International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT2390148

Chatbot Using Deep Learning

Shashank Joshi, Sandeep Shinde, Prerna Shinde, Neha Sagar, Sairam Rathod
Computer Science, Vishwakarma Institute of Technology, Pune, India

ARTICLEINFO

Article History:
Accepted: 01 Feb 2023
Published: 23 Feb 2023

Publication Issue
Volume 10, Issue 1
January-February-2023
Page Number
228-232

ABSTRACT

In this paper we have proposed the working of Assistant conversational agent(Chatbot) using deep learning concepts with the utilization of tensorflow library. The LSTM is used overhere, so that the input taken with more than 30-40 words in a sentence canbe replied or answerd with more adequate conversation. The movie dataset used to train the model is taken from Cornell. The model isdesigned to perform a movie dialogue prediction conversation between the user and chatbot. The main aim is to increase the accuracy and estimation of the model. In the proposed model, we have developed a Seq2Seq AI Chatbot with attention mechanism using LSTM and libraries like tenserflow.

Keywords : LSTM, tenserflow, Seq2Seq model, Attention mechanism

I. INTRODUCTION

Chatbots function as a medium for communication between humans and machines. ChatBot is a software system which will enable people to chat using AI. This software system is employed to perform tasks like immediately responding to its users, informing them, serving them to purchase products, and providing higher service to customers. During this paper, we have presented the working and the concepts of a deep learning-based chatbot using sequence to sequence model with attention using LSTM as well as their applications in various sectors such as telecommunication, banking, health, client call centers, and e-commerce. Experimenting with different algorithms we found out that this algorithm is best in comparison with their accuracies and correctness. We have used a movie dataset for training the model, this model can perform well with any type of dataset.

II. LITERATURE REVIEW

In a paper, G Krishna Vamsi, Akhtar Rasool, and Gaurav Hajela have planned an informal agent [(chatbot)computer software capable of communication with humans] using NLP. They mentioned that an important part of building any chatbot was the event of conversation. Despite several developments in NLP and AI, making an honest chatbot model remains a big challenge in this field even nowadays. An informal bot will be used for infinite errands. In general, they have to know the user's intent and deliver applicable replies. They proposed a program of an informal interface that enables a user to converse with the same manner one would address a person. Hence, these are used in

almost every client conversation platform, like social networks[1].

In another paper by Oriol Vinyals and Quoc
V. Le plenty of approaches have been proposed to include context or communication history into seq2seq models so as to create higher dialog agents. Maybe the most simple approach is to concatenate k previous occurrences by appending an end-of-occurrence symbol once every occurrence and feeding this long sequence of symbols into the encoder[3]. In the work by Sordoni et al. and Shang et al., the simplest approach that was used as a baseline is to use solely the first preceding occurrence because the context so as to create context-message-reply triples for training. A stronger approach was to concatenate the bag of word representations of the context and message occurrences rather than the particular occurrences. By using totally different representations for various occurrences higher results were achieved[5].

III. PRE-PROCESSING

Creating a dictionary that maps each line and its id. Creating a list of all of the conversations. Doing a first cleaning of the texts like substituting "i'm" with "i am" using regex for better accuracy. Adding tokens like GO,EOS,UNK,PAD.

1) GO - the first token and though vector is fed to the decoder in order to start generating tokens of the answer.
2) EOS - "end of sentence" - as soon as decoder generates this token answer is completed.
3) UNK - "unknown token" - is used to replace the rare words that are not in vocabulary. So your sentence My name is xyz123 will be translated into My name is _unk_.
4) PAD - GPU or CPU processes training data and ensures sequences have same length. If the length in sequence is 8,the sentence of my name is xyz123 will be padded as my name is xyz123 _pad_ _pad_pad_ _pad_.

IV. PROPOSED MODEL

RNNs have feedback loops which help them to maintain information in memory.It is difficult to train models when it comes to long sentences as RNN face vanishing gradient problem.

