A Survey on Online Review Analysis based on Sentiment Measurement

Reena Jose[1], Angitha Thomas[2], Geethumol M S[3], Nidhiya Sany[4]
[1]Assistant Professor, Dept of Computer Science
[2, 3, 4]Dept of Computer Science
[1, 2, 3, 4]MG UNIVERSITY, Depaul Institute of Science and Technology, Angamaly, India

Abstract- *Nowadays customers express their opinion about the product they consume through the reviews and ratings. The paper emphasis on product (textile) review analysis using sentiment measurement technique. Sentiment measurement in the sense to analyze the natural language text reviews by determining sentimental terms and finds whether it is positive, negative or neutral based on domain ontology. A sentiment score is generated and which can represent the emotional intensity of customers on a product. The reviews influence the customers and the enterprises over a product. This model can achieve high accuracy and provide decision support.*

Keywords- Sentiment Analysis, Lexicon-Base Approach, Domain Ontology, Product Reviews, Emotional Intensity.

I. INTRODUCTION

Reviews play a crucial role in online product consumptions. The consumers refer the reviews before consuming a product to know the product quality. The enterprises are eager to get reviews for analyzing it and market the products efficiently. The reviews are very valuable for the product marketing scale and decision supporting. The proposed model calculates the emotional intensity of the product review. As the reviews for the product grows widely it is difficult to evaluate and understand it. Therefore the scope of sentiment analysis in every sector is increasing widely. The existing sentiment analysis models are Lexicon based model, word alignment model (unsupervised) and Word alignment model (supervised). Sentiment measurement can be done in word level, sentence level and document level.

Sentiment analysis is done by analyzing the reviews, extracting the sentiment terms from the reviews and classify the terms as positive, negative or neutral categories based on sentiment dictionary data set. Now the data will be transformed in the form of scores or values. Therefore we can easily calculate the emotional intensity of the product review and represent it as patterns for better understanding. As the results it may help the customers and enterprise for decision making. The emotional analysis becoming a wide area of research and there are a lot of literatures available. The proposed system works based on a Domain ontology based on dictionaries and rules. Initially construct a domain ontology, negative dictionary and adverb dictionary. Collect the raw data, then extract the words by pattern matching with the domain ontology and dictionaries. Here emotional featured words are obtained and then analyzed based on rules, therefore gets emotional scores. Finally an average of these different emotional scores will be calculated and represent it using patterns like graphs and charts with data points.

A. Factors Effecting Sentiment Analysis:

The below shown [2] are some factors that affect the sentiment analysis:

i) Analysis level

In sentence level analysis words and emoticons are not taken into consideration. In document level all words and emoticons are taken into consideration for the analysis.

ii) Domain Dependency

The semantics of words and emoticon may change according to different domains.

iii) Intensifiers

Amplifiers are the exclamations and dots used in the text. It may add more sentiments to words associated with them.

iv) Count of Emoticons

By using emoticons people can express their feelings in a wide way.

v) Emoticon Position

Nowadays emoticons used anywhere in a sentence while so the semantics will be different.

vi) Classification

The classification of data into different categories like positive, negative and neutral. Some researchers have also classified their works into sub categories

Sentiment analysis is had achieved a lot but, still there are some imperfections. The paper constructs a domain ontology of online products from multiple perspective such as customers and enterprises. It provide decision support and provide more insight for consumers and enterprises about the product from the customer reviews.

II. RELATED WORKS

Development of web2.0 made internet users generate their opinions through different web applications. Sentiment analysis is also known as opinion mining. The important tasks of sentiment analysis are sentiment extraction, sentiment classification, sentiment retrieval and summarization [2]. This paper introduces a SSAM which satisfy these tasks. Here an LMR template and Maximum Entropy are combined to identify the sentient terms. Domain lexicon method is used for sentiment related element extraction. Maximum Entropy algorithm is proposed for classifying sentences. The text sentiments are calculated by text orientation. And finally sentiment text retrieval is presented.

In [3] it depicts different classifications of sentiment approach and tools used for sentiment analysis. The sentiment classifications included are the machine learning, lexicon based and hybrid approach. To predict the polarity of emotions based on the trained as well as test data sets is machine learning approach. The lexicon based approach uses a predefined list of words where each word is associated with sentiments and raw data is analyzed. In hybrid approach the combination of both machine learning and lexicon based approach is used for the improvement of sentiment analysis. The major tools included to detect the emotional polarity are Emoticons, LIWC, SentiStrengh, Senti WordNet, SenticNet, Happiness Index, AFINN, PANAS-t, Sentiment140, NRC, EWGA and FRN.The main task of sentiment analysis is to convert unstructured text into meaningful information.

The traditional methods used texts to detect sentiment polarity and assume that microblogs are independent and ignored microblogs are networked data. Therefore the traditional methods and performance were not satisfactory. In [4] it showcased a method of combining social context and topic context for microblog sentiment analysis. Here it uses a method to measure structure similarity by using structure similarity into social contexts. The semantic relations between the microblogs are modeled using topic context. Then Laplacian matrix of the graph is built by combining the social context and topic context. And also added Laplacian regularization into this sentiment analysis model. Adding topic context and social context were helpful for improving the sentiment classification accuracy.

