International Journal of Future Generation Communication and Networking
Vol. 13, No. 1s, (2020), pp. 138- 146

MULTILINGUAL CHATBOT WITH HUMAN CONVERSATIONAL ABILITY

[1] Aradhana Bisht, [2] Gopan Doshi, [3] Bhavna Arora, [4] Suvarna Pansambal
[1][2] Student, Dept. of Computer Engineering, [3][4] Asst. Prof., Dept. of Computer Engineering, Atharva College of Engineering, Mumbai, India

Abstract

Chatbots - The chatbot technology has become very fascinating to people around the globe because of its ability to communicate with humans. They respond to the user query and are sometimes capable of executing sundry tasks. Its implementation is easier because of wide availability of development platforms and language libraries. Most of the chatbots support English language only and very few have the skill to communicate in multiple languages. In this paper we are proposing an idea to build a chatbot that can communicate in as many languages as google translator supports and also the chatbot will be capable of doing humanly conversation. This can be done by using various technologies such as Natural Language Processing (NLP) techniques, Sequence To Sequence Modeling with encoder decoder architecture[12]. We aim to build a chatbot which will be like virtual assistant and will have the ability to have conversations more like human to human rather than human to bot and will also be able to communicate in multiple languages.

Keywords: *Chatbot, Multilingual, Communication, Human Conversational, Virtual agent, NLP, GNMT.*

1. Introduction

A chatbot is a virtual agent for conversation, which is capable of answering user queries in the form of text or speech. In other words, a chatbot is a software application/program that can chat with a user on any topic[5]. Chatbots can be classified into two different categories: i) Task Oriented chatbots - These are designed to perform specific tasks. For example: A customer care chatbot can answer queries related to customer's problem regarding specific service/product.

ii) General Purpose chatbots – These can also be called as virtual assistants that can have open ended discussions with the users. For example: Amazon's Alexa.

These applications are designed to simulate human-to-human interactions. Chatbots are extensively used in businesses and organizations including government, non-profit and private sectors. Their functioning can range from customer service, product suggestion, product inquiry to personal assistant. Many of the chatbots are built using various methods such as retrieval techniques, rule based techniques or simple machine learning algorithms. Chatbots built using retrieval based techniques scan for keywords within the inputs and retrieves pertinent answers based on the query string.

Chatbots are being used very frequently by people as a means for communication, and the study predicts that most of all businesses will use this technology by 2020[13]. About 69% of consumers prefer these systems for quick communication, and 40% do not care if an interaction is with a human or chatbot as long as they receive the assistance they require. Since talking and texting with mobile apps is so natural, many consider these conversational agents as a preferable interface to applications[3].

Chatbots system not only allow a more natural interface with applications to retrieve information but also allow users to communicate for therapy, entertainment, or even foreign language practice[8]. Nowadays, developers are integrating multilingual translation technology[2] with these systems to enable users to have communication in their native languages.

But only being able to communicate in multiple languages is not sufficient for an effective system, the chatbot should also have the ability to understand and respond to human behaviors.

2. Existing System Overview

There are some chatbots that are freely available for the public to use on the web. For example: ELIZA, Rose and Mitsuku that are very popular and have performed well in the Loebner Prize competition[3][10]. We have examined the working of Rose chatbot. Rose chatbot represents a female, which is a computer hacker living in

ISSN: 2233-7857 IJFGCN

San Francisco, representing her as a cartoon image on the page [3].
We also took another example for chatbot which is Google Assistant. Google Assistant is also known as Google's voice assistant. Google Assistant was developed to be a personal assistant while expanding on Google's existing "OK Google" voice controls which help users to get answers to their queries in text as well as in speech, it can perform specific tasks or also it can guide users to reach various sources

2.1) Rose chatbot:

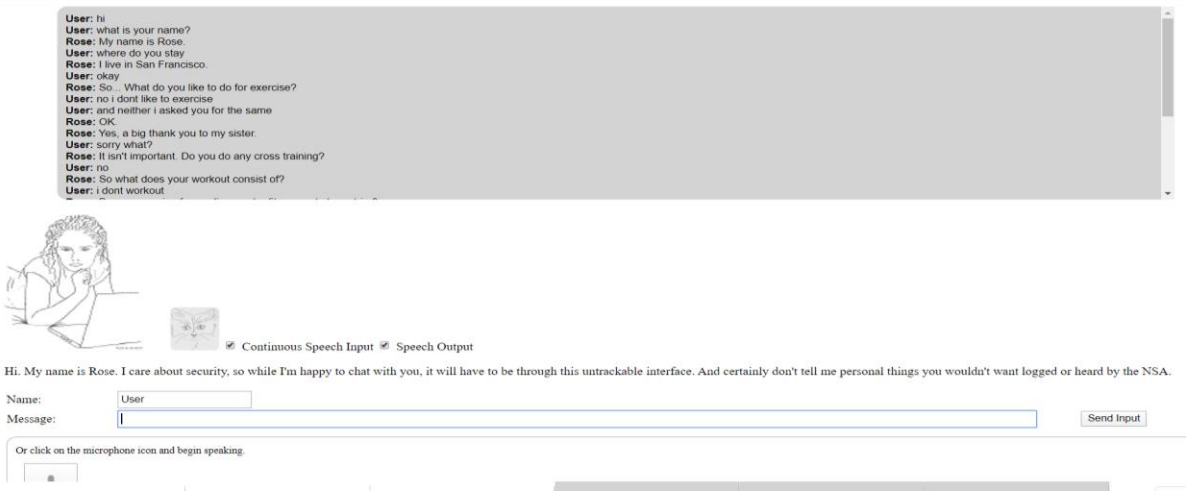


Figure 1. Conversation with Rose chatbot

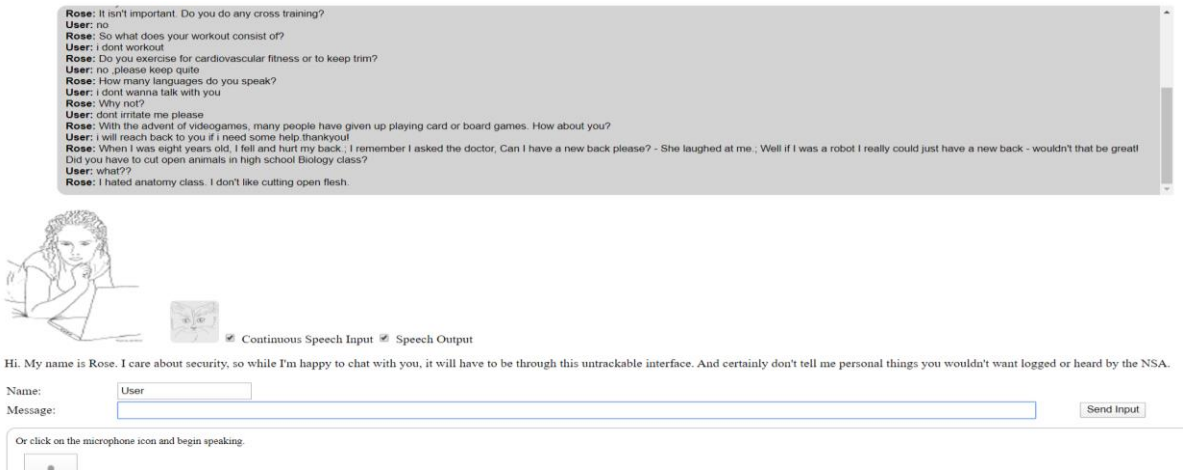
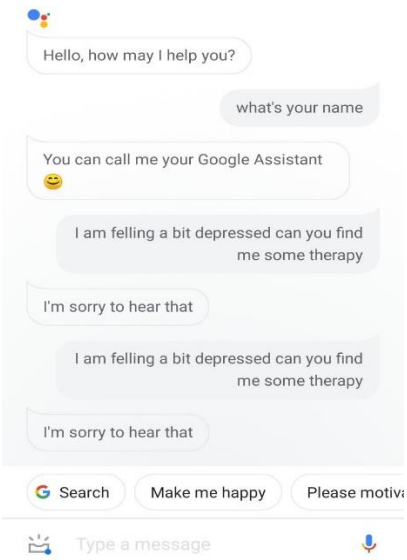


