International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

ChatBot for College Website

Shubham Sonawane, Shanmughasundaram R

Abstract*: *In any service-providing organization, it is necessary to answer current or future customer's questions for their information and for maintaining good relationship with the customers. These organizations usually have call centers with hundreds and thousands of employees to give replies to customers through chat (i.e. text) or phone calls. These kind of traditional systems need large number of employees which eventually increases the overall expenditure and manpower requirement of that organization. This paper is aimed at minimizing the manpower requirement i.e. decreasing the number of employees by implementing a chat service. This concept of chatbot has been proposed for implementation in College Website that will automatically reply to the queries of stakeholders like students, parents, etc. without any human interaction. For implementation of chatbot sequence to sequence model is used which is a deep neural network.

Index Terms*: *Deep learning, Long-short term memory (LSTM), Recurrent neural network (RNN).

I. INTRODUCTION

There is a rapid growth in an organization's technological advancement with improvement in design and automation. Nowadays, technologies are developed with integration of deep learning and machine learning for improving the performance of the systems.

Chatbots are the interfaces between the available information and the end user. Any information to the end user is easily available with natural language keywords and the entire information regarding the input data is available on a single click. The main advantage of chatbot is to share information within a short span of time. End user need not go through the entire website for minute details as they are easily available via chatbot. It is unlike traditional applications where end user needs to interact with them through simple and structured language such as, "place order", "cancel", "return order", etc.

Machine learning is not only used in data science[1] but also in automobile industry[2] biomedical industry, robotics[3], sports[4] etc,. For the past few years, chatbot[5],[6] implementation is a trending concept in service providing industries, e.g., Microsoft's XiaoIce and cortana, Apple Siri, Google Assistant, Amazon's Alexa, etc.

Manuscript published on 30 August 2019.
*****Correspondence Author(s)
Shubham Vijay Sonawane, PG Student, Department of Electrical and Electronics Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India.
Shanmughasundaram R, Assistant Professor, Department of Electrical and Electronics Engineering, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India.

II. SYSTEM OVERVIEW

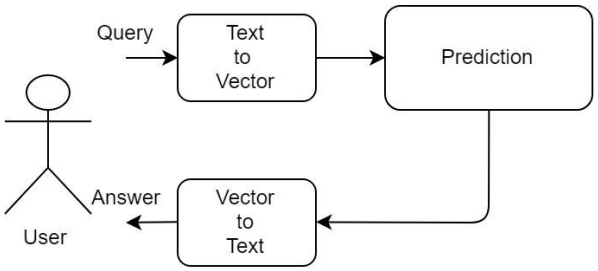

Fig. 1. System overview

The basic idea behind the project is to develop a chat service which is user friendly and can provide all the necessary information about the College. A chat window was created on a local host web server in which one can enter his/her queries related to the College that will be converted to an array of vectors. Then the model will look for the most similar vector for the given array of vectors using probability. Then the predicted array of vectors will appear in chat window as reply to the query in text form.

III. METHODOLOGY

For implementation of chatbot, three things are mainly considered viz. Sequence to Sequence model, Word embedding and Dataset.

A. *Sequence-to-Sequence model*

Sequence-to-sequence model is a deep learning neural network which is extensively used in chatbot applications, image and video captioning, neural machine translation, etc. It consists of an encoder and decoder as shown in figure 2. Encoder and decoder consists of Long-short term memory (LSTM) cells [7], [8], [10].

Encoder helps to encode text input for chatbot and then output of encoder is given to decoder which will predict the output for given text input. Sequence of text data is given to encoder and decoder while training, hence the name is 'Sequence-to-sequence model'.

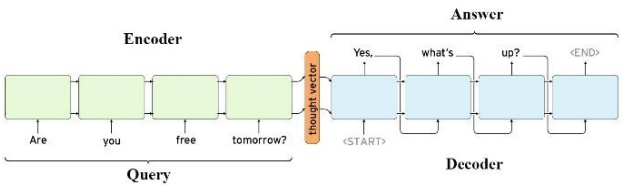

Fig. 2. LSTM Network

LSTM cell as shown in figure 3 consists of three gates.

a) Input gate:
Input gate is used for deciding whether to update cell state or not.

$$I_t = \sigma(A_i[P_{t-1}, u_t] + D_i \quad (1)$$

b) Forget gate:

Forget gate is used for deciding how much of cell state output from previous time step is allowed for the next time-step.

$$F_t = \sigma(A_f[P_{t-1}, u_t] + D_f) \quad (2)$$

c) Output gate:

Output gate is used to decide how much of current cell state is visible for next time step.

where,

$$O_t = \sigma(A_o[P_{t-1}, u_t] + D_o) \quad (3)$$

I_t is input gate.

F_t is forget gate.

D_x is biases for respective gate(x).

O_t is output gate.

σ is sigmoid function.

A_x is weight for the respective gate(x) neurons.

P_{t-1} is output of the previous LSTM block (at time-step $(t-1)$.

u_t is input at current time-step.

