



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: IV Month of publication: April 2021

DOI: <https://doi.org/10.22214/ijraset.2021.33584>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Classification Approach and Analysis for Predicting Social Emotion on User Review

Bhushan R. Chincholkar¹, J. M. Patil²

¹PG Scholar, ²Assistant Professor, Computer Engineering, Shree Sant Gajanan Maharaj College of Engineering, Shegaon, India

Abstract: Classification of opinions or sentiments is the core task in opinion mining. To accomplish this task, often Bag-Of-Words (BOW) is used as a feature for training a classifier in statistical machine learning. (pp3) BOW (Bag- of-Words) is the most common machine learning method used for sentiment classification. However, this model does not address the problem of polarity shift due to fundamental limitations. This is one of the reasons which affect its overall accuracy. Based on the review of current prediction algorithms of network security situation, prediction algorithms. This paper discusses a method using dual training and dual prediction for sentiment classification while addressing the problem of polarity shift and also evaluates it. We propose a modified fuzzy approach with two stage training for dealing with text ambiguity and classifying types of hate positive, negative and neutral speech and compare its performance with those popular methods as well as some existing fuzzy approaches, while the features are prepared through the bag-of-words and word embedding feature extraction methods alongside the correlation based feature subset selection method. The experimental results show that the proposed fuzzy method outperforms the other methods in most cases. It extends the given framework from polarity (positive-negative) classification to 3- class (positive negative- neutral) classification, by taking the all reviews into consideration.

I. INTRODUCTION

The traditional technique used in sentiment classification is of BOW (Bag-of-Words) model. This is most commonly used for document sentiment classification. In bag-of-words model, the sentences are represented by bag of words keeping multiplicity of words but ignoring grammar and word order. After this, statistical machine learning algorithms such as Naive Bayes algorithm are used to train the classifier on the basis of features like frequency of words, opinion words and opinion phrases. This model is very simple and effective but it distorts word order and it also does not consider linguistic knowledge for sentiment classification. Polarity shift is a major problem that this model does not address which affects the overall classification accuracy of this approach [1]. Sentiment Analysis (SA) task is to label people's opinions as different categories such as positive and negative from a given piece of text. Another task is to decide whether a given text is subjective, expressing the writer's opinions, or objective, expressing. These tasks were performed at different levels of analysis ranging from the document level, to the sentence and phrase level [8]. Websites, Forums, blogs, social networks, and content sharing services help people to share their experiences, knowledge and opinions. Capturing public opinion about social events, political movements, company strategies, marketing campaigns, and product preferences is garnering increasing interest from the scientific community (for the exciting open challenges), and from the business world (for the remarkable marketing fallouts and for possible financial market prediction). The resulting emerging fields are opinion mining (OM) or sentiment analysis (SA) [8]. A good range of research has been done on sentimental analysis to find how we can improve the data analysis for getting more valuable business information to improve service. Sentiment analysis is a way to getting structured information from unstructured data. Informal language & emotions are commonly used in user made data in social sites. Sometimes it follows the specific pattern to understand sometimes it does not. To get the information it needs to train system by using lexicon or machine learning technology [2]. In sentiment detection, word/words expressing sentiment or consumer opinions are separated from available text. Then we study the polarity of these words (positive or negative). One of the main difficulties in accurate opinion classification is the polarity shift problem. Polarity shift is a phenomenon in natural language processing which reverses the real or intended opinion of the reviewer about a particular subject [3]. In process system analysis, we focus on cyber-hate classification, since the spread of hate speech using social media can have disruptive impacts on social cohesion and lead to regional and community tensions. Automatic detection of cyber-hate has become a priority research area. In particular, we propose a modified fuzzy approach with two stage training for dealing with text ambiguity and classifying types of hate positive, negative and neutral speech and compare its performance with those popular methods as well as some existing fuzzy approaches, while the features are prepared through the bag-of-words and word embedding feature extraction methods alongside the correlation based feature subset selection method.

The experimental results show that the proposed fuzzy method outperforms the other methods in most cases. The rest of the paper is organized as follows: Section I Introduction. Section II discusses Background and Related Work. Section III discusses existing methodologies. Section IV discusses proposed method. Finally section V Conclude this review paper.

II. BACKGROUND AND RELATED WORK

The computational study of opinions of people towards entities and their attributes is defined as Sentiment Analysis. A product, person, service, topic or an individual can be referred to as an entity whereas an opinion is a positive, negative or neutral point of view towards an entity. Sentiment Analysis can be of many types – document based, sentence based, phrase based, and aspect based sentiment analysis.

In the paper [1], a method is discussed which uses dual training and dual prediction for sentiment classification while addressing polarity shift. The terms dual training and dual prediction refer to usage of both original review sample and opposite review sample for training and prediction.

