International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Text Classification using Recurrent Neural Network

Dr. Sankhya N Nayak [a], Mr. Devaraj F V [b], Mr. Nandish M [c], Mr. Mohan H G [d]

[a] *Dr. Sankhya N Nayak, Associate Professor, CSE, JNNCE Shimoga, 577204, India*
[b] *Mr. Devaraj F V, Assistant Professor, CSE, PESITM Shimoga, 577204, India*
[c] *Mr. Nandish M, Assistant Professor, CSE, JNNCE Shimoga, 577204, India*
[d] *Mr. Mohan H G, Assistant Professor, CSE, JNNCE Shimoga, 577204, India*

A B S T R A C T

Text can have rich data, yet separating experiences from it very well may be hard and tedious, because of its unstructured nature. Text classification is an AI method that allocates a bunch of predefined classifications to open-finished content. Text classifiers can be utilized to put together, structure, and sort basically any sort of text – from documents, medical studies, and files. Text classification is one of the major undertakings in natural language processing with large applications like sentiment analysis, topic labeling, spam detection, and intent detection.

Keywords: *Machine Learning, RNN, Recurrent Neural Networks, LSTM, text classification.*

1. Introduction

A chatbot is a software application which is based on AI and NLP to get the input from humans about their needs and deliver the desired results with the least amount of work the end customer can actually expect, which is very much similar to a virtual assistant, It is used in dialogue systems for a variety of purposes such as customer care, data management, personal assistant and collection of data. The AI, NLP and ML are extensively used in complex word classification processes in chatbot applications to search for predominantly used common buzzwords and generate responses using regular phrases pulled from relevant libraries or databases. There are also chatbot applications that just spawn. Clinical chatbots reduce responsibility for health services professionals by reducing clinic visits, reducing clinic visits and readmissions, and improving symptom information.

There are two types of chatbots. The first is a Retrival based chatbot and the another one is an Generative based chatbot.

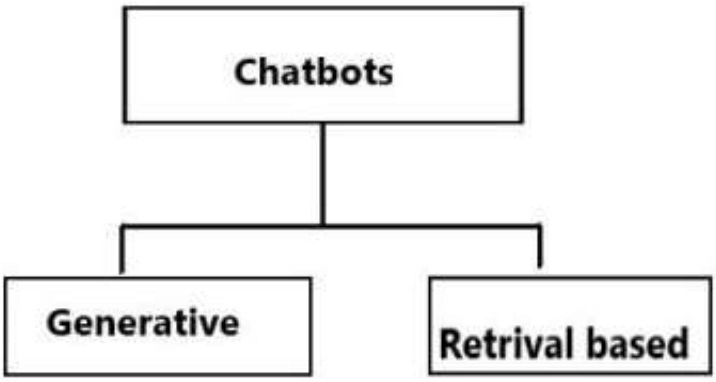

Fig.1- Categories of Chatbots.

1.1 Retrieval Based Chatbot (RBC)

The RBC is primarily based Chatbot, it is ready thru the pre- characterized responses or alternatively from the database that contains the response generated previously. Here the responses created rely on the current statistics. This type of Chatbots makes use of extraordinary techniques for offering the first-class response from the contemporary data given. The strategies applied are keyword matching, gadget gaining knowledge of, or Deep studying. The RBC mechanism do not produce any new output as they are primarily depending on the database of predefined records and accordingly produces the results [4]. The above fig indicates the working of Context and customer message is given to the model alongside the predefined responses so one can, as a result, classify the cause and bring the consumer message which can be visible thru the web interface or transportable application, and so on.

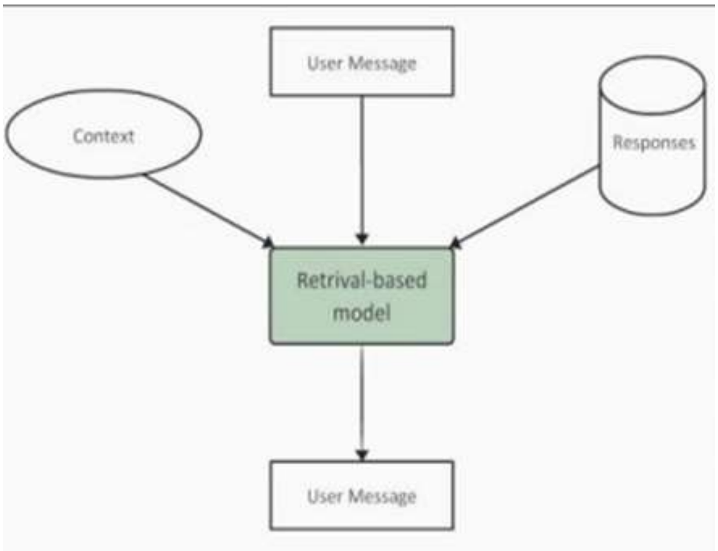

Fig.2- Retrieval based Chatbots.