Figure 2. Conversation with Rose chatbot

Figure-1 and Figure-2 shows a sample conversation with Rose chatbot but the system should also have the ability to i) Understand and respond to human behaviors, like the chatbot should also consider what humans say and respond accordingly. But in case of Rose chatbot we didn't get to see that, this is where Rose chatbot lacks behind. ii) And this kind of chatbot sometimes performs well but they are limited to English, and only a few chatbots like Mondly which is able to converse in multiple languages (i.e. 30 languages), Watson in 21 languages. So, providing a chatbot with multilingual capabilities enables people to use chatbots in their native language[7], although people can speak multiple languages but they mostly prefer communicating in their native language.

ISSN: 2233-7857 IJFGCN

2.2) Google Assistant:



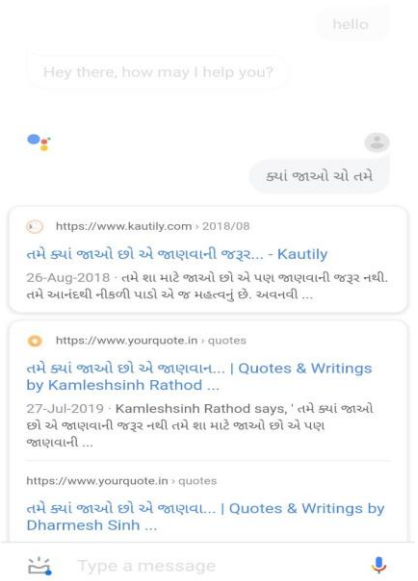


Figure 3. Conversation with Google Assistant Figure 4. Conversation with Google Assistant

Figure-3 and Figure-4 shows a sample conversation with Google Assistant, here the system has the ability to detect multiple languages as we have written a text in a Gujarati language which means -Where are you going, so it gave me a search result stating that text. But although it has the ability to detect multiple languages it cannot respond in that language unless and until you change the language in the settings and also it can allow the user to text in only three languages, one is your Android language and other two are Assistant languages.
Also it doesn't have the ability to converse according to human behavior, as one can refer figure (3) where instead of suggesting a therapy or simply redirecting me to a web search, it kept on saying the same thing.

So here, our idea is to propose a system which overcome the flaws of the above mentioned virtual agents by combining the two features which are being able to communicate according to human behaviors and listen to humans and then respond accordingly(unlike Rose chatbot) and most importantly it will have an ability to detect multiple languages and communicate in that particular language as given in the input.
Here is a glimpse of how our proposed system will respond as compared to other virtual agents, to the following inputs which are in English, Gujarati and Korean Languages.

INPUT TEXT	**ROSE RESPONSE**	**GOOGLE ASSISTANT**	**IDEAL CHATBOT RESPONSE**
Hi ,what is your name ?	Hi. My name is Rose	My name is Google Assistant	Hello, my name is xyz and what is your name?
Where do you stay?	I live in San Francisco	I am in Google's cloud	I am from USA, and you?

ISSN: 2233-7857 IJFGCN

I am feeling a bit depressed	Take two aspirin and pay me $2	I am sorry to hear that	Sorry to hear that, take care. If something is bothering you then you can talk to me about it.
તમે ક્યાં જાવ છો (where are you going)	I think you are just saying junk.	(gives search results on the web)	તમને તે માહિતીની કેમ જરૂર છે (why do you need that information)
너랑 얘기하고 싶지 않아 (I don't want to talk to you)	What is this junk	(gives search results on the web)	죄송합니다. 내가 뭐 잘못했어요? (Sorry. What did i do wrong?)

