Sentiment Analysis By Using NN and CNN

A Data Science Project in Python

Lin Xu

Master of Science in Computer System Engineering

College of Engineering, Northeastern University

Boston, USA

Xu.lin1@husky.neu.edu

Abstract

Sentiment analysis or opinion mining is used to automate the detection of subjective information such as opinions, attitudes, emotions, and feelings. I examine sentiment analysis on Sentiment 140 datasets. I use two different model to try the result and see the different between them. Most of these kind of analysis are using RNN but in my test, I try the NN and CNN model instead of RNN. In order to make this easily to understand and only classify the reviews into two classes: Negative and Positive.

Finally I evaluatie the accuracy and loss of different models and test the result of an input. The comparison depends on two data sets: training data set and test data set.

1 INTRODUCTION

Microblogging websites have evolved to become a source of varied kind of information. This is due to nature of microblogs on which people post real time messages about their opinions on a variety of topics, discuss current issues, complain, and express positive sentiment for products they use in daily life. In fact, companies manufacturing such products have started to poll these microblogs to get a sense of general sentiment for their product. Many times these companies study user reactions and reply to users on microblogs. One challenge is to build technology to detect and summarize an overall sentiment[1]

In this paper, I look at sentiment 140 which allows you to discover the sentiment of a brand, product, or topic on Twitter.

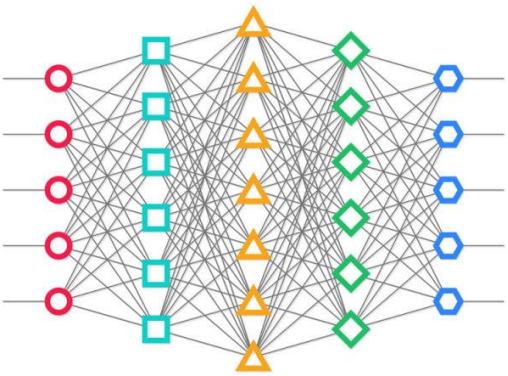
There has been a large amount of research in the area of sentiment classification. Traditionally most of it has focused on classifying larger pieces of text, like reviews[3]. But reviews represent summarized thoughts of authors, tweets are more casual and limited to 140 characters of text.

The work of Pang et al. has served as a baseline and many authors have used the techniques provided in their paper across different domains.[2]

In order to train a classifier, supervised learning usually requires hand-labeled training data. In this dataset, the data had already been classified into two sets, one is train data and the other is test data. Even though the data had been classified the style and format need to be handled, so it can be used to train a model.

Recent days RNN with LSTM is a good way to process natural language, but in my paper, I focus on normal neural network and convolutional neural network, so it can help me understand the difference between them, and it will help me with my further study.

1.1 Introduction to Neural Network



Neural networks are a type of machine learning models which are designed to operate similar to biological neurons and human nervous system. These models are used to recognize complex patterns and relationships that exists within a labelled dataset.

Key concepts in a Neural Network

A. Neuron:
A Neuron is a single processing unit of a Neural Network which are connected to different other neurons in the network. These connections repersents inputs and ouputs from a neuron. To each of its connections, the neuron assigns a "weight" (W) which signifies the importance the input and adds a bias (b) term.

B. Activation Functions
The activation functions are used to apply non-linear transformation on input to map it to output. The aim of activation functions is to predict the right class of the target variable based on the input combination of variables. Some of the popular activation functions are Relu, Sigmoid, and TanH.

C. Forward Propagation
Neural Network model goes through the process called forward propagation in which it passes the computed activation outputs in the forward direction.

$Z = W*X + b$
$A = g(Z)$

- g is the activation function
- A is the activation using the input
- W is the weight associated with the input
- B is the bias associated with the node

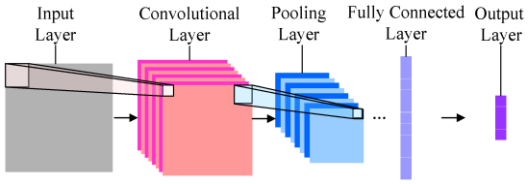
D. Error Computation:
The neural network learns by improving the values of weights and bias. The model computes the error in the predicted output in the final layer which is then used to make small adjustments the weights and bias. The adjustments are made such that the total error is minimized. Loss function measures the error in the final layer and cost function measures the total error of the network.

Loss = Actual_Value - Predicted_Value

Cost = Summation (Loss)

E. Backward Propagation:
Neural Network model undergoes the process called backpropagation in which the error is passed to backward layers so that those layers can also improve the associated values of weights and bias. It uses the algorithm called Gradient Descent in which the error is minimized and optimal values of weights and bias are obtained. This weights and bias adjustment is done by computing the derivative of error, derivative of weights, bias and subtracting them from the original values.

1.2 Introduction to Deep Neural Network - Convolutional Neural Network



Key components of Convolutional Neural Network.

A. Convolutional layer

In this layer, a kernel (or weight) matrix is used to extract low level features from the images. The kernel with its weights rotates over the image matrix in a sliding window fashion in order to obtained the convolved output. The kernel matrix behaves like a filter in an image extracting particular information from the original image matrix. During the colvolution process, The weights are learnt such that the loss function is minimized.

B. Stride

Stride is defined as the number of steps the kernel or the weight matrix takes while moving across the entire image moving N pixel at a time. If the weight matrix moves N pixel at a time, it is called stride of N.