A novel sentiment classification algorithm is included in [5] to detect the sentiment polarity. The word emotion is dynamic tendency in Chinese.The different emotional words, the field, targets or object will be misclassified, therefore common emotional dictionaries were replaced and here they included different set of dictionaries and they are Absolute emotional dictionary (AbED), the relative emotional dictionary (ReED), the field of emotional dictionary (FiED) and the field of targets and opinion words dictionary (TfED). Multi-string pattern matching algorithm is used as an effective algorithm to discriminate sentiment polarity with those emotional dictionary, negative dictionary and modified dictionary. It has made a higher precision and recall rate.

Sentiment classifications for online product reviews[6] are discussed using a set of experiments which uses different machine learning algorithms. Here it is much simplified and classified the reviews into positive and negative classes. Sentiment classification is depicted by, (i) represent the entire data set in a summarized manner, like a label or score, and (ii) the data is fragmented into two classes. Here included a study on multiple sentiment classification algorithms such as Winnow, a generative model based on language modeling and a discriminative classifier that employs projection based learning. Experimental comparison of a discriminative model and generative model showed that these can achieve 75%-80% of precision and recall on polarity prediction. While winnow model only achieved 43% for positive and 16% for negative reviews of recall. Here the major classifiers used were Passive-Aggressive (PA) Algorithm Based Classifier, Language Modeling (LM) Based Classifier, Winnow Classifier. It showcased a experiments on large set of data. Discriminating classifiers combined with high order n-grams as feature provided better performance in the experimental result.

To obtain the sentiments of the user from the reviews, the sentences are to be reduced into words in which the adjectives, verbs and adverbs are processed using Parts Of Speech Tagging (POST) [7]. The negative expressions are detected and processed using Negation Phrase Identification algorithm. Here the sentences are classified into negative or positive. The quality of online products are analyzed using the positive and negative separation of comments.

Sentiment analysis is the study area in Natural language processing (NLP) [8], emphasis on sentiment analysis using different methods of natural language processing. It process using different steps: (i) data collection with labelling, (ii) data mining and (iii) sentiment analysis. The reviews are analyzed based on text based and emoticon based approaches.

For the analysis of results Weka data mining tool is used. By using Weka to measure the effectiveness of data, different classifiers are used, Such as Support vector machine (SVM), Naïve byes (NB) and Logistic Regression (LR). As a result of this study, SVM provides maximum accuracy and efficient outcomes.

III. CONCLUSION

As the opinions and comments to online products increase, the sentiment analysis method provide a way to measure the reviews based on emotions, attitudes of people to the product.This paper shows the available methods of sentimental analysis and proposed a semi supervised approach to measure the sentiment score based on Domain ontology. Here the reviews are tagged positive or negative based on the Domain ontology, negative dictionary and adverb dictionary. As though sentiment analysis have limitations but it can provide high accuracy on review evaluations. The proposed system provide decision-support for the customers and enterpriser can use the result for the product sales or improvement of product.

REFERENCES

[1] Y. Du and L. Yang, *"A Sentiment Measurement Model for Online Reviews of Complex Products",* 2019 International Conference on Communications, Information System and Computer Engineering (CISCE)*,* Haikou, China, 2019,pp. 199-202.

[2] Ms. Payal Yadav and Prof. Dhatri Pandya, "*SentiReview: Sentiment Analysis based on Text and Emoticons* ", Research Gate, Conference Paper, February 2017

[3] Hui He and Bo Chen, "*A Smart Sentiment Analysis System in Word,Sentence and Text Level*", Advanced Engineering Forum, Trans TechPublications, 2011, 1, pp.338-342.

[4] Alessia D'Andrea, Fernando Ferri, Patrizia Grifoni and Tiziana Guzzo "*Approaches, Tools and Applications for Sentiment Analysis Implementation*", International Journal of Computer Applications Volume 125 – No.3, September 2015.

[5] Xiaomei Zou, Jing Yang and Jianpei Zhang,"*Microblog sentiment analysis using social and topic context*.", PLoS ONE, February 2, 2018.

[6] Jiao J , Zhou Y , "*Sentiment Polarity Analysis based multi-dictionary*",Physics Procedia, 2011, 22, pp.590-596.

[7] Hang Cui, Vibhu Mittal and Mayur Datar," Comparative Experiments on Sentiment Classification for Online Product Reviews",Google Research,Proceedings of the 21st National Conference on Artificial Intelligence, AAAI, Boston, MA (2006).

[8] Raheesa Safrin, K.R.Sharmila, T.S.Shri Subangi and E.A.Vimal,"*SENTIMENT ANALYSIS ON ONLINE PRODUCT REVIEW*",International Research Journal of Engineering and Technology (IRJET),Vol: 04, Apr -2017.

[9] Hang Cui, Vibhu Mittal and Mayur Datar, " *Comparative Experiments on Sentiment Classification for Online Product Reviews*", Google Research,Proceedings of the 21st National Conference on Artificial Intelligence, AAAI, Boston, MA (2006).