Table of comparison

3. Proposed System

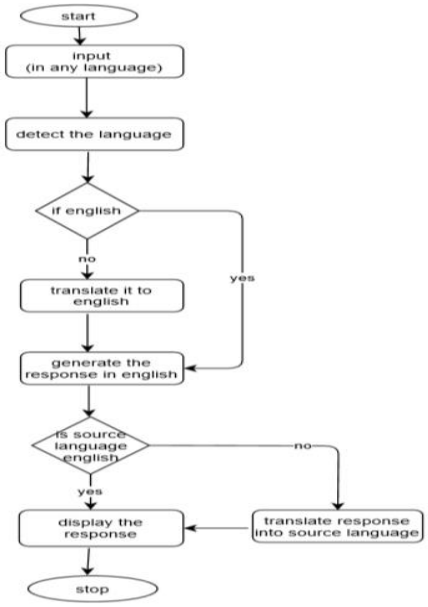


Figure 5. Flowchart of the proposed system

Implementation of a proposed system will be done by considering the step wise execution of the above flowchart.

i) Initially the input is detected, which can be in any language. We identify the language by using Language Identification Algorithm which comes under supervised learning algorithms in Machine Learning and also NLP can do the same.

ii) Once the Language is detected it checks for the particular language library in the Google Translate API (application programming interface). If the language detected is English then the text is forwarded for next operation and if the Language detected is some language other than English then it is translated into English language with the help of Google Translate[13] which is a free multilingual and machine translation service developed by Google, which is used for translating text from one language into another[1].

Text in Source Language > *translated into English language*
(if not in English) *(with the help of Google Translate)*

iii) After the Translation, apply Natural Language Processing (NLP) techniques or Sequence To Sequence

ISSN: 2233-7857 IJFGCN

Modeling with encoder-decoder architecture, to generate the response for the input text (in English Language).
iv) Now, if the source language was English then directly output the generated response whereas, if the input text was in some language other than English then Translate the response generated in English language to the input/source language and then output the response.

Generated response > Translated into source language
(in English language) (if source language is not English)

Here is the elucidation of the technologies using which our idea can be implemented. Firstly we describe the technique of how a chatbot basically generates responses[9]. Here the role of machine learning comes into picture where we initially provide a dataset containing numerous combinations of human-to-human conversations (only in English language). This will help the system in learning, understanding and responding to various inputs in multiple ways which will be more like a humanly conversation. Models can be trained using either one of the following technologies. Both of these techniques are extensively used for developing a chatbot.

3.1) Natural Language Processing (NLP)

Natural Language Processing is a reference to the artificial intelligence methods of communicating with an intelligent system by using the natural language[11]. Implementation of chatbot is one of the most crucial applications of NLP here we are also taking into consideration the speech recognition, the voice assistants like Siri, Google assistant and Cortana[6]. Now, the other use case of natural language processing is the machine translation and the most common example for it is Google Translate. It uses NLP to translate data from one language to another and this action is done in real time.

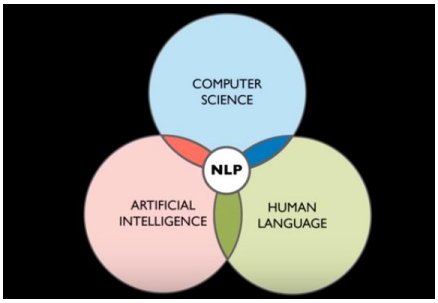


Figure 6. Natural Language Processing(NLP)

There are two major categories in which the NLP is divided : 1)Natural Language Understanding. 2)Natural Language Generation. This refers to mapping the given input into natural language which can be converted into useful representation and analyzing those characteristics of the language whereas Generation is the process in which meaningful phrases and sentences are produced in the form of natural language by utilizing some representation.