1.2 Generative Based Chatbot

Another way to build a chatbot is to use generative models. Generative models do not have predefined answers as specified in pull-based models. From now on, you will create your own answers. These types of chatbots are really advanced, require a lot of information to prepare, and use complex algorithms to provide answers. As such, it is very complex and rarely used.[3] The figure above shows the generative model. Custom messages and historical data are passed to the model. This will generate a response based on the user's message. Finally, the chatbot generates a response based on the user's message and responds through a web or mobile interface.

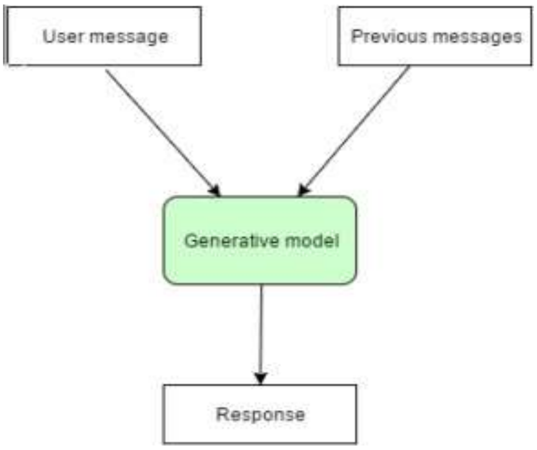

Fig. 3- Generative based Chatbots.

2. Related Work

This section talks about the previous works on the Recurrent Neural Networks and LSTM.

2.1 Recurrent Neural Networks

A recurrent Neural Network (RNN) is a kind of Neural Network where the output from the previous step is given to the contribution of the current step. In conventional neural networks, every one of the information and yield states is autonomously functional, yet there are situations when the model requirements to anticipate the following grouping of words in a sentence, due to which model requirements to recall the last words. Thus, RNN came into a presence that tackled the issue by including hidden layers. The dominating the highlight of the RNN model is the hidden layer which recalls some substance of the sequence. Basic Feed Forward Network additionally recalls what they have realized during the training. However, the distinction from the FNN is that RNN additionally recalls things gained from the old information arrangement. The RNN works on the principle of hidden layers which is the reason for the model to recollect the past inputs given to the model during training. For instance, as displayed in the figure, there are two information sources being associated with each hidden layer. Further, the hidden layers are then associated with the output layer which gives the predicted output from the model. The previous sentences are learned during the training and remembered. They are used in the next cycle in the hidden as shown hence the words are learned recurrently and remember. Once the model is done with the training it creates the anticipated output from the model for the given unique sentence. Henceforth, they are called recurrent neural organizations.

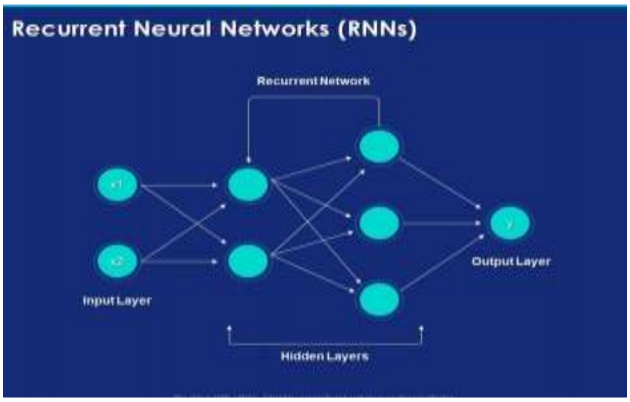

Fig. 4 - Recurrent Neural Networks.

3. Methodology

In this section, the usage of LSTM for the text classification is provided. These classified texts are helpful in chatbots for the conversations.