In the research [2], represented noun as sentimental words and which has a good impact on sentiment detection additionally some words has duel sentiment base on its application, mostly those words are a noun. It introduced the possible solution for these issues and experiment on those.

Finally, we maximize the accuracy by 3% by applying our offered solution. A dual prediction algorithm [3] is used for classification of test reviews by considering both positive and negative sides of each review. At the end we build a pseudo-opposites dictionary using a corpus based method.

By this we tackle the problem of having to depend upon an external opposites dictionary for opposites of reviews. By doing this we also get a domain adaptive dictionary for training a classifier which increases the accuracy of the dual opinion mining model. Based on the review of current prediction algorithms of network security situation, prediction algorithms based on Kalman filtering are studied [5].

A prediction algorithm of network security situation based on grey correlation entropy Kalman filtering is presented, hoping to be more helpful to network administrators through providing them information more effectively. We propose a dual training algorithm to make use of original and reversed training reviews in pairs for learning a sentiment classifier, and a dual prediction algorithm to classify the test reviews by considering two sides of one review. We also extend the DSA framework from polarity (positive-negative) classification to 3-class (positive negative- neutral) classification, by taking the neutral reviews into consideration. Finally, we develop a corpus-based method to construct a pseudo-antonym dictionary, which removes DSA's dependency on an external antonym dictionary for review reversion.

III. EXISTING METHODOLOGIES

A. Machine Learning

Train up the machining is the primary requirement for sentiment analysis. Machine learning is most popular technology besides this lexicon is also use for sentiment analysis. Without any of that sentiment analysis is not possible. There are many algorithm and approaches for machine learning. Some are supervised where some are unsupervised. Machine learning methods and term-counting methods are generally used for document-level opinion classification as well as sentence level opinion classification [2].

B. Opinion Mining and Polarity Shift

The opinion orientation is calculated by summation of orientation scores, based on manually collected or lexical resources. In the machine learning methods, opinion classification is considered to be a statistical problem. A structure in which sentence is broken into its words and stored which resembles a bag of words is used to store opiniated text. The previously trained machine learning algorithms are applied as classifiers. However, these traditional models resembling bag of words prove inefficient in dealing with polarity shifting of the text [3].

C. Building Predictive Model with Naive Bayes

The Naive Bayes algorithm is a classification technique based on the Bayes Theorem with an assumption of independence among predictors. This stage aims to classify the reviews into categories that help or do not help properly. The data used to build the model is labelled data with two categories: 280 reviews with help categories and 50 reviews with categories not helpful. Prediction is done by using probabilistic methods with Naive Bayes algorithm to classify each of the reviews into one category [6].

D. Task Scheduling Prediction Algorithms

At present, most procedures of signal processing present a certain periodicity, such as the frame-based audio video signal processing, macro block-based image processing. In the periodic signal processing, the processing procedures in each cycle are basically the same, or they have the same task scheduling. So, the task scheduling prediction can be implemented through period detection. However, the period detection here has the following characteristics, which are different from the period detection of time series data mining [7].

E. Lexicon-based Approach

The lexicon-based approach depends on finding the opinion lexicon which is used to analyse the text. There are two methods in this approach. The dictionary-based approach which depends on finding opinion seed words, and then searches the dictionary of their synonyms and antonyms. The corpus-based approach begins with a seed list of opinion words, and then finds other opinion words in a large corpus to help in finding opinion words with context specific orientations. This could be done by using statistical or semantic methods [8].

IV. PROPOSED METHODOLOGY

In proposed system there are basically two modules are available .First module indicate text analysis framework in detail. And second module indicate the prediction user review rating on user previously review data.In this system, we focus on the detection of online hate speech (cyberhate) in short informal text posted to social media platforms. This has become a priority research topic due to the concern that the spread of online hate speech could lead to antisocial outcomes. In particular, we deal with four types of online hate speech, namely, religion, race, disability, and sexual orientation, by proposing a novel fuzzy approach grounded in generative learning, especially for dealing with text ambiguity, which could result from the following cases: 1) the same word may be used in different contexts leading to different semantic meanings and 2) that similar instances are assigned different labels by different annotators due to their different opinions. The proposed fuzzy approach is different from existing fuzzy systems in two aspects.