Now, the natural language understanding is usually more complex as compared to the natural language generation. The reason behind this is, it requires a massive amount of data and considerable time to usually understand a particular language specifically when you're not a human being. There are series of steps which needs to be followed in order to implement the Natural Language Processing that are :

Tokenization - It is the process in which strings are divided into tokens which in turn are small structures/units that are utilized for tokenization.

Stemming - It is a process which ordinarily refers to generalizing the words into its base or the root form. The working of stemming algorithm involves cutting off the end or the beginning of the word and estimating a list of common prefixes/suffixes that can be found in an inflected word. This algorithm can be successful only sometimes but not always. For instance - Affectation, Affecting, Affection, Affected can be cut down to the root word Affect.

Lemmatization - This process takes into consideration the lexical analysis of the word, doing so requires a detailed dictionary for the algorithm to refer and link the form back to its original word or the root word which

ISSN: 2233-7857 IJFGCN

is also known as lemma. The principal task of lemmatization is to
grouping together different inflected forms of a word, called Lemma. Stemming and lemmatization are moderately identical but the property which differs them is that, the output of lemmatization is a proper word. For instance - a lemmatizer should map coming, came, comes into come.
POS tags - POS tags is an acronym for Parts Of Speech tags. The POS tags are a reference to the morphological type of a word. It represents the functions of a word which consists of its meaning and grammatical presence within the sentence. It is probable for a word to have more than one part of speech depending on what context it is used. For instance - "Google something on the Internet" this sentence uses Google as a verb although it is a proper noun. These are few restrictions that may occur while processing natural language.
Named Entity Recognition(NER) – In this phase it detects the named entities of the person or the company, quantities or the location. This consists of three steps which includes - the noun phrase identification, the phrase classification and entity disambiguation.
Chunking - After we've executed all the above steps, it is time for us to group it together and make sense out of it and for that we have Chunking. It is the process in which we pick discrete pieces of information and group them together into bigger pieces and these bigger pieces are what we typically call chunks. This process helps to extract meaningful information and insights from the given text.

3.2) Google Neural Machine Translation (GNMT)
Google's Neural Machine Translation (GNMT)[4] is essentialy used for machine translation. But, GNMT contains Sequence to Sequence module with many augmenting approaches which can help in building a good dialogue generator. For translation, GNMT does not apply traditional phrase-based translation systems in which the way of performing translation is by breaking up source sentences into multiple chunks and then translating phrase-by-phrase. It rather uses a translation approach which is more like human.

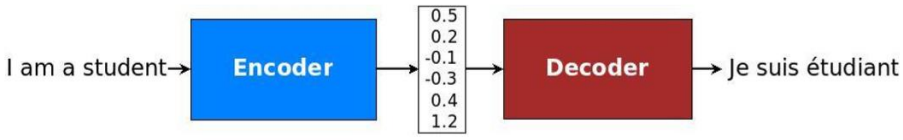


Figure 7. Example of encoder-decoder architecture

Figure-7: Encoder-Decoder architecture - example of a general approach for NMT. An encoder converts a source sentence into a "meaning" vector which is passed through a decoder to produce a translation.
GNMT is based on Seq2Seq architecture composed of encoder-decoder units. GNMT can be used with multiple variants of architecture. The Seq2Seq module is composed of an encoder and a decoder. The source text is taken as input by the encoder and the text is processed to generate intermediate representation of input text called thought vector. The decode input unit is then filled with thought vector. The thought vector is now processed by the decoder and the output is generated. In case of dialogue generation problem the output is a response and for machine translation problem the output is a target text. This architecture is shown to be capable of processing long-range dependencies and producing better fluency of translation or responses.
Encoder-Decoder units of Seq2Seq model can have diverse architectures. Both encoder and decoder are composed of Recurrent Nueral Network(RNN). But the encoder and decoder unit can be composed of different cell types other than vanilla RNN cell, including Long Short-Term Memory(LSTM), or a Gated Recurrent Unit(GRU). Also, unidirectional amd bidirectional RNN can be used for composing encoder and decoder. But it has been empirically found that, LSTM works well for dialogue generation and other language problems as full sequence text can become information bottleneck and in many cases complete sequence is not required for efficient dialogue generation. Also, bidirectional RNN can help make the performance better.

