|| Volume 6 || Issue 7 || July 2021 || ISSN (Online) 2456-0774
INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH
AND ENGINEERING TRENDS

NOVEL TECHNIQUE FOR SENTIMENT ANALYSIS USING TEXT AND FACIAL RECOGNITION

Nishiket Bansode[1] , Prof.H. U. Joshi[2]

Student, Department Of Computer Engineering JSPM's, Imperial College Of Engineering Research Wagholi,Pune-412207[1]

Professor, Department Of Computer Engineering JSPM's, Imperial College Of Engineering Research Wagholi,Pune-412207[2]

-- ***---

Abstract In recent successes of deep learning in the many fields of natural language processing, past studies of emotion recognition of Twitter user for the most part fo- cused around the utilization of lexicons and basic classifiers on pack-of-words mod- els. The central question of this study is whether it can enhance their performance utilizing deep learning. To this end, it exploits hash tags to make three extensive emotion-labeled data sets comparing to various orders of emotions. At that point analyze the performance of a few word and character-based repetitive and convo- lutional neural systems with the performance on pack of-words and latent semantic indexing models. Moreover details check the transferability of the last hidden state representations between various classifications of emotions and whether it is con- ceivable to assemble a unison model for predicting every one of them utilizing a common representation. It is demonstrate that repetitive neural systems, particularly character-based ones, can enhance over pack of-words and latent semantic indexing models. While classify the tweet emotion, semantic of the token needs to be con- sidered. Semantic of the token stored in hash map to get searched easily. In spite of the fact that the exchange abilities of these models are poor, the recently proposed training heuristic delivers a unison model with execution similar to that of the three single models.

Keywords: *: Sentimental Analysis, Emotion Recognition, Twitter, Social Media, Hashtag, Tweet.*

---***---

I INTRODUCTION

The development of social network platforms has given people a new way to generate and consume a great deal of information on the web. In the past, people used to get information from portal websites. A large number of websites provide a long list of topics varying from politics to entertainment. These traditional online information sources are useful but less efficient because they often contain redundant information. However, since the arrival of online social network platforms, people tend to get information from these platforms because of their fast and efficient fea- tures. These platforms are available for users to choose the information source they are interested in. And also a large number of social network platforms such as Twit- ter, Google+, and Facebook provide information for users.

Twitter is the most popular microblogging platform in the world. It is also the fastest growing social network platform and has a dominant position in the area of mircroblogging. More than 500 million registered users post 340 million twit- ter messages every day, sharing their opinions and daily activities. Compared with regular microblogging platforms, Twitter messages are much shorter. You are only allowed to post 140 characters or less in one Twitter message. This feature makes Twitter easier for people to get the main point from the massive amount of informa- tion available online. Depending on the need of the users, Twitter users can follow whichever people and information source they prefer. With all of the advantages mentioned above, Twitter thus has become a powerful platform with many kinds of information from worldwide breaking news to purchasing products at home. [1] [3] In the last few years, the information streams on Twitter have experienced an unbelievable increase in the popularity of this social network. The users dispose a massive amount of information about different aspects. However, not all of the in- formation is useful for users and each user has their own interests and preferences.

There is urgency for users to have personalized services. Nowadays, more and more personalized services are provided to benefit the users. People need this personal- ized service to make their fast-paced lives more efficient. Every day, a large amount of information is published by users on the Twitter platform. These data relate to users behavior and many research studies therefore focus on Twitter and this data collection. One of the research studies in the field of Twitter is user modeling. In order to provide a personalized service, researchers started to explore ranking and recommendations of web resources referenced from Twitter. A large amount of re- search focus on modeling users interests based on users published tweets data. [5]Regardless of the tweets content and Twitters potential use, researchers also no- ticed that tweets often convey pertinent information about the users emotional states. Emotion analysis on Twitter has thus become an important research issue in the mi- cro blogging area. Most research related to emotion focuses on the sentiment clas- sification on Twitter. A number of features and methods for training classifier for sentiment on Twitter platform have been researched in the past few years with vary- ing results. There are also some other research studies related to emotion analysis on Twitter. One of the studies in this area is about getting feedback about products by extracting the customers emotion on the Twitter platform. Also, investigating public attitudes by extraction of emotions from Twitter messages has been the focus of pre- vious studies.

