**Text Emotion Classification Using LSTM**

Ahmad Nadhirsyah Jaman, Alexander Maximillian, Alif Brazali, Richard Aldy and Ali Abdurrab

1 Computer Science, Binus University, DKI Jakarta, Indonesia

2 Computer Science, Binus University, DKI Jakarta, Indonesia

3Computer Science, Binus University, DKI Jakarta, Indonesia

4Computer Science, Binus University, DKI Jakarta, Indonesia

5Computer Science, Binus University, DKI Jakarta, Indonesia

E-mail: [ahmad.jaman@binus.ac.id](mailto:ahmad.jaman@binus.ac.id)

[alexander.maximillian@binus.ac.id](mailto:alexander.maximillian@binus.ac.id)

[richard.aldy@binus.ac.id](mailto:richard.aldy@binus.ac.id)

[alif.brazali@binus.ac.id](mailto:alif.brazali@binus.ac.id)

[ali.abdurrab@binus.ac.id](mailto:ali.abdurrab@binus.ac.id)

Received xxxxxx

Accepted for publication xxxxxx

Published xxxxxx

Abstract

There are various human emotions, for example, sad, happy, afraid, etc. Humans tend to show their emotions when communicating, whether through direct or indirect communication. However, sometimes it is difficult to understand the emotion conveyed through text. In this paper we try to solve this problem by making a text emotion classification model using LSTM. We make the model using python with the help of several libraries, such as NLTK, Keras, scikit-learn, etc. The dataset that we used contains 7,029 data for each class and total 42,174 data that is split into 72% training dataset, 8% validation dataset, and 20% testing dataset. The results show that the LSTM model is able to handle the text classification with very high accuracy. This result conclude that this model is ready to be used for real world data and may even be used to help in research studies if it is ever needed.

Keywords: *text classification, emotion classification, classification, information retrieval, LSTM*

1. Introduction

Human emotion is a complex component in this world, because there are so many emotion that human have, such as happy, sad, angry, disgust, fear, etc. When we talk to another person we can easily recognize their emotion by their body language, the most easy way to recognize their emotion is by looking into their eyes, by doing that we can determine their emotion. We can recognize their emotion by listening to their voice, for example when people are sad they tend to speak in a low tone and sounded gloomy. But it will be hard to recognize an emotion through a text, because we can’t see their face, eyes, lips, or hear how they sound. That is why through this paper we will experimenting a text emotion classification using LSTM.

It is important to make a computer to understand and recognize human emotion, we need to understand the sources of emotion inside our body[1]. As I said in the first paragraph, emotion could be expressed verbally by understanding some known words they talk or non-verbally like their tone of voice, body languages including facial expression.

Recurrent Neural Networks (RNNs) are one of the most prevalent architectures because of the ability to handle variable-length texts. Humans can't analyze from scratch every second. Any human can understand each word based on his understanding of previous words. He doesn’t throw everything away and start thinking from scratch again. His thoughts have persistence. Traditional neural networks can’t do this, and it seems like a speed process is coming. For example, imagine a human want to classify what kind of event is happening at every point in a movie. It’s not clear how a traditional neural network could use its reasoning about previous events in the film to inform later ones. Recurrent neural network addresses can face this type of issues. They are networks with multiple loops in them, allowing information to continue. Though RNNs are capable of modeling long sequential data theoretically they fail to represent long sequences in real time applications [2].

Recently, LSTM is most popular to deal with sentiment classification. LSTM is proposed by Hoch Reiter and Schmid Huber in 1997 and was refined and popularized by many people in the following work. They work tremendously well on large different types of problems and are now widely used. LSTMs are explicitly designed to ignore the long-term dependency problem [3]. Remembering information for a long time is practically their default behavior, not something they struggle to learn. All recurrent neural networks have the form of a chain of repeating modules of the neural networks. In the level of RNNs, this repeating module having a very simple structure, such as a single tanh layer. The IMDB benchmark dataset is used for our experimental studies that contain movie reviews that are classified as being positive or negative.

By using Long Short Term Memory (LSTM) we hope we can classify human emotion. LSTM is one of Recurrent neural network (RNN), LSTM was created with a purpose to overcome the problem of long-term dependency in RNN[4]. Long-term dependency can be difficult to learn from RNN because it is trained by back-propagation through time (BPTT).

2. Theory

## Recurrent Neural Network

Recurrent Neural Network (RNN) is a chain-like neural network where each node stores information from the previous node on the chain and uses it in calculating the output.

The input and output of an RNN are sequences, where each node would receive one input from the input sequence with information from the previous node and send one output to the output sequence with information to the next node.

RNN can be used to process data that have a connection between them, such as words in a sentence and daily weather. An example of RNN is prediction of exchange rate based on previous rates.

## Long Short-Term Memory

Long Short-Term Memory (LSTM) is an RNN architecture used in deep learning. In LSTM, information sent to the next node contains two values; the output of the node and "memory", which is data stored in the cell state. The cell state runs through the network, and in each node, the memory will be affected by the input and also affect the output. LSTM has been successful in handling problems such as speech recognition, handwriting recognition, and music composition.

