ISSN: 2277-2685
IJESR/Jan.2020/Vol-10/Issue-1/1- 7

A Study of Deep Learning Classification Methods vs. Conventional Machine Learning for Sentiment Analysis

Mr. G Ahmed Zeeshan[1], Dr. T Narayana[2], Dr. E Mohan[3]

Assistant Professor[1], Associate Professor[2], Professor[3]
Department of CSE
Global Institute of Engineering and Technology

Abstract:

People nowadays are more and more turning to online social networks and services as a means of communication and advocacy. In order to ascertain the general public's stance on a certain issue, product, or subject, sentiment analysis involves cataloguing and classifying these views. With each passing day, sentiment analysis becomes more and more crucial. The capacity for computers to learn new things without being explicitly taught is the result of machine learning, a subfield of computer science. Among the many branches of machine learning, "Deep Learning" focuses on algorithms that use neural implementations, such as neural networks, neural beliefs, and so on. Evaluating feelings for a certain set of data requires the use of the most practical and precise method possible as this has implications for both producers and consumers. Various machine learning, deep learning, and hybrid approaches are compared in this project's proposed research. After comparing their accuracy for sentiment analysis, it's clear that deep learning algorithms often provide higher outcomes. On the other hand, there are situations when the disparity in accuracy between the two approaches is not significant enough to warrant using one machine learning method over the other, especially when other, more practical methods are available.

I. Introduction

To find out whether a product, service, news item, etc. is getting good, neutral, or negative reactions, sentiment analysis looks at how people feel about it or reads reviews written about it. The intensity of one's feelings regarding a service or product is considered alongside its polarity. On the other hand, sentiment analysis provides customers with a more accurate picture of the product's advantages and disadvantages. Resources that provide personal opinions, such as review sites, blogs, etc., are becoming more accessible and popular in the modern day. Reviews posted on social media platforms like Twitter, Facebook, and Instagram are subject to sentiment analysis. Reviewers may have had personal experiences with a product that is for sale on an e-commerce site like Amazon. Sentiment analysis is useful for dealing with comments that include a lot of language or conflicting opinions. Companies may learn more about customer sentiment and their place in the market as a result. On top of that, they may assist their company expand while also better satisfying customers' demands and desires. There are several degrees of sentiment analysis. Levels of sentiment analysis are selected according to available time and the significance of the results for a subsequent job. We begin with analysis at the sentence level. Every sentence needs an object and a subject. We get our own impressions of a thing via subjectivity. Therefore, the subjective aspect of a statement may be categorized as positive, negative, or neutral. A different tier is the feature tier. Due to its emphasis on phrases and words, this level is superior. It takes an emotionally charged word and finds out who it's meant to be spoken to. The system then determines whether the word is neutral, positive, or negative. In this way, we eliminate as many deceptive remarks as possible. Document Level sentiment analysis is the third level. This technique determines if the document is

generally favorable or negative by looking at it in its entirety. The nuanced aspects of facts and feelings are ignored. This means that feature level is usually the one that gets picked.

II. LITERATURE SURVEY

This chapter describes the research literature relevant to the primary aspects of this thesis. The core aspects of this thesis are machine learning applications to natural language processing and deep learning classification techniques. Both these fields have received a lot of attention in the past years and there are a number of popular texts with relevant background material. As there is an enormous amount of literature available on both these aspects, these works can be described along several dimensions.

1) Research on text sentiment analysis based on CNNs and SVM

Sentiment analysis of Internet reviews is a hot research topic in Web information mining. The traditional text sentiment analysis method is mainly based on emotion dictionary or machine learning. However, its dependence on emotion dictionary construction and artificial design and extraction features makes the generalization ability limited. In contrast, depth models have more powerful expressive power, and can learn complex mapping functions from data to affective semantics better. In this paper, a Convolutional Neural Networks (CNNs) model combined with SVM text sentiment analysis is proposed. The experimental results show that the proposed method improves the accuracy of text sentiment classification effectively compared with traditional CNN, and confirms the effectiveness of sentiment analysis based on CNNs and SVM