In our daily life Fake news is an important issue on social media. Using fake news the more criminal activity are happing in the world it causes defect on hu- man life to avoid and stop criminal activities by using these techniques Our work considers crowd signals for detecting fake news and is motivated by tools recently introduced by Twitter that enable users to flag fake news. Since emotion plays an indispensable role in Twitter, I may expect a better strategy for constructing the user profile when I take users emotions into consideration. Combining emotion and user modelling is not a totally new idea. Some previous studies already focused on this combination and they proved that the combination of emotion and user model- ing could improve the quality of user profile. However, most of these studies have a user-interaction part to collect the emotive response from users and no one has com- bined emotion and user modeling on the Twitter platform. The main goal behind the study of this thesis is to analyze the emotion features in Twitter and add these users emotion features in user modeling strategies.Social media has become a dominant networking channel and vent for people to express their thoughts and feelings in the form of reviews, tweets, posts, discus- sions, etc. The wide spread use of social media services has also led to a wide spread availability of opinionated and emotional posts. The exponential growth of subject-Novel Technique For Sentiment Analysis Using Text And Facial Recognitiontive and emotional textual information in digital, unstructured format on resources publicly available on Twitter means that it is paramount to find ways of extracting valuable information from it. Emotion analysis on tweets concern the application of automatic methods for predicting the sentiment content and emotional state of a tweet, respectively. Even with huge capabilities of computation, understanding the sentiments and emotions embodied in the analysed text has remained as one of the challenging tasks on Twitter. Hence, this motivates to propose emotion prediction, using tweets dataset. [2]. Majority of previous studies predict either Ekman's or Plutchiks's classifications, while POMS's adjectives had only been used in simple keyword spotting algorithms [1].We are not aware of any studies that tackle the problem of predicting POMS's categories from the text. Methodologically, they mainly used simple classification algorithms, like logistic regression or support vector machines, on top of word and n-gram counts, and other custom-engineered features (capturing the use of punctuation, the presence or absence of negation, and counts of words from various emotion lexicons).

Deep learning has recently shown a lot of success in the field of NLP: it has been used for improving sentiment analysis, opinion mining and other tasks like part of- speech tagging, chunking, named entity recognition, and semantic role labelling. Radford et. al even reported on discovering the sentiment unit while training to generate reviews. Despite these successes, deep learning approaches for emotion recognition are almost non-existent.

Previous work generally studied only one emotion classification. Working with multiple classifications simultaneously not only enables performance comparisons between different emotion categorizations on the same type of data, but also allows us to develop a single model for predicting multiple classifications at the same time. The next section describes three psychological classifications of emotions. The third section outlines the data collection and filtering. To obtain an emotion-labelled training data set, we exploited the presence of hashtags, a self-annotation mechanism frequently used on Twitter. In particular, we used the hashtags that, according to the three classifications, directly correspond to particular emotions. We further filtered the tweets with heuristics aimed at increasing the chances that the hashtag found in the tweet actually reflects the author's emotions.

The fourth section focuses on algorithms. We start with commonly used classifiers on bag-of-words and latent semantic indexing models that serve as the base- line that we try to surpass with deep learning. We use recurrent and convolutional neural network models, into which we feed either words or individual characters. We describe how to transfer models trained with one classification of emotions to another classification. Finally, we show a novel algorithm for fitting a common net- work for the three classifications with different complexity and very different data set sizes. The fifth section shows empirical results and debate on ideas for future work.

II.LITERATURE REVIEW

"Emotional states of individuals, also known as moods, are central to the `- sion of thoughts, ideas and opinions, and in turn impact attitudes and behavior". In this paper we have proposed a method which detects the emotion or mood of the tweet and classify the twitter message under appropriate emotional category. Our approach is a two-step approach, it is so called as it uses two approaches for the clas- sification process, one is Rule Based approach and the other is Machine Learning ap- proach. The first approach is the Rule Based Approach (RBA), our minor contribu- tions in this approach are pre-processing, tagging, feature selection and Knowledge base creation. Feature selection is based on tags. Our second approach is Machine Learning Approach (MLA), in this the classifier is based on supervised machine learning algorithm called Na¨ıve Bayes which requires labeled data. Na¨ıve Bayes is used to detect and classify the emotion of a tweet. The output of RBA is given to MLA as input because MLA requires labeled data which we have already created through RBA. We have compared the accuracies of both the approaches, observed that, with the rule based approach we are able to classify the tweets with accuracy around 85% and with the machine learning approach the accuracy is around 88%. Machine learning approach performance is better than rule based approach, the per- formance has been improved as we have removed the error data while training the model. The approaches are involved with the concepts of Natural Language Process- ing, Artificial Intelligence, and Machine Learning for the development of the system. Our major contributions in this paper are detection of emotion for non hashtagged data and the labeled data creation for machine learning approach without manual creation. [1]