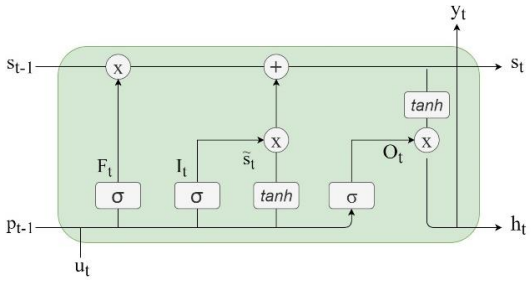

Fig. 3. LSTM cell

$$\tilde{S_t} = tanh(A_c[O_{t-1}, u_t] + D_c \quad (4)$$

$$S_t = (F_t \times S_{t-1}) + (I_t \times \tilde{S_t}) \quad (5)$$

$$P_t = O_t \times \tanh(S_t) \quad (6)$$

where,

S_t is cell state

$\tilde{S_t}$ is candidate for cell state

B. *Word embedding*

Word embedding [9] is species of string or word depiction that is used for words that have relatively identical meaning to have an equivalent depiction. It is a distributed depiction for text to enhance the realization for deep learning methods on natural language transform dilemma. Essentially, it is a type of technique where each word is represented using a real-valued vector in a pre-defined vector space. Every single word is outlined to a unique vector and the assigned vector values are well-informed in such a fashion that simulate to a neural network, so the approach is best when clubbed together with deep learning. The dense-distributed depiction is used for individual word. Every individual word is depicted by a real-valued vector, mostly 10's or 100's of dimensions. Word embedding allows to capture natural meaning of words. The size of the vector space is stated as an element of the model viz. 50, 100, or 300 dimensions.

C. *Dataset*

The dataset is created according to frequently asked questions. The dataset consists of lines of questions and answers in '.txt' format. It is created in the format that first line is question after that respective answer is written, like that whole data set is created of questions and their answers. While using dataset for training at first two python lists are created one for questions and the other for answers, then that lists are tokenized and then used for LSTM network training. To implement chatbot for college website, it is essential to provide the data regarding information of the college. By keeping this in perspective, the data-set has been collected by student survey as well as from the queries raised on the college website.

For example-

What is the duration of integrated science course?

5 years

Where is ATM in campus?

near academic block 2 and main gate

D. *Training of model for prescribed data*

Fig. 4. Training model

As described earlier, the data set was created. Next process is to train the model with the created dataset. Within this sequence, the data set is required to be pre-processed before training. While pre-processing, the data set need to follow the sequence of operations mentioned in flow chart viz. separation of data, word embedding, etc.

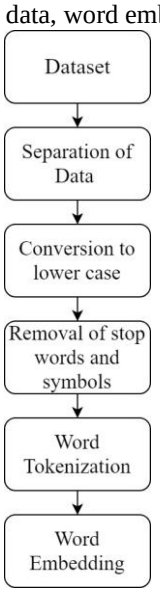

Fig. 5. Data pre-processing

a. *Data pre-processing*

Pre-processing of data is required to make entire data uniform which will be easy to interpret while training a model. Initially, data set is divided into two python lists of questions and answers.

After data separation, the entire data will be converted to lower case to make the data uniform. For better interpretation of data during training, stop words and symbols are removed. The next step is to convert each sentence into a list of words i.e. word tokenization. The final process is word embedding which converts words into vectors and this simplified processed data is used for training a model.

b. *Training of Model*

Training is a supervised learning process in which model is exposed to inputs and their respective outputs to derive relation between inputs and outputs. The data is simultaneously used for model training as well as testing. The entire data is split into 90 % for training and 10 % for testing. The accuracy of model is dependent on number of iterations.

IV. IMPLEMENTATION

The code has been written in python language under virtual environment which is a tool used for keeping dependencies for python projects. PyCharm IDE is used for coding. Encoder-Decoder model was created using TensorFlow APIs which is a library for deep neural network. Parameters used for creating model is shown in Table 1.

Table I. Model parameters

No. of RNN layers	2
No. of hidden units in RNN cell	512
Batch size	256
Learning rate	0.002
Dropout	0.9
No. of epochs	150
Activation function	Softmax

During training of the model, loss was calculated for each epoch using loss function. At the 150th epoch, loss was 0.012.

Loss function is calculated at each epoch and it calculates the error between predicted output and the target value. The error value will be used by the model while training to update the weights to minimize the loss. The loss function is shown in figure 6. To execute the model and to interact with end user, the interface between model and the end user must be in simplified manner. For creating the user friendly experience, the model has been integrated with Flask python package which will render the html page. In html page, a chat window was created to interact with end user.

