International Journal of Computer Science and Engineering (IJCSE)
ISSN (P): 2278-9960; ISSN (E): 2278-9979
Vol. 8, Issue 3, Apr-May 201; 29-36

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

REVIEW ANALYSIS USING SPARSE VECTOR AND DEEP NEURAL NETWORK

T. Thiyagu, Anant Vatta, Sneh Upadhyay & Rajat Chaudhary

Research Scholar, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India

ABSTRACT

In the modern world of social networks, sentiment analysis has become one of the most important research topics in the field of deep learning. The Sentiment Analysis is the name of the problem that with a sentence or text the machine gets capable to analyze and predict with the maximum precision possible. The sentiment that will be obtained by a person when reads it or the contextual opinion related to something. In this project, we are implementing deep neutral network and sparse vector algorithm to implement sentiment analysis. The above algorithm can perform high data analysis and gives enhanced accuracy of sentiments implemented. The approach can reduce the training time of the neural network model through a regional embedding. At the same time, sparse vector algorithm uses a sentence to extract sentiment features of the whole sentence and controls the transmission of information through different weight matrices, which can effectively infer the sentiment polarities of different targets in the same sentence. Finally, experimental results of data show that our approach yields better performance than SVM and several other neural network models. Experiments on different data and different domains show that the proposed method can solve the high dimensional problem with good performance and in various review units.

KEYWORDS: *Machine Learning, Deep Neural Network, Embedding, Sentiment Analysis, Sparse Vector Algorithm*

Article History

Received: *19 Apr 2019* **|** ***Revised:*** *04 May 2019* **|** ***Accepted:*** *21 May 2019*

INTRODUCTION

Sentiment analysis or opinion mining is the computational study of people's opinions, sentiments, emotions, appraisals, and attitudes towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. 1 The inception and rapid growth of the field coincide with those of the social media on the Web, for example, reviews, forum discussions, blogs, micro-blogs, Twitter, and social networks, because for the first time in human history, we have a huge volume of opinionated data recorded in digital forms. Since early 2000, sentiment analysis has grown to be one of the most active research areas in natural language processing (NLP). It is also widely studied in data mining, Web mining, text mining, and information retrieval. In fact, it has spread from computer science to management sciences and social sciences such as marketing, finance, political science, communications, health science, and even history, due to its importance to business and society as a whole. This proliferation is due to the fact that opinions are central to almost all human activities and are key influencers of our behaviors. Our beliefs and perceptions of reality, and the choices we make, are, to a considerable degree, conditioned upon how others see and evaluate the world. For this reason, whenever we need to make a decision we often seek out the opinions of others. This is not only true for individuals

but also true for organizations.

Since about a decade ago, deep learning has emerged as a powerful machine learning technique4 and produced state-of-the-art results in many application domains, ranging from computer vision and speech recognition to NLP. Applying deep learning to sentiment analysis has also become very popular recently. This paper first gives an overview of deep learning and then provides a comprehensive survey of the sentiment analysis research based on deep learning.

LITERATURE SURVEY

- To incorporate the methods of measuring the favourability content of phrases into tools by Tony Mullen and Nigel Collier. National Institute of Informatics (NII), Tokyo Using Semantic Orientation with PMI, Osgood Semantic Differentiation with Word net. The System has benefits but it's depict the result in the star rating.
- To use different machine learning techniques which are more reliable by Oscar Romero Llombart. Univers it at Autonoma De Barcelona Using the Naive Bayes, Neural Networks, Support Vector Machines. It uses the different machine learning methods are optimized to yield maximum performance. It's Data is optimized for efficiency but degrades the quality of data.
- Comprehensive Survey of Sentimental Analysis based on Deep Learning by Lei Zhang, Shuai Wang, Bing Liu. University of Illinois a Chicago. This research uses Deep Learning, Word Embedding, Recurrent Neural Network shows the result for almost the data and don't provide reliable results for test cases.
- Proposal of CNN-RLSTM by Siyuan Chen, Chao Peng, LinsenCai, LanyingGuo. East China Normal University Shanghai, China This research has used Convolutional Neural Network, Regional Long Short Term Memory.
- Depict the results in the graph but it's hard o understand and Graphs still use numbers instead of word by using CNN.

