, we developed an interactive bot which is based on 'hybrid technology' and interacts with the people in regional language (Malayalam).

The main objective of this project was to make the people aware of the disease COVID-19 and the protective measures and precautions to be taken during this period. This chatbot provided all the information related to the pandemic including Government Orders, FAQs and details regarding the COVID-19, authenticated by the State Kerala of Government which is available 24x7 and people do not need to make telephonic contact with the authorities to get the information regarding the virus in this pandemic situation. This chatbot was rolled out through the official website of the State Government of Kerala and was one of the major tools to fight against the situation. The project helped the State Administration as well as the general public in controlling the anxiety among the people. The solution also has the facility of a semi-manual intervention for handling queries with the help of the State Government's Health Department, which the Chatbot is not able to respond to.

2. Methodology

On chatting with the bot, it says, 'കൊറോണ വൈറസുമായി ബന്ധപ്പെട്ടുള്ള വിവരങ്ങൾ അറിയാനായി ചോദിക്കുക', (Ask for information related to the coronavirus). When the user asks about his queries or FAQs,

the bot assists in autocompleting of his queries or FAQ's. If the user finds any of the autocomplete suitable for his query, he can select that.
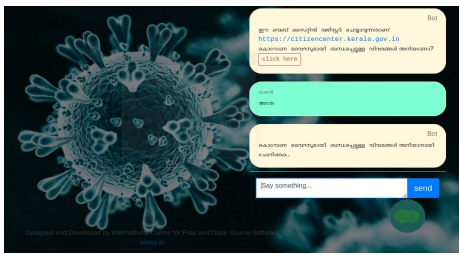

Figure.2

In some cases the user may frame his own questions. The bot should be able to reply in such situations also. But the reply can be relevant or not. In order to make the bot generate a relevant answer, Hybrid approach is used. Hybrid chatbot employs the concepts of both AI and rule based bots and can handle these situations. In this approach, the user query undergoes some preprocessing stages like tokenization, stop words removal and root extraction of keywords. If a match is found, then all the queries regarding that keyword will be the bot's reply and if a match is not found, the system will use similarity measurement to find the most relevant keyword, and respond back.
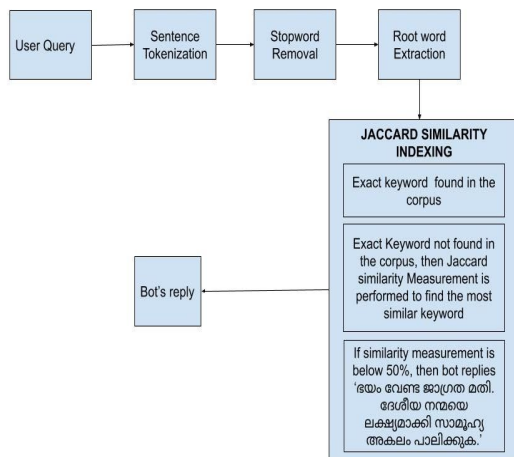

Figure.3

3. Result

When a user asks about the various queries regarding the coronavirus, bot generally categorizes it into three categories (a) exact query available in the database, (b) if exact query not found, then search for the keywords in that query, (c) if exact keyword is not found then use Jaccard indexing. For the above three situations, the chatbot managed to provide relevant responses. For unknown queries a team from Directorate of Health service associated with this work to answer these questions.

4. Conclusion

Already there are so many works going on for the chatbot in English. So there are a lot of libraries and platforms available for building the chatbots in English. Malayalam language is an Indian language that has its own characteristic features, which makes it entirely different from other languages. So there are no readily available modules or platforms for building a chatbot in Malayalam. So developing a chatbot in such a language is a tedious process. Due to the lack of dataset in Malayalam, the accuracy of the chatbot was low in the initial stage. But with the increase in dataset, the accuracy also increased. Thus with the increase in the dataset, the accuracy of the chatbot can be increased.

References

Weizenbaum, Joseph. 1966 "ELIZA—a computer program for the study of natural language communication between man and machine." *Communications of the ACM9.1,* 36-45,, .

Bayan AbuShawar, Eric Atwell. 2015 "ALICE Chatbot: Trials and Outputs" *Computación y Sistemas, Vol. 19, No. 4,,* pp. 625–632 ISSN 2007-9737

Kalaiyarasi, T., Ranjani Parthasarathi, and T. V. Geetha. 2003 "Poongkuzhali-an intelligent tamil chatterbot." *Sixth Tamil Internet 2003 Conference. Vol. 1. sn,.*

S. Chaitrali, Kulkarni, U. Amruta, Bhavsar, Savita Chaitrali S Pingale. 2017. "Bank Chatbot –an Intelligent Assistant System Using Nlp And Machine Learning", *International Research Journal Of Engineering And Technology (Irjet), Volume: 04 Issue: 05, 2017.*

Shah, Rishabh, Siddhant Lahoti, and K. Lavanya. 2017, "Anintelligent chat-bot using natural language processing." *International Journal of Engineering Research6.5 : 281-286. 2017.*