Emotional states of individuals, also known as moods, are central to the ex- pression of thoughts, ideas and opinions, and in turn impact attitudes and behavior. Social media tools like twitter is increasingly used by individuals to broadcast their day-to-day happenings or to report on an external event of interest, understanding the rich 'landscape' of moods will help us better to interpret millions of individuals. This paper describes a Rule Based approach, which detects the emotion or mood of the tweet and classifies the twitter message under appropriate emotional category. The accuracy with the system is 85%. With the proposed system it is possible to understand the deeper levels of emotions i.e., finer grained instead of sentiment i.e., coarse grained. Sentiment says whether the tweet is positive or negative but the pro- posed system gives the deeper information of tweet which has adverse uses in the field of Psychology, Intelligence Bureau, Social and Economic trends. [2]

We examine if common machine learning techniques known to perform well in coarsegrained emotion and sentiment classification can also be applied successfully on a set of fine-grained emotion categories. We first describe the grounded theory approach used to develop a corpus of 5,553 tweets manually annotated with 28 emo- tion categories. From our preliminary experiments, we have identified two machine learning algorithms that perform well in this emotion classification task and demon- strated that it is feasible to train classifiers to detect 28 emotion categories without a huge drop in performance compared to coarser-grained classification schemes.

Automatic fine-grained emotion detection is a challenging task but we have demonstrated that it is feasible to train a classifier to perform decently well in clas- sifying as many as 28 emotion categories. Our 28 emotion categories is an exten- sion to the six to eight emotion categories commonly-used in the state-of-the-art (Alm et al., 2005; Aman & Szpakowicz, 2007; Mohammad, 2012). Some of the 28 emotion categories overlap with those found in existing emotion theories such as Plutchik's (1962) 24 categories on the wheel of emotion and Shaver et al.'s (2001) tree-structured list of emotions. Existing emotion theories in psychology are not de- veloped specifically based on emotions expressed in text. Therefore, our emotion categories offer a more fitting framework for the study of emotion in text. [3]

Social media and microblog tools are increasingly used by individuals to express their feelings and opinions in the form of short text messages. Detecting emotions in text has a wide range of applications including identifying anxiety or depression of individuals and measuring well-being or public mood of a community. In this paper, we propose a new approach for automatically classifying text messages of individuals to infer their emotional states. To model emotional states, we utilize the well-established Circumplex model that characterizes effective experience along two dimensions: valence and arousal. We select Twitter messages as input data set, as they provide a very large, diverse and freely avail- able ensemble of emo- tions. Using hash-tags as labels, our methodology trains supervised classifiers to detect multiple classes of emotion on potentially huge data sets with no manual ef- fort. We investigate the utility of several features for emotion detection, including unigrams, emoticons, negations and punctuation. To tackle the problem of sparse and high dimensional feature vectors of messages, we utilize a lexicon of emotions. We have compared the accuracy of several machine learning algorithms, including SVM, KNN, Decision Tree, and Naive Bayes for classifying Twitter messages. Our technique has an accuracy of over 90%, while demonstrating robustness across learn- ing algorithms. [4]

Authors explore the properties of byte-level recurrent language models. When given sufficient amounts of capacity, training data, and compute time, the repre- sentations learned by these models include disentangled features corresponding to high-level concepts. Specifically, it find a single unit which performs sentiment analysis. These representations, learned in an unsupervised manner, achieve state of the art on the binary subset of the Stanford Sentiment Treebank. They are also very data efficient. When using only a handful of labeled examples, this approach matches the performance of strong baselines trained on full datasets. Authors also demonstrate the sentiment unit has a direct influence on the generative process of the model. Simply fixing its value to be positive or negative generates samples with the corresponding positive or negative sentiment. [5]

Discourse Parsing and Sentiment Analysis are two fundamental tasks in Nat- ural Language Processing that haqve been shown to be mutually beneficial. In this work, authors design and compare two Neural models for jointly learning both tasks. In this approach, authors first create a vector representation for all the text segments in the input sentence. Next, it apply three different Recursive Neural Net models: one for discourse structure prediction, one for discourse relation prediction and one for sentiment analysis. Finally, authors combine these Neural Nets in two different joint models: Multi-tasking and