2) Convolutional Recurrent Deep Learning Model for Sentence Classification

As the amount of unstructured text data that humanity produces overall and, on the Internet, grows, so does the need to intelligently to process it and extract different types of knowledge from it. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been applied to natural language processing systems with comparative, remarkable results. The CNN is a noble approach to extract higher level features that are invariant to local translation. However, it requires stacking multiple convolutional layers in order to capture long-term dependencies, due to the locality of the convolutional and pooling layers. In this paper, we describe a joint CNN and RNN framework to overcome this problem. Briefly, we use an unsupervised neural language model to train initial word embeddings that are further tuned by our deep learning network, then, the pre-trained parameters of the network are used to initialize the model. At a final stage, the proposed framework combines former information with a set of feature maps learned by a convolutional layer with long-term dependencies learned via long-short-term memory. Empirically, we show that our approach, with slight hyperparameter tuning and static vectors, achieves outstanding results on multiple sentiment analysis benchmarks. Our approach outperforms

several existing approaches in term of accuracy; our results are also competitive with the state-of-the-art results on the Stanford Large Movie Review data set with 93.3% accuracy, and the Stanford Sentiment Treebank data set with 48.8% fine-grained and 89.2% binary accuracy, respectively. Our approach has a significant role in reducing the number of parameters and constructing the convolutional layer followed by the recurrent layer as a substitute for the pooling layer. Our results show that we were able to reduce the loss of detailed, local information and capture long-term dependencies with an efficient framework that has fewer parameters and a high level of performance.

3. Deep Learning for Sentiment Analysis on Google Play Consumer Review

In recent years, there has been an increasing interest in sentiment analysis on consumer reviews to understand the opinion polarity on social media. However, little attention has been paid to the development of deep learning for sentiment analysis on consumer reviews in Chinese. The research objective of this paper is to explore the impact of deep learning for sentiment analysis on Google Play consumer reviews in Chinese. A web mining technique was implemented for collecting 196,651 reviews on Google Play. We used Long Short-Term Memory (LSTM) deep

learning model, Naïve Bayes (NB), and support vector machine (SVM) approaches for sentiment analysis on consumer reviews and compared the experimental results. The experimental results suggest that the accuracy of deep learning for sentiment analysis on Google Play consumer review achieves 94% and deep learning approach outperforms Naïve Bayes (74.12%) and Support Vector Machine (76.46%) in the present study. Our finding confirmed that sentiment analysis on Google Play consumer review with deep learning is outstanding. The contributions of this paper are three-fold. First, the present study confirmed sentiment analysis with deep learning on Google Play consumer review may improve the accuracy of prediction. Second, we create a sentiment dictionary named iSGoPaSD for Google Play review. Third, the study compared the result of average sampling data and non-average sampling data. We found that deep learning method with non-average sampling data reached the better performance.

4. Application of Deep Learning to Sentiment Analysis for recommender system on cloud

Sentiment analysis of short texts like single sentences and reviews available on different social networking sites is challenging because of the limited contextual information. Based on the sentiments and opinions available,

developing a recommendation system is an interesting concept, which includes strategies that combine the small text content with prior knowledge. In this paper, we explore a new application of Recursive Neural Networks (RNN) with deep learning system for sentiment analysis of reviews. The proposed RNN-based Deep-learning Sentiment Analysis (RDSA) recommends the places that are near to the user's current location by analyzing the different reviews and consequently computing the score grounded on it. Deep Learning is used to optimize the recommendations depending on the sentiment analysis performed on the different reviews, which are taken from different social networking sites. The Experiments performed indicate that the RNN based Deep-learning Sentiment Analysis (RDSA) improvises the behavior by increasing the accuracy of the sentiment analysis, which in turn yields better recommendations to the user and thus helps to identify a particular position as per the requirement of the user need.

5. Comparative study of machine learning techniques in sentimental analysis

Sentimental Analysis is reference to the task of Natural Language Processing to determine whether a text contains subjective information and what information it expresses i.e., whether the attitude behind the text is positive, negative or neutral. This paper focuses on the several machine learning techniques which are used in analyzing the sentiments and in opinion mining. Sentimental analysis with the blend of machine learning could be useful in predicting the product reviews and consumer attitude towards to newly launched product. This paper presents a detail survey of various machine learning techniques and then compared with their accuracy, advantages and limitations of each technique. On comparing we get 85% of accuracy by using supervised machine learning technique which is higher than that of unsupervised learning techniques

III.SYSTEM ANALYSIS

1.1. EXISTING SYSTEM

In existing approach sentiment analysis is done by using pure NLP for analyzing the textual data in a statistical way. Many Natural Language Processing tasks are being implemented to analyses massive amounts of data. The analyzed data is then classified into positive or negative sentiment. But analyzing the sentiment of every user behavior will be very difficult. We can’t predict the behavior of a specific person more accurately using NLP only.