C. Pooling Layer

Pooling layers are used to extract the most informative features from the generated convolved output.

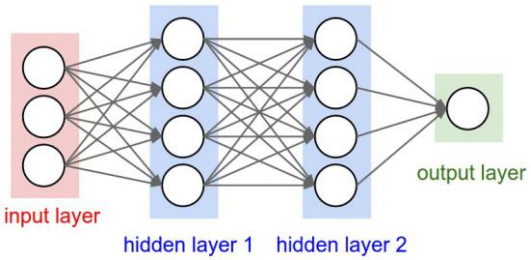
D. Output Layer

To generate the final output, a dense or a fully connected layer is applied with the softmax activation function. Softmax function is used to generate the probabilities for each class of the target variable.[6]

2. APPROACH

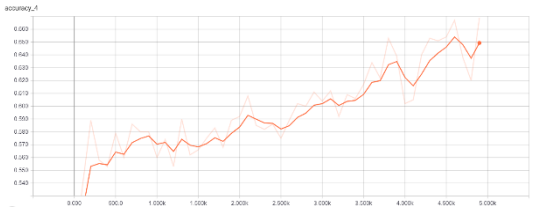
2.1 Train Neural Network Model

Designed a two hidden layer neural network model, input size is 7739 which represents the number of words in the wordlist, 1500 neuron in each hidden layer.

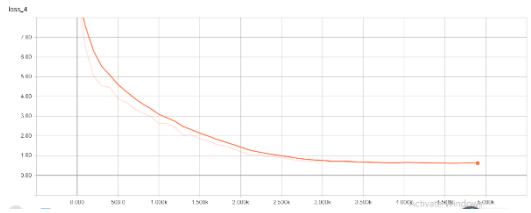


After train the model, I get the accuracy and loss.

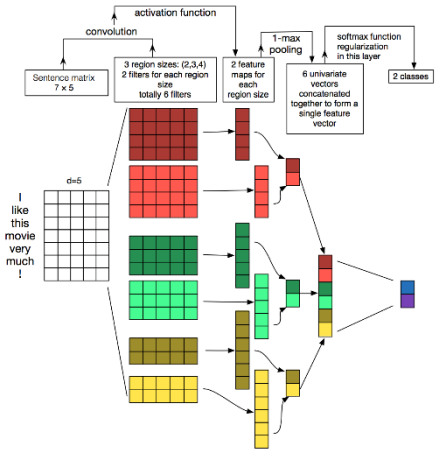
Accuracy



Loss



2.2 Train Convolutional Neural Network Model



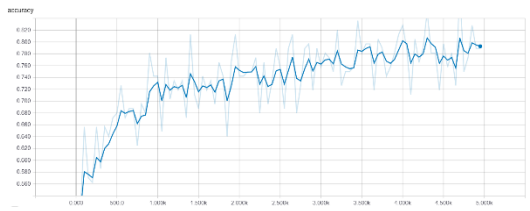
input: X's shape is [None, 8057], Dictionary's Length is 8057. By embedding_lookup, Convert X to [None, 8057, 128]

After expande_dims get result [None, 8057, 128, 1]

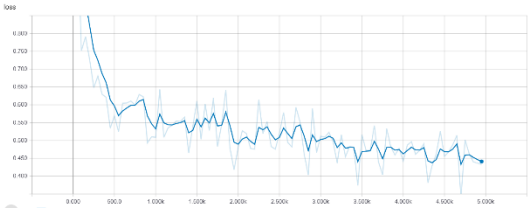
Convolution and pooling : Convolution Kernel shape is [3/4/5, 128, 1, 128], length, width, channel and number of kernel

Result after convolution [None, 8057-3+1, 1, 128], width change to 1 so the pooling layer shape is [1, 8055, 1, 1] after [None, 1, 1, 128]

Accuracy of CNN:



Loss of CNN



3. RESULTS

After train the model, I test the result by using NN model. It can produce right answer, but sometimes it also produced some wrong answer. We can understand it, since the accuracy is only 0.66. That's why CNN appear and now RNN is commonly used in NLP.

4. FUTURE WORK

Machine learning techniques perform well for classifying sentiment in tweets. In this Project, I classify the overall sentiment of a tweet. These days more and more app have the function of review, sentiment analysis should be more accuracy and can be applied in different length of text. Even though there is a better model named RNN, but also there are some defects in different models. So next I will try to understand them and may make them better.

5. REFERENCES

[1] Mohey El-Din, Doaa. (2016). Analyzing Scientific Papers Based on Sentiment Analysis (First Draft). 10.13140/RG.2.1.2222.6328.

[2] Go, Alec & Bhayani, Richa & Huang, Lei. (2009). Twitter sentiment classification using distant supervision. Processing. 150.

[3] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up Sentiment classification using machine learningtechniques. In Proceedings of the Conference onEmpirical Methods in Natural Language Processing (EMNLP), pages 79–86, 2002.

[4] Text classification with movie reviews https://www.tensorflow.org/tutorials/keras/basic_text_classification

[5] Creating training and test data sets and preparing the data https://blog.griddynamics.com/creating-training-and-test-data-sets-and-preparing-the-data-for-twitter-stream-sentiment-analysis-of-social-movie-reviews/

[6] A Very Comprehensive Tutorial : NN + CNN https://www.kaggle.com/shivamb/a-very-comprehensive-tutorial-nn-cnn