Pre-training. The results on two standard corpora indicate that both methods result in improvements in each task but Multi-tasking has a bigger impact than Pre-training. Specifically for Discourse Parsing, shows im- provements in the prediction on the set of contrastive relations. [6]

In this paper, authors go one step further and develop a novel method for senti- ment learning in the MapReduce framework. their algorithm exploits the hash tags and emoticons inside a tweet, as sentiment labels, and proceeds to a classification procedure of diverse sentiment types in a parallel and distributed manner. Moreover, it utilize Bloom filters to compact the storage size of intermediate data and boost the performance of the algorithm. Through an extensive experimental evaluation, It prove that this solution is efficient, robust and scalable and confirm the quality of sentiment identification. [7]

III. RESEARCH AND METHODOLOGY

With the development of networks, social platforms play an indispensable role in people's daily lives. As the most popular micro blogging platform, Twitter has a vast amount of information available in the form of tweets shared by millions of users. Since this data stream is constantly growing, it is difficult to extract relevant information for users. More and more people want to benefit from these data and get a personalized service from Twitter. Extracting the semantic meaning of Twitter and modeling the interests of users allows people to enjoy a personalized service on Twitter. Meanwhile, research shows that people tend to express their emotions on Twitter. These emotional tweets usually clearly express the users preferences compared with other normal tweets. [4] [7] With the development of networks, social platforms play an indispensable role in people's daily lives. As the most popular micro blogging platform, Twitter has a vast amount of information available in the form of tweets shared by millions of users. Since this data stream is constantly growing, it is difficult to extract relevant information for users. More and more people want to benefit from these data and get a personalized service from Twitter. Extracting the semantic meaning of Twitter and modeling the interests of users allows people to enjoy a personalized service on Twitter. Meanwhile, research shows that people tend to express their emotions on Twitter. These emotional tweets usually clearly express the users preferences compared with other normal tweets. [3].

The role of social media in our day to day life has increased rapidly in recent years. It is now used not only for social interaction, but also as an important platform for exchanging information and news. Twitter, Facebook a micro-blogging service, connects millions of users around the world and allows for the real-time propagation of information and news. Fake news has become a major problem in these social networks. Fake news has become a major problem in these social networks Fake news has vast impact in our modern society. Detecting Fake news is an important step. This work purposes the use of machine learning techniques to detect Fake news, using Support Vector Machine (SVM) algorithm. The normalization method is important step for cleansing data before using the machine learning method to classify data. This model count the credibility of content and user reputation. This method develops a method for automating fake news detection on social media by learning to predict accuracy assessments in credibility-focused Twitter dataset.

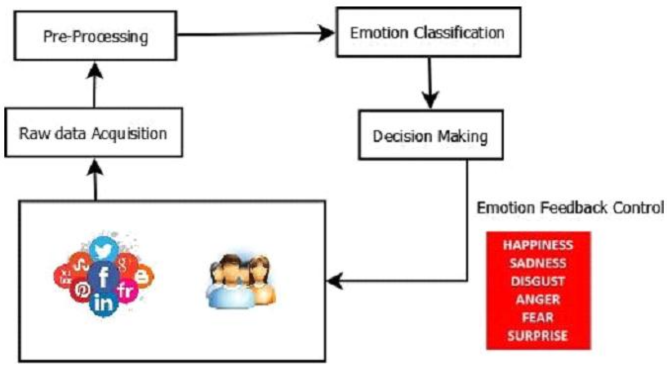

Fig 1. Architecture Diagram

Therefore, the goal of this work is to design some emotion-based user modeling strategies which exploit these emotional data. This work introduce and analyze the approaches for detecting emotion on Twitter. First it evaluate and compare the per- formance of proposed approaches of emotion detection. Then use these approaches of emotion detection to analyze Twitter sample dataset for the purpose of user mod- eling. Also proposed set of emotion-based user modeling strategies on the Twitter platform based on these detected emotional data. Furthermore, It evaluate emotion- based user modeling strategies and investigate their impacts on normal user profiles. Proposed system results show that emotion-based user profiles enhance the quality of user profiles and have a better performance. [1]

