International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0343

e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2371-2374

Classification of Sequential Data in Deep Learning Using LSTM Network

Nilam Deepak Padwal[1], *Dr. Kamal Alaskar*[2]

[1]*Research scholar, Computer Science and Application, Bharati Vidyapeeth (Deemed to Be University) Institute of Management Kolhapur, Thane, India.*

[2]*Research Guide, Computer Science and Application, Bharati Vidyapeeth (Deemed to Be University) Institute of Management Kolhapur, India.*

Emails: *nilam17aug@gmail.com*[1], *kamal.alaskar@bharatividyapeeth.edu*[2]

Abstract

In the empire of natural language processing the study of sentiment analysis has important significance. Sentiment analysis is a very challenging task with great requirements in every field. Understanding public opinion from a web social network is a rapid growth development technology to identify the sentiment analysis from comments obtained from social media, to obtain specific opinions from the same. The cumulative growth of social media provided big data in front of text that can immeasurably augment its specialty. Deep learning is part of the broader family of machine learning methods based on artificial neural networks with representation learning. Deep learning is famous because of its applicability and performance hence it can be called state of the art and should be an automated process. So in this paper, we discussed the attention mechanism of LSTM with neural networks has achieved good results in semantic association and classification.

Keywords: *ANN, CNN, deep learning, LSTM, machine learning, neural network, RNN.*

1. Introduction

Currently the research in sentiment analysis with various, prominent, scientific, technologies in directions, particularly in applying natural language processing and text mining technologies. Sentiment analysis finds a mixed bag of domains. Sentiment analysis is a method that includes machine learning and deep learning while the functionality of the machine learning model is relatively straightforward and feature extraction is a complex process. Deep learning uses multiple layers to progressively extract higher-level features from raw input. As deep learning is a subfield of artificial intelligence and machine learning that has been inspired by the structure of the human brain, machine learning depends upon statistical technical focus and it represents logical structure. The neural network where algorithms are ANN (Artificial Neural Network): A simple neural network layer architecture that is fully connected but does not handle sequential time series based on chronological order. A perceptron connected through arrows has weights with an input layer, hidden layer, and output layer. There are many hidden layers so deep learning of the consists of different types of neural networks. Convolutional Neural Network (CNN): It is responsible for image specification Recurrent Neural Network (RNN): It is responsible for Speech and text specification. Each hidden unit gives feedback to itself and other states but the sequences of length are larger. Through the various layers of primitive features the edges can detect complex feature-identified shapes as huge data for deep learning as it's data-hungry for the security of data it’s linearly updated. So, it is a powerful graphical processing unit with more memory for matrix calculations. As the ANN is not handling chronological and sequential data. While in the RNN sentences, it converts to vectors and states. That means the hidden unit gives feedback but it's problematic when the length of sequences is more and decision-making depends upon the starting stage so we switch to LSTM which is long short-term memory. Decision making of any text, if it depends upon the first word then maybe a vanishing ingredients problem arises. Now word-by-word processing of any text maintains two types of context as a core concept of LSTM means they ca

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0343
e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2371-2374

LSTM means they can communicate with each other through Short-term context and long-term contexts. So the block diagram of LSTM is, Figure 1 shows the Block diagram of LSTM Architecture

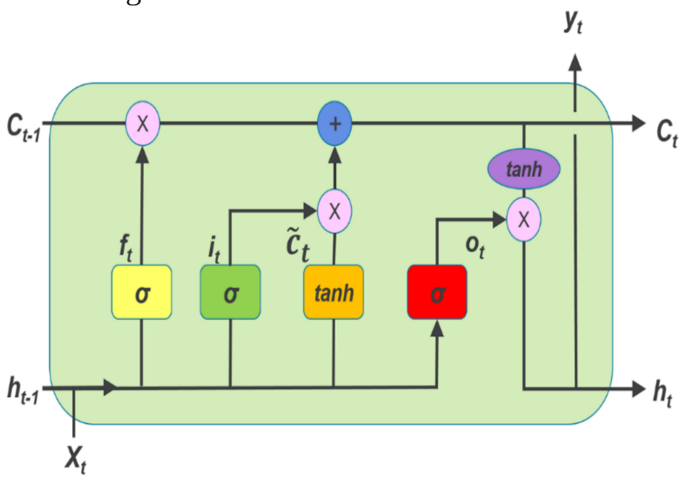


Figure 1 Block Diagram of LSTM Architecture

LSTM network with two parallel lines one is long term which represents the cell state and another is short term which represents the hidden state through which they can interact with each other. The gates are the complex architecture of LSTM, which is to update the cell state and calculate the hidden state. In LSTM layer accepts the input vector x which may be notified and gives output y. Train the input vector in sequences of vectors, any error occurs then it is the sum of deviations of all target signals from corresponding computed by the network.