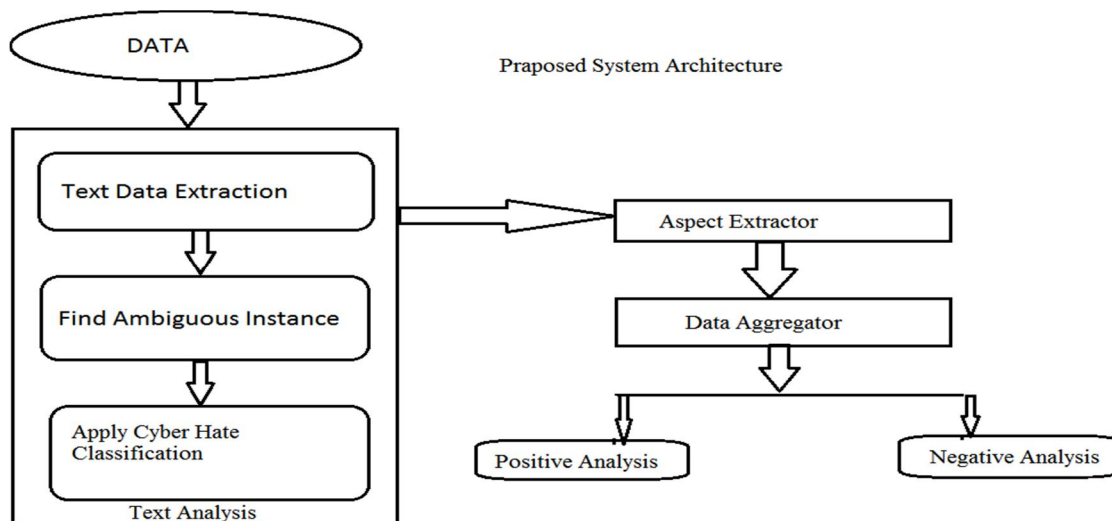


Fig 2: Proposed System Architecture

V. CONCLUSION

Sentiment Analysis is very important research because Sentiment Analysis help in summarizing opinion and reviews of public. They consider as research filed. However, Sentiment Analysis still needs to improve and progress. Moreover, there are many challenges like the polarity in a complex sentence. In this system, we proposed a modified fuzzy approach for cyber hate classification. In particular, we argued that fuzzy approaches are more suitable than previously used non-fuzzy approaches that are known to perform well on hate speech data, due to the advantages of fuzzy approaches in dealing with fuzziness, imprecision, and uncertainty of text. We conducted experiments using four data sets on four types of hate speech, namely, religion, race, disability, and sexual orientation.



REFERENCE

- [1] Ravinder Ahuja, Willson Anand," Sentiment Classification of Movie Reviews Using Dual Training and Dual Perdition", Fourth International Conference on Image Information Processing (ICIIP),2017.
- [2] Prosanta Kumar Chaki, Ikrum Hossain, Probhas Ranjan Chanda, Shikha Anirban," An Aspect of Sentiment Analysis: Sentimental Noun with Dual Sentimental Words Analysis", International Conference on Current Trends in Computer, Electrical, Electronics and Communication (ICCTCEEC-2017).
- [3] Mrs. Mukta Y. Raut, Mrs. Mayura A. Kulkarni," Polarity Shift in Opinion Mining", IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT) Rajarshi Shahu College of Engineering, Pune India. Dec 2-3, 2016.
- [4] Jianghao Lin, Yeli Gu, Yongmei Zhou, Aimin Yang, Jin Chen and Xinguang Li," Combining Convolutional Neural Networks and Word Topic Features for Chinese Short Text Sentiment Analysis", IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), Nov. 2018.
- [5] Zhang Lin ; Liu Xiujie ; Ma Jing ; Sun Wenchang ; Wang Xiufang," The prediction algorithm of network security situation based on grey correlation entropy Kalman filtering", IEEE 7th Joint International Information Technology and Artificial Intelligence Conference, Dec 2014.
- [6] Fetty Fitriyanti Lubis, Yusep Rosmansyah, Suhono Harso Supangkat," Improving Course Review Helpfulness Prediction through Sentiment Analysis", The International Conference on ICT for Smart Society (ICISS),2017.
- [7] Quan Haojun, Zhang Tao, Guo Jichang," Task Scheduling Prediction Algorithms for Dynamic Hardware/Software Partitioning", Fifth International Symposium on Parallel Architectures, Algorithms and Programming,2012.
- [8] Zohreh Madhoushi, Abdul Razak Hamdan, Suhaila Zainudin," Sentiment Analysis Techniques in Recent Works", Science and Information Conference, July 28-30, 2015
- [9] Daniel Ansari," Sentiment Polarity Classification using Structural Features", IEEE 15th International Conference on Data Mining Workshops, 2015.
- [10] Han Liu, Pete Burnap, Wafa Alorainy, Matthew L. Williams" A Fuzzy Approach to Text Classification With Two-Stage Training for Ambiguous Instances", IEEE TRANSACTIONS ON COMPUTATIONAL SOCIAL SYSTEMS, 2019.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)