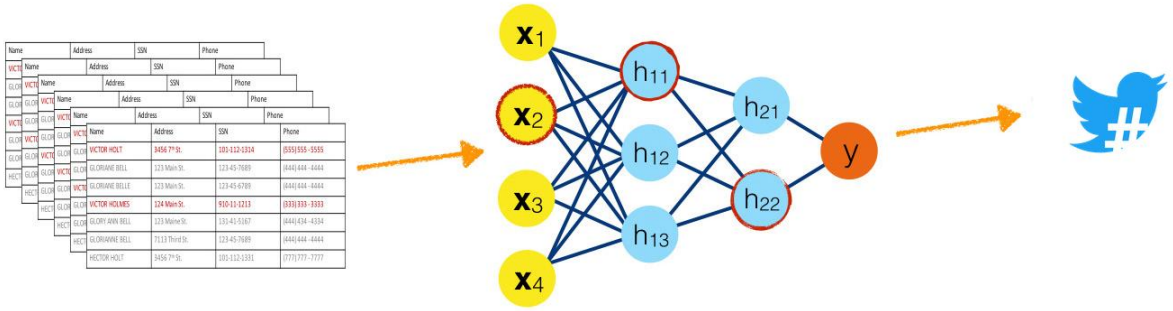
1.2. PROPOSED SYSTEM

In this proposed application we are using NLP along with Machine Learning and Deep Learning for user behavior analysis. Proposed sentiment analyzer will extract the sentiment from online textual documents for a specific user. The sentiment classification is a binomial classification in which we categorize the statement in two criteria i.e. positive, negative or neutral. In this study, we compare several machine learning as well as deep learning techniques to perform sentiment classification on the dataset. The dataset used for this task is a corpus of English language Tweets. Our research work is split into two stages where the data is pre-processed in the first stage using various text-mining techniques. Subsequently in the second stage, we built multiple models with the use of machine learning and deep learning techniques to identify the sentiment of the proclamations on the Twitter data. The overarching goal of this research is to discover the most suitable approach for any individual or organization that may want to perform sentiment analysis on the textual data and get desired result.

IV.SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture of Comparative Evaluation for Traditional Machine Learning and Deep Learning Classification Techniques for Sentiment Analysis.



4.1. System Architecture

V. SYSTEM IMPLEMENTATION

5.1. MODULES

- User
- Admin
- Data Preprocessing
- Sentiment Classification
- Data Visualization

Module description:

5.1.1 User:

Sentiment analysis is the analysis of the emotions or reviews of the consumers or users to determine whether a product, service, news, article, etc. is generating positive, neutral or negative responses. Not just the polarity, but also the depth of the feeling towards the particular service or product is taken into account. Alternatively, sentiment analysis also helps the consumers get a better idea of the pros and cons of the product or service.

5.1.2 Admin:

The aim of admin is to approve the machine learning users. the entire data must be gathered to admin. Sentiment analysis is done at many levels. Depending on the amount of time that can be dedicated to

sentiment analysis and its importance for a proceeding task, the level of sentiment analysis is chosen. First is sentence level analysis. A sentence has a subject and an object. Subjectivity gives us the subjective views of the object. Thus, the subjective part of the sentence is classified into positive negative or neutral. Another level is the feature level. This is the best level as it takes into account words and phrases. It takes a word with an emotion and determines the target of the word. It then classifies the word into positive, negative or neutral. Thus, maximum misleading comments are removed. The third level is known as Document Level sentiment analysis. This method takes the complete document as a whole and classifies it as an overall negative or positive document. It does not take into account the finer details of facts and emotions. Thus, in most cases feature level is chosen above the rest.

5.1.3 Data Preprocessing:

In this module we are going to collect the twitter based textual information then we are going to filter and clean the unnecessary information using data pre-processing and we can achieve fully formatted dataset

In this module we divide dataset into two parts for performing algorithmic training.

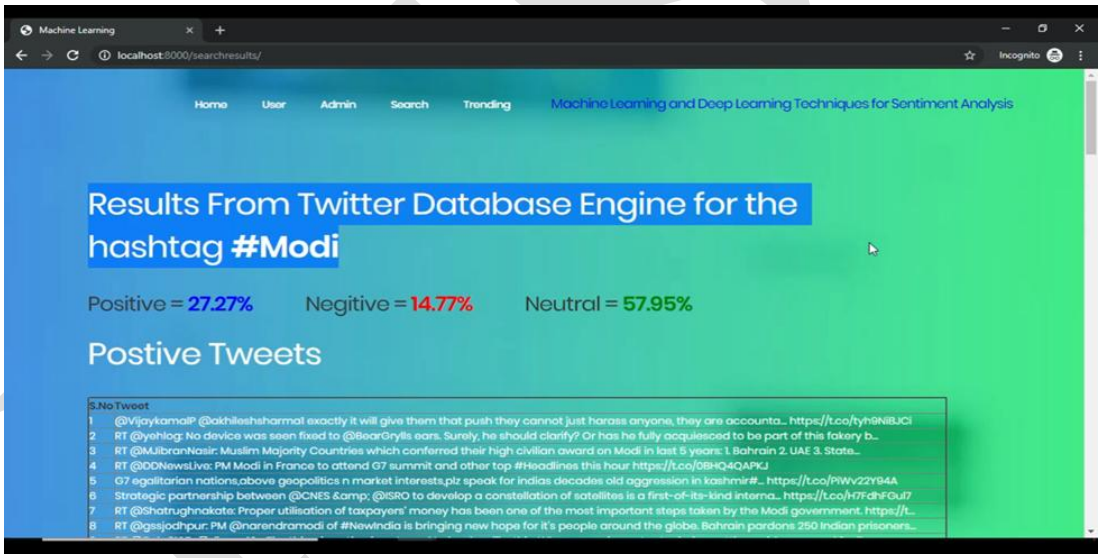
5.1.4 Sentiment Classification:

In this module we are selecting an algorithm called SVM classifier. It is used to classify the sentiment behavior of a user with the help of NLP.

5.1.5 Data Visualization:

In this module we are visualizing the all-algorithmic accuracy report for better understanding purpose.

VI. RESULTS



6.This figureshowsthe Tweets page of the project

Fig. 6.1This figure showsthe Naïve Bayes Results

VII.CONCLUSION AND FUTURE WORK

Sentiment analysis is a growing field with many applications. Based on the result of the sentiment analysis, not only are customer needs met, but also manufacturers, producers etc. get an idea about the response of the user or customer, thus ensuring that they can make the required changes. We have observed after the execution of project that the techniques of machine learning and deep learning give comparatively superior and substantial results. Machine learning methods are more basic in nature and easier to implement. They give substantial results. Deep learning methods as well as the combination of deep learning and machine learning methods are superior and complex in nature. In most cases they give better results than traditional machine learning algorithms. In some rare cases however, the difference in the accuracy of the two techniques are not very high and, in such cases, the deep learning method only increases the complexity of solving.

REFERENCES :

The paper "Research on Text Sentiment Analysis based on CNNs and SVM" was presented at the 2018 IEEE 13th Industrial Electronics and Applications Conference (ICIEA) by Yulin Chen and Zhi Zhang.
The authors of the 2018 article "Convolutional Recurrent Deep Learning Model for Sentence classification" are Abdalraouf Hassan and Ausif Mahmood.

[3] "Deep Learning for sentiment Analysis On Google Play Consumer Review" by Min-Yuh Day and Yue-Da Lin

Conference on Information Reuse and Integration (IRI) 2017, organized by the IEEE.

Using a deep learning technique and decision trees, Arman S. Zharmagambetov and Alexander A. Pak conducted sentiment analysis on a paper at the 2015 12th International Conference on Electronics, Computers, and Computation (ICECCO).

A Novel Deep Learning Architecture for Sentiment Classification, by Rahul Ghosh, Kumar Ravi, and Vadlamani Ravi, presented at the 3rd International Conference on Recent Advances in Information Technology (RAIT-2016).

[6] "Review Sentiment Analysis Based on Deep Learning" presented at the 2015 IEEE 12th International Conference on e-Business Engineering by Thongchai Hu, Jianqing Hu, Weifeng Ding, and Xiaolin Zheng.

In their 2017 paper "Application of Deep Learning to Sentiment Analysis for Recommender System on Cloud," G. Preethi, P. Venkata Krishna, V. Saritha, Mohammad S. Obaidat, and Sumanth Yenduri presented at the IEEE International Conference on Computer, Information and Telecommunication Systems (CITS).

In their 2017 paper "Sentiment classification: Feature selection based approaches versus deep learning," AlperKursatUysal and Yi Lu Murphey presented their findings at the IEEE International Conference on Computer and Information Technology.

The paper "Sentiment Analysis using Neural Network" was presented at the 2nd International Conference on Next Generation Computing Technologies (NGCT-2016) by Akshi Kumar and RituRani.

10. "Sentiment analysis in Twitter using machine learning techniques" (M S Neethu and R Rajashree, 2013) in Proceedings of the Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT).

In their 2014 paper presented at the Seventh International Conference on Contemporary Computing (IC3), GeetikaGautam and Divakar Yadav discuss sentiment analysis of Twitter data using machine learning algorithms and semantic analysis.

An overview of sentiment analysis using machine learning techniques, by Peng Yang and Yunfang Chen, [12]

IEEE 2017 Second International Conference on Networking, Electronics, and Automation Control (ITNEC).

[13] "Application of machine learning techniques to sentiment analysis" by Anuja P. Jain and Padma Dandannavar

The 2016 Second International Conference on Theory and Practice of Computer and Communication Technologies (iCATccT).

In a 2017 paper presented at the International Conference on Innovative Communication and Computational Technologies (ICICCT), B.K. Bhavitha, Anisha P. Rodrigues, and Niranjan N. Chiplunkar compared several machine learning methods for sentiment analysis.