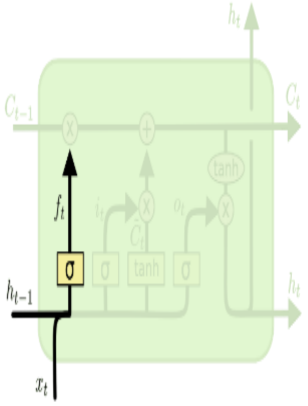
Figure 2 Simple Structure of LSTM

Consider the above diagram in which input can be from the previous input state, previous hidden state, or input for the current timestamp 'h'. In processing state update the cell state, can be updated based on current input $x_t \rightarrow ct_0$ i.e. content removed from the long-term memory and x_t which added necessary information on correct input. Calculate the h_t i.e. output of hidden state. Then we get two outputs current hidden state and the current self-state which means we perform processing as three inputs and obtain two outputs. Based on the current input we decide which things vanish from the long state or which necessary information can be added to the current input state. One kind of the complex architecture of LSTM is divided into three gates. Figure 2 shows the Simple structure of LSTM.

2. Forget Gate

To remove something from the cell state. Where h_t and c_t are vectors, collections of numbers, with an equal number of dimensions. tx is also one kind of vector. To calculate an x_t when we want to solve the sentiment analysis problem to obtain results in 1(positive) or 0(negative) by using one-hot encoding, Bag of words, TF – IDF, these all are text vectorization techniques. Now f_t c_t and o_t are vectors with dimensions are shape and side are same as h_t and c_t. Pointwise operations are addition, multiplication, and tan h which are held between two vectors. Neural network layer which carries multiple nodes in which they carry activation functions. The activation function is basically Linear or identical activation function in which $f(x) = x$ means input and output are the same. Nonlinear activation function consists of four types Sigmoid or logistic activation function, softmax activation function, tan h or hyperbolic activation function, and ReLU (Rectified Linear unit) activation function. The nonlinear function ranges between 0 and 1 so, sigmoid $f(x) = 1 \div (1 + e\bar{}x)$.In the classification problem, we use the sigmoid (whose performance like a logit is very small then the value closes to zero and the logit is a very large value close to one) or softmax nonlinear function (that represents softmax output vector to be probability distribution gives a better idea about confident prediction).Now neural networks have the same number of nodes so in forget gate with pointwise multiplication that means context is removed from the cell state that may be in the long-term state. In point-wise operation$(Ct - 1)$ removes the things we call removal of things as mentioned in the diagram below. Figure 2 shows the Forgot gate. , it is a powerful graphical processing unit with more memory for matrix calculations. At the ANN not handling sequential data.

e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2371-2374

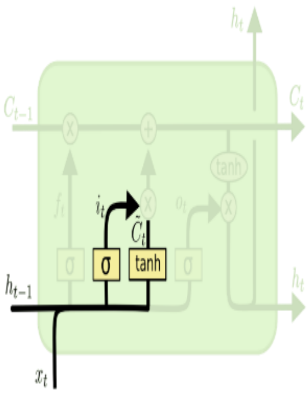


$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right)$$

Figure 2 **Forget Gate**

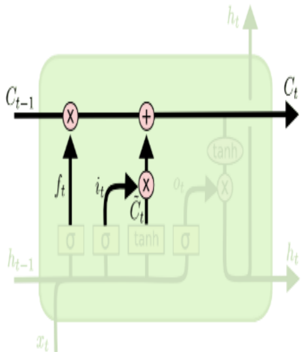
3. Input Gate

when we want to add some new information to the cell state from the previous hidden State then we need to calculate the candidate cell state based on the current input and previous hidden State, some candidate values are evaluated to add to the cell state.ie $C\overline{t}$ it has the potential to be added in a long-term context. the value of i_t that decides which value from the candidate cell state is added to the cell state is one kind of filtering process. Filtering from C_t and adding the neural network activation function is sigmoid. Input form h_t-1 and tx with Fully Connected weight matrix and bias b_i. Calculate c_t is also called the current cell state so, it carries forward the Information means carrying long-term context. So, Figure 3 shows the Input gate Calculate the Candidate cell state with the filtering process. Figure 4 shows the Input gate Calculate the current cell state