ISSN: 2233-7857 IJFGCN

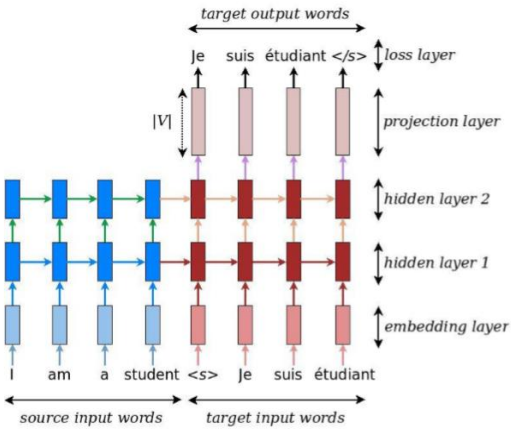


Figure 8. Example of Neural machine translation

Figure-8: Neural Machine Translation - example of a deep recurrent architecture proposed by translating a source sentence "I am a student" into target sentence "Je suis étudiant". Here, "< s >" marks the start of the decoding process whereas, "< /s >" tells the decoder to stop.

After the language is detected, the translation part comes into picture which is done using Google Translator. In this, the input/source language is translated into English language. Google translator supports translation for a total of 109 languages. For instance:

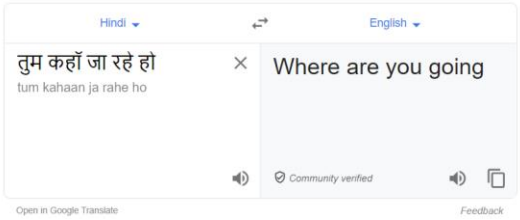


Figure 9. Translation(Hindi to English)

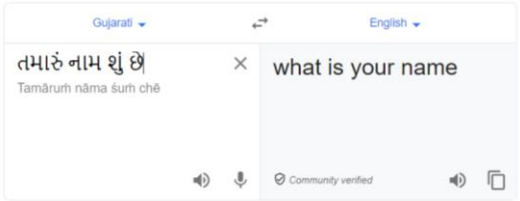


Figure 10. Translation(Gujarati to English)

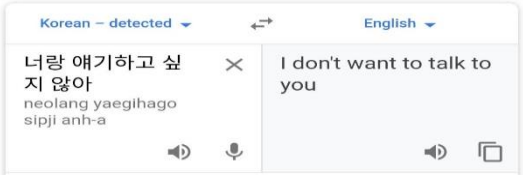


Figure 11. Translation(Korean to English)

Figure-9 shows the translation from Hindi to English. Figure-10 shows the translation from Gujarati to English.

ISSN: 2233-7857 IJFGCN

Figure-11 shows the translation from Korean to English language.
In the same way it can translate any source language into English(if the source is other than English). After the response is generated by NLP/GNMT in English, it can again be translated to the source language using Google Translate.

Figure 12. Translation(English to Hindi)

Figure 13. Translation(English to Gujarati)

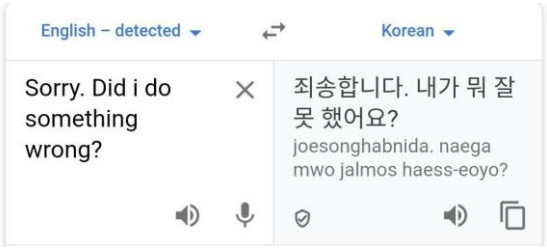


Figure 14. Translation(English to Korean)

4. Future Scope

The virtual agent proposed in our idea can be implemented with the use of robust, high quality and real-life conversational datasets processed using technologies like Natural Language Processing/Google's Neural Machine Translation which would have a better ability to emulate human interaction. Also the system can be capable of communicating in multiple languages(as many languages as supported by the Google Translator).
This system can be developed using various languages such as Python, Java, Ruby, Clojure and PHP. The system built using the proposed idea will be useful to people around the globe as an entertainer, a listener and for executing basic tasks.