 LSTM networks are well-suited to classifying, processing, and making predictions supported statistic data since there may be lags of unknown duration between important events in a very statistic. LSTMs were developed to accommodate the exploding and vanishing gradient problems that may be encountered when training traditional RNNs. Relative insensitivity to gap length is a bonus of LSTM over RNNs, hidden Markov models, and other sequence learning methods in numerous applications. There are several architectures of LSTM units. a typical architecture consists of a cell (the memory a part of the LSTM unit) and three "regulators", usually called gates, of the flow of knowledge inside the LSTM unit: an input gate, an output gate and a forget gate. Some variations of the LSTM unit don't have one or more of those gates or even produce other gates. as an example, gated recurrent units (GRUs) don't have an output gate.

LSTM with a forget gate

The compact forms of the equations for the forward pass of an LSTM unit with a forget gate are:



Where the initial values are c0=0 and h0=0 and the operator o denotes the Hadamard product (element-wise product). The subscript t indexes the time step. In this model, σ is the sigmoid activation function, tanh the hyperbolic tangent activation function, Xt the input at time t, Wi, Wc, Wf, Wo, Ui, Uc, Uf, Uo are weight matrices to regulate the input and bi, bc, bf , bo are bias vectors.

3. Method

For this project, we used Python as the programming language as it has a lot of libraries to help with machine learning and deep learning. The libraries used are pandas and NumPy to process the dataset, NLTK for data cleaning, TensorFlow and Keras for model creation, scikit-learn to pre-process data and provide model metrics, and Matplotlib to help with plotting data.

The dataset used in this project is a dataset of sentences with emotion label, containing 21,405 sentences with 6 emotion labels. The dataset can be downloaded from Kaggle over at https://www.kaggle.com/ishantjuyal/emotions-in-text.

## Data Preprocessing

From the 6 possible labels, the dataset are skewed where the most common label appears 7,029 times and the least common label appears 879 times. To balance the data, we do oversampling method to raise the data of each class into 7,029 data. In total there are 42,174 data to be processed. The data is then cleaned by applying lowercase and removing symbols and stopwords. The cleaned sentences are then tokenized to be prepared as input for the neural network. The dataset is then split into training and testing dataset, with the split being 80% of the dataset for training and 20% for testing.

## Modeling

The first part of the neural network is an embedding layer, where each word in the input sentence is turned into a vector. After the embedding layer is an LSTM layer of 100 nodes. At the end is a fully-connected layer of 6 nodes to determine the output label.

## Testing and Evaluation

The main metric used to evaluate the model is accuracy. Accuracy is calculated from dividing the amount of correct predicted labels by the total amount of test labels. The higher the model accuracy, the better the model will be.

4. Result

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Training Accuracy** | **Validation Accuracy** | **Testing Accuracy** |
| BPNN | 99.7% | 95.6% | 94.4% |
| LSTM | 99.5% | 96.7% | 96.8% |

Table 1. Models Accuracy

We also made a BPNN model for comparison purpose. Based on the table we can see that both of the models have a slight overfit. However, the BPNN model have a worse overfit than LSTM model. The reason is that BPNN is not meant for sequential data. We found that both models have very high training accuracy.

We can see that the validation and testing accuracies are a bit lower than training accuracy, however they are still very promising. Overall, LSTM model has higher accuracies on all occassions. This shows that for this classification problem, LSTM have an advantage over BPNN, although both are still very much usable.

5. Conclusion

In this paper, we have proposed an LSTM based sentiment classification approach for text data. Users from all over the world express and openly share their opinions on various topics. Manual analysis of such large amounts of data is very difficult, so a reasonable need for their computer processing has arisen. Sentiment analysis processes people's opinions and attitudes towards products, services, politics, social events, and company strategies. The text method using LSTM has better accuracy than BPNN with better sentiment classification performance with an accuracy of 96.8% compared to BPNN when the amount of training data is more. In the future we plan to expand this study to a greater extent where different embedding models can be considered across a wide variety of data sets.

References

1. Alhagry, Fahmy and El-Khoribi 2017. Emotion Recognition based on EEG using LSTM Recurrent Neural Netowrk
2. Sepp Hochreiter and Jurgen Schmidhuber. Long short-term memory. Neural computation, 9(8):1735–1780, 1997
3. Dr. G. S. N. Murthy, Shanmukha Rao Allu, Bhargavi Andhavarapu and Mounika Bagadi, Mounika Belusonti, "Text based Sentiment Analysis using LSTM", International Journal of Engineering Research and, vol. 9, no. 05, 2020. Available: 10.17577/ijertv9is050290.
4. PRANJAL SRIVASTAVA, Essentials of Deep Learning: Introduction to Long Short Term Memory, DECEMBER 10, 2017.
5. Hoch Reiter S, Schmidhuber, J. (1997). Long short-term memory. In: Neural Computation 9(8): 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735