LSTM have memory cell that keep information in memory for long periods of time as well as gates that are used to control when information enters, when its output and when its forgotten. So we are using LSTM in our model

A) Seq2Seq Model

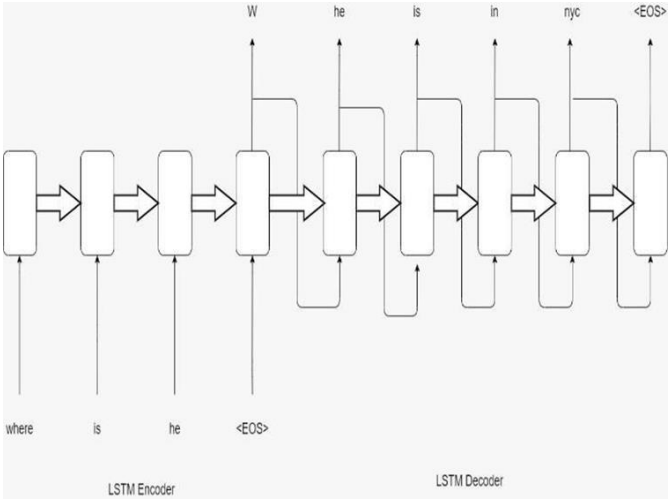

Fig 1: Seq2Seq Model

In long sentences there is a high probability that initial context will be lost by the end of sequence as the output sequence heavily relies on context by hidden state in the endmost output of the encoder.

B) Seq2Seq model with attention

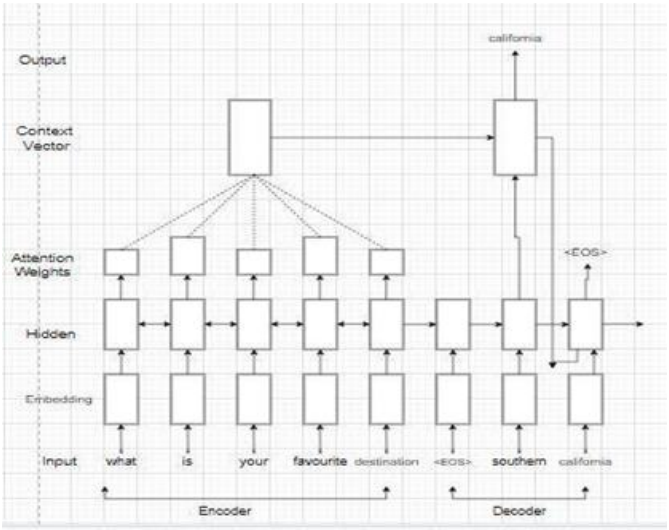

Fig 2: Seq2Seq model with attention

1) Alignment score

The score is calculated from the previous hidden state of the decoder. Instead of heavily relying on context by hidden state in endmost output of the encoder the decoder decides which part of input sentence to pay attention to The alignment vector is computed at every time step of the decoder

2) Attention weights

After applying a softmax function on alignment scores, attention weights are obtained. Softmax function gives the probabilities between 0 and 1 higher the probability of the input sequence higher will be its influence in predicting target words.

3) Context Vector

Weighted sum of attention weights and hidden states of the encoder gives context vector.

By Context vector,Decoder's output from the previous time ,Previous decoder's hidden state the target word is predicted.

V. Results

Model was trained for 500 epochs achieving an accuracy of 90 percent. The model was also deployedusing flask.

```
row 2
QUESTION: who am i
REAL ANSWER: doctor livingston
PREDICTED ANSWER: doctor livingston

row 3
QUESTION: where are you from
REAL ANSWER:  southern california
PREDICTED ANSWER:  southern california

row 4
QUESTION: the bullets
REAL ANSWER: oh the bullets
PREDICTED ANSWER: oh the bullets

row 5
QUESTION:how old are you
REAL ANSWER: nineteen or
PREDICTED ANSWER: nineteen or

row 6
QUESTION: i know
REAL ANSWER: she works with she
PREDICTED ANSWER: oh
```