EXISTING SYSTEM

Sentiment Analysis Using Support Vector Machines with Diverse Information Sources

In the earlier days, there was Star Based Review System in which the object was being judged on the basis of the stars they receive. The star-based review system couldn't be that much appreciated as it was the decision of a particular group of people who could review it and give stars on the basis of their opinion. By using sentimental analysis the whole concept of review system can be changed. Here the review can be taken from each individual who has experienced the thing and he can provide his opinion, by this the opinions of the audience can be used and displayed in the vector form which will provide the better reviews as compared to that of Star Based Review System.

PROPOSED SYSTEM

Review Analysis Using Sparse Vector and Deep Neural Network

Introduction

Deep learning architectures such as deep neural networks, deep belief networks, and recurrent neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation where they have produced results comparable to and in some cases superior to human experts.

Sentiment Analysis refers to the use of natural language processing, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information.

Experiments on different data scale and different domains will be conducted and the ability to solve the high dimensional problem with good performance is expected as an outcome.

Problem Statement

In this proposed review system we have used Sparse Vector and Deep Neural network for getting review more precisely using Sentimental Analysis. This can be further used in the deep review of product on an E-commerce platform. This can also be implemented on the movie rating review system, to get more precise reviews of movies on Sparse vector machine and represent using graph.

METHODOLOGY

The working of the system can be divided into 2 major modules: Training and Testing.

These two modules along with other steps can be divided into 5 basic modules:

MODULE 1: Importing dependencies and downloading training test data

MODULE 2: Configure and import data into Tensor-Flow model

MODULE 3: Building and executing Linear Classifier

MODULE 4: Use Embedding with DNN model

MODULE 5: Visualizing Embedding.

Modules

Importing Dependencies and Downloading Training Test Data

In supervised Deep learning concept, the machine always needs to train with demo data and then only its starts predicting the outputs with maximum precision.

Import matplot lib.pyplot as plt

Import numpy as np

Import pandas as pd

Import tensorflow as tf

Matplotlib is the python library used for plotting framework.

Numpy is the python library used for the mathematical application.

Pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with “relational” data.

TensorFlow™ is an open source software library for high -performance numerical computation.

Configure and Import Data into Tensor Flow Model

```
ds = tf.data.TFRecordDataset(input_filenames)

ds = ds.map(_parse_function)

Features = {

"terms": tf.VarLenFeature(dtype=tf.string),

  "labels": tf.FixedLenFeature(shape=[1], dtype=tf.float32)

 }

Parsed_features = tf.parse_single_example(record, features)

Terms = parsed_features['terms'].values

Labels = parsed_features['labels']
```

The tf.data API supports a variety of file formats so that you can process large datasets that do not fit in memory and enables you to build complex input pipelines from simple, reusable pieces.

The filenames argument to the TFRecordDataset initializer can either be a string, a list of strings, or a tf.Tensor of strings.

The tf.data.TextLineDataset provides an easy way to extract lines from one or more text files.

Following the above code parse the training and test data (which is in TFRecord format) and return a dict of the features and the corresponding labels.

Building and Executing Linear Classifier

By using tensorflow class tf.LinearClassifier inheriting from the class estimator, build a model with 50 terms to train a linear model to classify instances into one of the multiple possible classes.

```
Terms = ("enjoyed", "favorite", "horrible", "brilliant", "highly" ……)

Terms_feature_column=tf.feature_column.categorical_column_with_vocabulary_list(key="terms",
vocabulary_list=terms)

Classifier = tf. estimator. Linear Classifier

Classifier. train

Evaluation_ metrics = classifier. evaluate

Print("Test set metrics:")

For m in evaluation_ metrics:

Print(m, evaluation_ metrics[m])
```

classifier. train and classifier. evaluate is used to retrieved dataset for training the machine and evaluate the program by executing test code respectively.

Use Embedding with DNN Model

By using tensorflow class tf.DNNClassifier inheriting from the class estimator, it trains and evaluate the same as linear classifier but in an efficient way because it uses embedding column. An embedding column takes sparse data as input and converts it into a lower dimensional vector as output.