Profile of Mood States [6] is a psychological instrument for assessing the individ- ual's mood state. It defines 65 adjectives that are rated by the subject on the five-point scale. Each adjective contributes to one of the six categories. For example, feeling annoyed will positively contribute to the anger category. The higher the score for the adjective, the more it contributes to the overall score for its category, except for relaxed and efficient whose contributions to their respective categories are nega- tive. POMS combines these ratings into a six-dimensional mood state representation consisting of categories: anger, depression, fatigue, vigour, tension and confusion. Since POMS is not publicly available, we used the structure from Norcross et al. [7], which is known to closely match POMS's categories. We supplemented it with additional information from the BrianMac Sports Coach website1. Comparing to the original structure, we discarded the adjective blue, since it only rarely corresponds to an emotion and not a color, and word-sense disambiguation tools were unsuccessful at distinguishing between the two meanings. We also removed adjectives relaxed and efficient, which have negative contributions, since the tweets containing them would represent counter-examples for their corresponding category.

For each category, we used the following adjectives:

•anger: angry, peeved, grouchy, spiteful, annoyed, resentful, bitter, ready to fight, deceived, furious, bad tempered, rebellious,

•depression: sorry for things done, unworthy, guilty, worthless, desperate, hope- less, helpless, lonely, terrified, discouraged, miserable, gloomy, sad, unhappy,

•fatigue: fatigued, exhausted, bushed, sluggish, worn out, weary, listless,

•vigour: active, energetic, full of pep, lively, vigorous, cheerful, carefree, alert,

•tension: tense, panicky, anxious, shaky, on edge, uneasy, restless, nervous,

•confusion: forgetful, unable to concentrate, muddled, confused, bewildered, uncertain about things.

IV.CONCLUSION

The main idea of the paper is to expand the leverage of deep learning and data mining techniques for emotion recognize in tweeter. Proposed system is directly connected to the tweeter's public account to get the tweets stream. Experiment shows the result in three algorithms name as SVM, Random Forest and CNN. In experiments needs to tokenize the tweets and get semantic of the token to get better result. Accuracy of the proposed system using semantic with preprocessing is much better than existing system. Proposed system is taken word as input instead of char- acters.

This proposed System works on probably the largest data set for emotion pre- diction, using tweets from years. With the aim of developing a universal emotion detection algorithm, I did not restrict ourselves only to one domain, but rather tested its usefulness for different classifications of emotions. Since the training data was annotated automatically and since I use character-based approaches, our solution is language independent and could easily be adapted for other languages. I believe this work is beneficial for the user modeling on the Twitter platform and seeks to com- bine two hotspots, the emotion and user modeling.

In this project identifying misinformation is authoritative in online social media platforms, because information is circulated easily across the online community by unsupported sources. To be able to automatically detect fake news and stop misin- formation circulation,can be useful in analyzing activist movements. User accounts who include many URLs, @username mentions and hashtags in their tweet.This pa- per also solve the problem of assessing information credibility on Twitter. The issue of information credibility has come under scrutiny, especially in social networks that are now being used actively as first sources of information.Future scope of the system is to recognize the emotion based on uploaded multi-media (i.e. images) by user.

REFERENCES

1.Niko Colneric and Janez Demsar, "Emotion Recognition on Twitter: Com- parative Study and Training a Unison Model", IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, FEBRUARY 2018.

2.A. Radford, R. Jozefowicz, and I. Sutskever, "Learning to Generate Reviews and Discovering Sentiment", 2017.

3.B. Nejat, G. Carenini, and R. Ng, "Exploring Joint Neural Model for Sen- tence Level Discourse Parsing and Sentiment Analysis", Proc. of the SIG- DIAL 2017.

4.N. Nodarakis, S. Sioutas, A. Tsakalidis, and G. Tzimas, "Using Hadoop for Large Scale Analysis on Twitter: A Technical Report", arXiv preprint arXiv: 1602.01248, 2016.

5.Y. Zhang and B. C. Wallace, "A Sensitivity Analysis of (and Practitioners' Guide to) Convolutional Neural Networks for Sentence Classification", ArXiv preprint arXiv:1510.03820v4, 2016.

6.J. Guo, W. Che, H. Wang, and T. Liu, "A Universal Framework for Inductive Transfer Parsing across Multi-typed Treebanks", Proc. of the 26th Int. Conf. on Computational Linguistics (COLING-16), pp. 12–22, 2016.

7.S. M. Mohammad and S. Kiritchenko, "Using Hashtags to Capture Fine Emo- tion Categories from Tweets", Computational Intelligence, vol. 31, no. 2, pp. 301–326, 2015.