Fig 3 : Real Answer and Predicted Answer

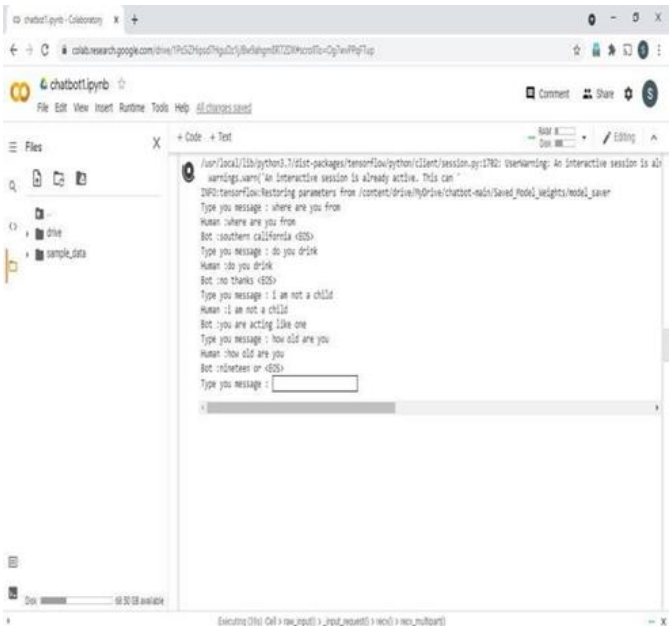

Fig 4: Chatbot using Google Colab

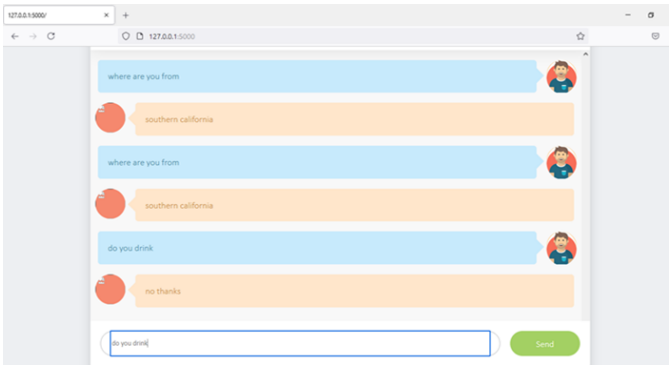

Fig 5: Chatbot Deployed On Flask

VI. FUTURE ENHANCEMENTS

We can implement this proposed model in other domains like medical, forensic, sports, college domain etc. It will be beneficial in all fields without spending much time.

VII. LIMITATIONS

1) The blunt truth to know is, chatbots are machines, not people. Although you can attempt to equip them with a casual tone, chatbots will never truly sound human.

2) Chatbots are great at providing facts and data. But emotional bond cannot be created and building an emotional bond can be a make-or- break factor in today's competitive environment.

3) Predefined or Closed-domain: chatbots only answer the questions from a closed domain, or answer those questions, which are defined in the database.

4) Language Structure: The structure of sentence making differs from language to language. For example, every language has its own rules for punctuation, text structure, and use of spaces. While existing chatbots cannot distinguish it.

5) Accuracy: Chatbots should be designed in such a way that their conversation is like a human's to complete any task. But existing chatbots are bad at suddenly changing any subject and provide an unpredictable response. Thus, we cannot achieve a satisfactory level of accuracy.

VIII. APPLICATION

1) Health Care

By easing the burden on medical professionals Chatbots have made their way into health care. A Chatbot like Super Izzy can track menstrual cycles, dates and fertile windows.

2) Transportation

Uber in partnership with Facebook has enabled users to sign up for Uber and request a ride, without having to leave Messenger or download the Uber app. Ride status updates and ride receipts are delivered to a private conversation between the customers and Uber on Messenger, making it easy to track Uber rides and payment history.

3) Market Research

The way market researchers get in touch with respondents is changing as many respondents may prefer to be contacted on WhatsApp or the Facebook messenger. Instead of having a dedicated survey app, many research firms are developing

Chatbots to have personalized, engaging conversational surveys with respondents to improve the experience and increase completion rates.

4) E-commerce

The e-commerce industry is also improving the shopping experience with Chatbots. Customers can now search and shop more conveniently with the help of chatbots.

IX. CONCLUSION

In the proposed model we have developed a AI chatbot using seq2seq algorithm with attention in which the bot will predict the relatable dialogue answers from the movie dataset that are closely related to user's question.Also,we have experimented different algorithms for building chatbot and it was observed that this algorithm turned out to be best in comparison with their accuracies and correctness.The working of the algorithm remains same,so it can be used for other applications like health care ,transportation ,e-commerce, etc

X. REFERENCES

[1]. G. K. Vamsi, A. Rasool and G. Hajela, "Chatbot: A Deep Neural Network Based Human to Machine Conversation Model," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020.

[2]. H. Wang, Z. Lu, H. Li, and E. Chen, (2013). A dataset for research on short-text conversations. In EMNLP.

[3]. Oriol Vinyals and Quoc V. Le. “A Neural Conversational Model”, 2015

[4]. P. Kandpal, K. Jasnani, R. Raut and S. Bhorge, "Contextual Chatbot for Healthcare Purposes (using Deep Learning)," 2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 2020.

[5]. Sordoni et al. and Shang et al. “Recurrent neural network to model dialogue in short conversations”, 2015.

Cite this article as :

Shashank Joshi, Sandeep Shinde, Prerna Shinde, Neha Sagar, Sairam Rathod, "Chatbot Using Deep Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 1, pp.228-232, January-February-2023.
Available at doi :
https://doi.org/10.32628/CSEIT2390148
Journal URL : https://ijsrcseit.com/CSEIT2390148