$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right)$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Figure 3 **Input Gate Calculate the Candidate Cell State with The Filtering Process**

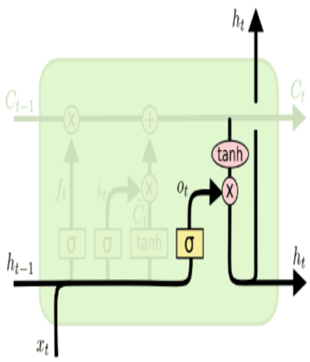


$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Figure 4 **Input Gate Calculate the Current Cell State**

4. Output Gate

A current time stamp to decide output or the value of hidden state i.e. tan h C_t value is got in the range [-1, 1], and filtering from tan h is not. So below Figure 5 shows the output of the current time stamp.



$$o_t = \sigma\left(W_o\, [h_{t-1}, x_t] + b_o\right)$$
$$h_t = o_t * \tanh(C_t)$$

Figure 5 **Output Gate**

Conclusion

In contemporary years, deep learning has driven a new wave of artificial intelligence playing an irreplaceable role in the field of NLP, it is a main concept involving constructing a neural network model for text sentiment classification. Sentiment analysis based on deep learning has attracted an increasing number of research due to powerful performance. So this paper proposed. LSTM as a memory cell and gate unit a multiplicative input gate unit introduced to protect the memory content stored and from perturbation by irrelevant input. Likewise, a multiplicative output gate unit is introduced to protect another unit from perturbation by currently irrelevant memory content stored. LSTM leads to many more successful runs and learning faster,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2024.0343
e ISSN: 2584-2854
Volume: 02
Issue: 07 July 2024
Page No: 2371-2374

solving complex artificial long-time leg tasks that have never been solved by previous recurrent network algorithms. LSTM extracts the information conveyed by temporal order widely separated input, so can learn all input expectations. For nontrivial tasks, the LSTM is recommended.

References

[1]. Guixian Xu, Yueting Meng, Xiaoyu Qiu, Ziheng Yu, And Xu Wu, “Sentiment Analysis of Comment Texts Based On Bilstm”, Special Section On Artificial Intelligence and Cognitive Computing for Communication and Network, Digital Object Identifier 10.1109/Access.2019.2909919.

[2]. Josh Patterson, Adam Gibson: “Deep Learning”, Printed in the United States of America. Published by O’Reilly Media, CA 95472. Copyright © 2017.

[3]. Nikhil Buduma, ”Fundamentals of Deep Learning”, Published by O’Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, © 2015

[4]. Sepp Hochreiter, “Long Short-term Memory”, article in Neural Computation, December 1997 DOI: 10.1162/neco.1997.9.8.1735 · Source: PubMed.

[5]. Sepp Hochreiter, “LSTM can solve hard long time lag problems”, Advances in Neural Information Processing Systems · January 1996.

[6]. Xavier Glorot, Yoshua Bengio, “Understanding the difficulty of training deep feedforward neural networks”, 13th International Conference on Artificial Intelligence and Statistics (AISTATS) 2010, Chia Laguna Resort, Sardinia, Italy.

[7]. Zhulu Chu, Xihan Wang, Meilin Jin, Ning Zhang, Quanli Gao and Lianhe Shao, “An Effective Strategy for Sentiment Analysis Based on Complex-Valued Embedding and Quantum Long Short-Term Memory Neural Network”, Axioms 2024, 13, 207. “https://doi.org/ 10.3390/axioms13030207”.

[8]. https://www.researchgate.net/publication/365903956_Review_on_sentiment_analysis_for_text_classification/link/6388c662658cec21049d4a43/download

[9]. https://www.google.com/search?q=block+diagram+of+lstm+network&rlz=1C1ONGR_enIN1086IN1086&sourceid=chrome&ie=UTF-8#vhid=HdlqNCjRTMIS_M&vssid=l