3.1 Recurrent Neural Networks

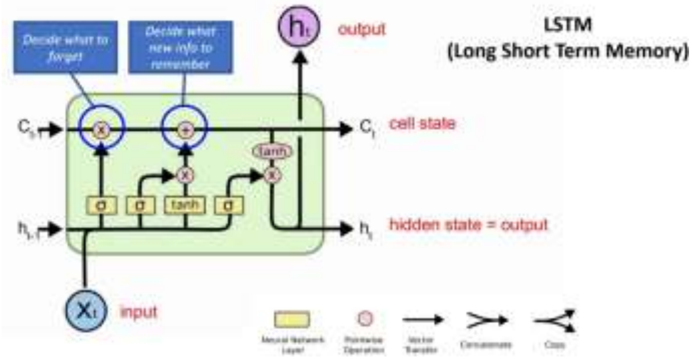

Fig. 5- Long Short Term Memory.

LSTM was intended to defeat the issues of basic Recurrent Network (RNN) by permitting the network to store informa- tion in a kind of memory that it can access at a later occasion. LSTM is an exceptional sort of Recurrent Neural Network (RNN) that can learn long haul designs.[4][5]. The way to LSTMs is the cell express, the even line going through the highest point of the chart. [5] The cell state is refreshed twice with few calculations that subsequent balance out calculations.

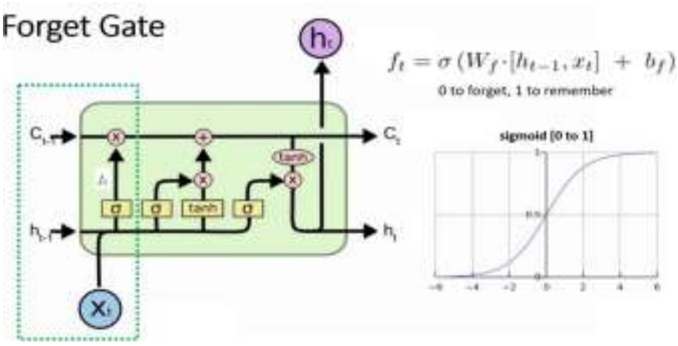

Fig. 6- Forget Gate.

Forget Gate*:* The initial step is to choose what data we will discard from the cell state. This choice is made by a sigmoid layer called the “Forget Gate”.[8][9]

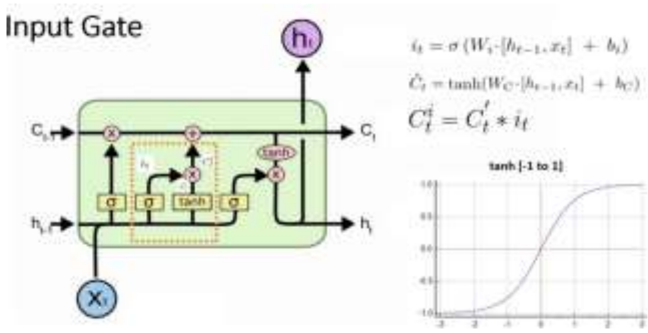

Fig. 7- Input Gate.

Input Gate: The upcoming stages, for a given cell state the new data to be stored will be selected. This process has two parts. Initial “Input Gate layer”, which is a sigmoid layer that will select the esteems that are refreshed. The second part is a vector of new applicant esteems is prepared by the tanh layer that can be added to the state [8][9].

The cell state is updated in the subsequent stage by combining these two parts.

Output Gate: At long last, the yield of the process is to be selected. It is selected based on the separated adaptation founded on the cell state. The process is started by selecting the part that yields in our cell state, it is done by running a sigmoid layer. Later on the tanh layer runs over the cell state (in order to normalize the data from - 1 and 1) and then it is duplicated as specified in [4][6].

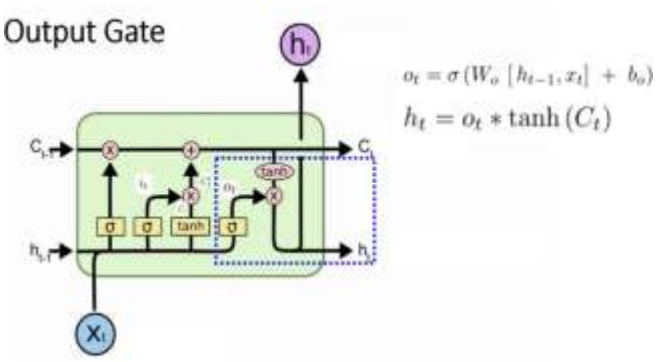

Fig. 8- Output Gate.