5. Conclusion

In this paper we have discussed the limitations of the existing chatbots/virtual agents and proposed an idea through which we can quash the flaws and implement a chatbot that will be Multilingual and will have the ability to communicate in a more humanly manner. We have also compared the response generated by the existing system and response that should be generated by an ideal system based on the same input which is represented

ISSN: 2233-7857 IJFGCN

in the form of a table. We have also explained the details of how an ideal chatbot system can be implemented with the help of a chatbot. The aim of this paper is to point out the inadequacy of the existing system and express how to overcome the existing limitations and build a system which is more developed and accurate.

References

[1]K. Imamura, (2020), Language Translation, 10.1007/978-981-15-0595-9_4.
[2] M. Vanjani, M. Aiken, and M. Park, "Chatbots for Multilingual Conversations", Journal of Management Science and Business Intelligence, 2019, 4-1 July. 2019, pages 19-24 doi:10.5281/zenodo.3264011.
[3] M. Vanjani, J. Posey and M. Aiken, "AN ELAVALUATION OF A MULTILINGUAL CHATBOT", Issues in Information Systems, Vol. 20, Issue 1,(2019), pp. 134-143.
[4] A. Tiha, "Intelligent Chatbot using Deep Learning", (2018), 10.13140/RG.2.2.14006.75841.
[5] J. Weizenbaum, (1966) ELIZA:A computer program for the study of natural language communication between man and machine. Communications of the ACM, volume 9, Issue 1, 36-45.
[6] A. Tulshan & S. Dhage. (2019). Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa: 4th International Symposium SIRS 2018, Bangalore, India, September 19-22, 2018, Revised Selected Papers. 10.1007/978-981-13-5758-9_17.
[7] L. Fryer, "Bots as Language Learning Tools", Language, Learning and Technology. (2006),10.
[8] B. Shawar and E. Atwell, Chatbots: Are they Really Usefull?. (2007), LDV Forum. 22. 29-49.
[9]M. Pereira, and L. Coheur, (2013), Just.Chat - a platform for processing information to be used in chatbots. Instituto Superior T'ecnico, Taguspark Lisboa, Portugal.
[10]D. Mladenic and L. Bradesko, (2012), A survey of chatbot system through a Loebner prize competition. Artificial Intelligence Laboratory, Jozef Stefan Institute, Ljubljana Slovenia.
[11]A. Gelbukh, "Natural language processing", *Fifth International Conference on Hybrid Intelligent Systems (HIS'05),* Rio de Janerio, Brazil, 2005, pp. 1 pp.-. doi: 10.1109/ICHIS.2005.79.
[12]A. Ali and M. Muhammad. (2019), Conversational AI Chatbot Based on Encoder-Decoder Architectures with Attention Mechanism. 10.13140/RG.2.2.12710.27204.
[13]H. Ghasemi and M. Hashemian, (2016), A Comparative Study of Google Translate Translations: An Error Analysis of English-to-Persian and Persian-to-English Translations. English Language Teaching. 9. 13. 10.5539/elt.v9n3p13.
[14] R. Elezaj, (2018). Translation chatbots and the US election. *Multilingual Insights,* November 6, https://multilingual.com/translation-chatbots-us-election/

Acknowledgements

We express our gratefulness to honorable Principal Mr. S. P. Kallurkar Sir for encouragement and facilities provided to us and we would like to express our deep sense of gratitude to Prof. Suvarna Pansambal, Head of Department of Computer Engineering, Atharva College of Engineering, Mumbai, For her generous support and suggestions.

ISSN: 2233-7857 IJFGCN