Terms_ embedding_ column = tf. feature_ column. embedding_ column(terms_ feature_ column, dimension=2)

Feature_columns=tf.feature_column.categorical_column_with_vocabulary_list(key="terms", vocabulary_ list=terms)

Classifier =tf. estimator. DNN Classifier(

Feature_columns=[tf.feature_column.indicator_column(terms_feature_column)],)

In this particular model, embedding_column are set to 2D which has number line attribute for x-y axis and shows accurate examined valued for positive or negative review by placing relevant words in the plot(x,y).

Visualizing Embedding

By using python library “numpy and matplotlib”, dataset can visualize in terms of graph.

Import num py as np

Import mat plotlib.pyplot as plt

importing data values from DNN classifier and create labeled matrix

Embedding_ matrix = classifier. get_ variable value(data path')

By using for loop, relevant data represented in the form of the plot after checking and comparing from user’s review.

Using the embedding algorithm will take sparse data as input and convert it into a 2D vector as output.

Forterm_index in range(len(terms)):

Term_vector = np.zeros(len(terms))

Embedding_xy = np.matmul(term_vector, embedding_matrix)

Plt.text(embedding_xy[0],embedding_xy[1],terms[term_index])

Plt.show()

Flow Diagram

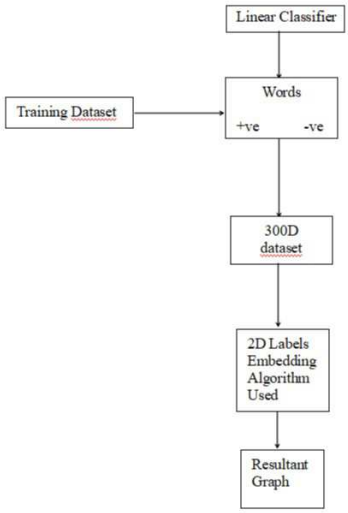

Figure 1: Flow Diagram

CONCLUSIONS

In this paper, we proposed a sentimental analysis system, which can be used to efficiently analyze the tone of a text. The system is made using Deep Neural Networks and Sparse Vector Algorithms. Few test cases show that the system is efficient in the analysis of any provided text. This system can be taken as a base for many future works involving sentimental analysis.

Future Enhancements

By using the proposed system with a few alterations many future enhancements can be brought. The system can be used for doing a deep review of a product on e-commerce platforms. This can also be used for the movie rating review system, providing precise reviews of movies using graphical representations.

REFERENCES

1. *Tony Mullen and Nigel Collier.National Institute of Informatics (NII), Tokyo 2004*
2. *To incorporate the methods of measuring the favourability content of phrases into tools*
3. *Analysing Sentiments Expressed on Twitter by UK Energy Company Consumers*
4. *Apply Word Vectors for Sentiment Analysis of APP Reviews*
5. *Aspect-Level Sentiment Analysis on E-Commerce Data*
6. *Business reviews classification using sentiment analysis*
7. *Exploring public sentiments for live-able places base on a crowd-calibrated sentiment analysis mechanism*

Fine-grained Sentiment Analysis Based on Sentiment Disambiguations

8. *B. Suneetha, S. Hemachandra, B. Gowri prasad & A. Jayanth,Application of Neural Network in Load Shedding and Some Predictable Functioning of Load Shedding Methods International Journal of Electronics, Communication & Instrumentation Engineering Research and Development(TJPRC : IJECIERD), Volume 3, Issue 2, May-June 2013, Pp 45-50*
9. *Sentiment Analysis: Arabic Sentiment Lexicons*
10. *http://www.aclweb.org/anthology/W04-3253*
11. *https://ieeexplore.ieee.org/document/8554619*
12. *https://ieeexplore.ieee.org/document/7811108*
13. *https://ieeexplore.ieee.org/document/8597286*
14. *https://ieeexplore.ieee.org/document/7426090*
15. *https://ieeexplore.ieee.org/document/7752312*
16. *https://ieeexplore.ieee.org/document/7976546*
17. *https://ieeexplore.ieee.org/document/8125054*