3.2 Embedding Layer

Embedding Layer The model starts with an implanting layer which trans- forms the info whole number lists into the relating word vectors. Word inserting is an approach to address a word as a vector. Word embeddings permit the worth of the vector’s component to be prepared. In the wake of preparing, words with comparative implications regularly have the comparative vectors.[9][10]

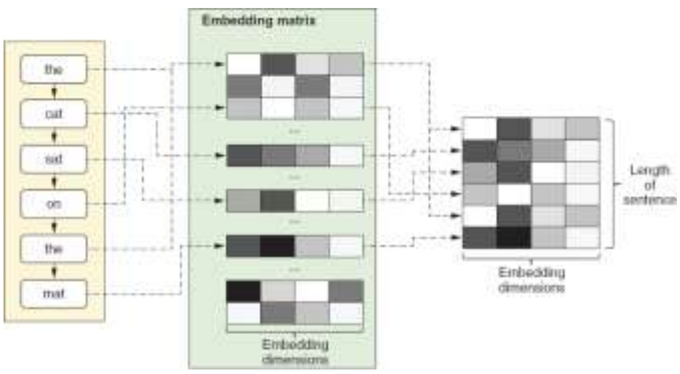

Fig. 9-Embedding Layer.

4. Results

4.1 Model Summary

Table 1: Results Obtained

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 100, 100)	10000000
spatial dropout1d	(None, 100, 100)	0
lstm (LSTM)	(None, 100)	80400
dense (Dense)	(None, 128)	12928
dropout (Dropout)	(None, 128)	0
dense 1 (Dense)	(None, 128)	16512
dense 2 (Dense)	(None, 128)	0

5. Conclusion

Multi-class text classification using RNN is found that the text classified suffers from vanishing gradient problem which minimizes the accuracy and predicting outcome. Due to this problem LSTM (Long Short-Term Memory) are used which will in turn solve vanishing problem, increases more accuracy and give proper predicting outcome. Then LSTM model can be improved by using GRU (Gated recurrent units) and Attention based graph.

References

Vagelis Hristidis, ‘Chatbot Technologies and Challenges’, In First International Conference on Artificial Intelligence for Industries, 2018.

[1] Peng Song,Chaoyang Geng,Zhijie Li, “Research on Text Classification Based on Convolutional Neural Network”, 2019 International Conference on Computer Network, Electronic and Automation (ICCNEA)

[2] Tobias Gentner, Timon Neitzel, Jacob Schulze, Ricardo Buettner, ‘A Systematic Literature Review of Medical Chatbot Research from a Behavior Change Perspective’, In IEEE 44th Annual Computers, Software, and Applications Conference, 2020

[3] Fan Zhang, Nicha Dvornek,Junlin Yang,Julius Chapiro,“Layer Embedding Analysis in Convolutional Neural Networks for Improved Probability Calibration and Classification”, IEEE Transactions on Medical Imaging PP(99):1-1

[4] LIANG ZHANG, JIE ZHOU, CHENGCAI CHEN, AND LIANG HE, “Retrieval-Polished Response Generation for Chatbot”, date of publication June 22, 2020.

[5] Myat Sapal Phyu, Khin Thandar Nwet,”A Study on a Joint Deep Learning Model for Myanmar Text Classification”, 2020 IEEE Conference on Computer Applications(ICCA)

[6] Yi-Chao Wu; Fei Yin; Zhuo Chen; Cheng-Lin Liu,” Handwritten Chinese Text Recogni-tion Using Separable Multi-Dimensional Recurrent Neural Network”, 2017 14th IAPR Interna-tional Conference on Document Analysis and Recognition (ICDAR)

[7] Jingjing Cai; Jianping Li; Wei Li; Ji Wang,” Deeplearning Model Used in Text Classifica-tion”, 2018 15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)

[8] Panitan Muangkammuen; Narong Intiruk; Kanda Runapongsa Saikaew,“ Automated Thai-FAQ Chatbot using RNN-LSTM”, 2018 22nd International Computer Science and Engi-neering Conference (ICSEC)

[9] Jenq-Haur Wang, Ting-Wei Liu, Xiong Luo, Long Wang “An LSTM Approach to Short Text Sentiment Classification with Word Embeddings”, The 2018 Conference on Computational Linguistics and Speech Processing.