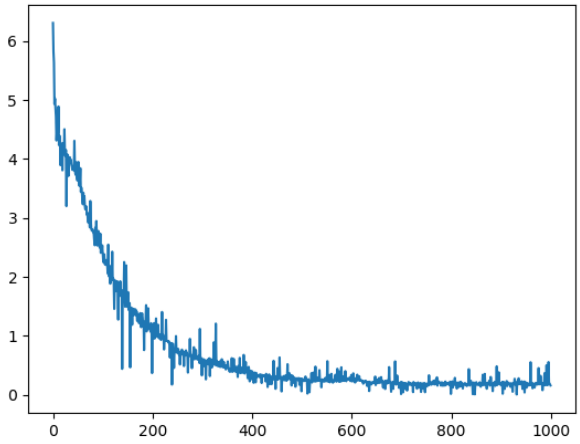

Fig. 6. Loss function

V.RESULTS AND DISCUSSION

After deploying the deep learning model, the accuracy obtained for chatbot is 92.89%. To compute the accuracy, all the questions from the dataset were asked to chatbot and the prediction was compared with the target answer. The model is deployed on web. During interaction between end user and chatbot, a chat window pops up as shown in figure 7. Also the waiting time had been shown with animation during interaction. Along with that, the chat time is also displayed. All these features are shown in figure 8 and figure 9.

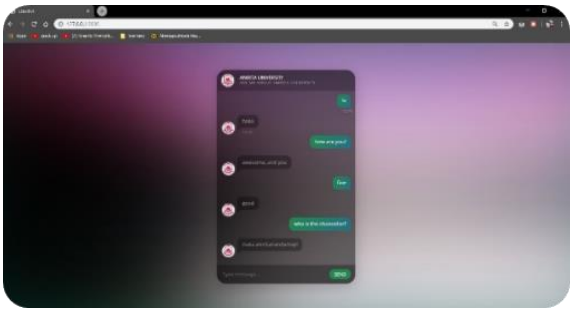

Fig. 7. Web Interface

Fig. 8. Chat Time

Retrieval Number: J88670881019/19©BEIESP
DOI: 10.35940/ijitee.J8867.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Fig. 9. Waiting time

Fig. 10A. Chatbot interaction-1

Fig. 10B. Chatbot interaction-2

VI. CONCLUSION

A chatbot for college website is implemented using deep learning algorithm. It has been observed that the chatbot requires a large dataset. The model has to be trained for more number of epochs which increases computational time. So for testing purpose, small amount of data has been provided with proper question and answer. The accuracy of the model 92.89%.

REFERENCES

1. M. Mohammadi, A. Al-Fuqaha, S. Sorour and M. Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey," in *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2923-2960, Fourthquarter 2018. doi: 10.1109/COMST.2018.2844341
2. A. Luckow, M. Cook, N. Ashcraft, E. Weill, E. Djerekarov and B. Vorster, "Deep learning in the automotive industry: Applications and tools," *2016 IEEE International Conference on Big Data (Big Data)*, Washington, DC, 2016, pp. 3759-3768. doi: 10.1109/BigData.2016.7841045
3. Xiyang Song, Huangwei Fang, Xiong Jiao and Ying Wang, "Autonomous mobile robot navigation using machine learning," *2012 IEEE 6th International Conference on Information and Automation for Sustainability*, Beijing, 2012, pp. 135-140. doi: 10.1109/ICIAFS.2012.6419894
4. I. Zelic, I. Kononenko, N. Lavrac and V. Vuga, "Diagnosis of sport injuries with machine learning: explanation of induced decisions," Proceedings of Computer Based Medical Systems, Maribor, amritaSlovenia, 1997, pp. 195-199. doi: 10.1109/CBMS.1997.596433
5. B. Setiaji and F. W. Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling," *2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, Bangkok, 2016, pp. 72-77. doi: 10.1109/ISMS.2016.53
6. H. N. Io and C. B. Lee, "Chatbots and conversational agents: A bibliometric analysis," *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Singapore, 2017, pp. 215-219. doi: 10.1109/IEEM.2017.8289883
7. Yu Wang, "A new concept using LSTM Neural Networks for dynamic system identification," *2017 American Control Conference (ACC)*, Seattle, WA, 2017, pp. 5324-5329. doi: 10.23919/ACC.2017.7963782
8. J. Kim, J. Kim, H. L. T. Thu and H. Kim, "Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection," *2016 International Conference on Platform Technology and Service (PlatCon)*, Jeju, 2016, pp. 1-5. doi: 10.1109/PlatCon.2016.7456805
9. M, Akumar, and Dr. Soman K. P., "Amrita_CEN at SemEval-2016 : Complex Word identification using word embeddings", in International Workshop on Semantic Evaluation (SemEval 2016), 2016, 2016.
10. R. Vinayakumar, K. P. Soman and P. Poornachandran, "Long short-term memory based operation log anomaly detection," 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, 2017, pp. 236-242. doi: 10.1109/ICACCI.2017.8125846

AUTHORS PROFILE

Shubham Sonawane, M.Tech (Embedded Systems Engg.), Department of Electrical and Electronics Engg., Amrita School of Engineering, Coimbatore
shubhamsonawane21@gmail.com

Shanmughasundaram R, M.Tech, Assistant Professor, Department of Electrical and Electronics Engg., Amrita School of Engineering, Coimbatore
rshanmugha@gmail.com