

A SURVEY ON PRODUCT RATING PREDICTION USING DATA MINING FOR POLARITY CHECK

^{#1}Sonwalkar Komal, ^{#2}Shaikh Sabiya, ^{#3}Shinde Rinku, ^{#4}Gaikwad Priti



¹komal.sonwalkar@gmail.com

²shaikhsabiya5@gmail.com

³rinkushinde1318@gmail.com

⁴priti.gaikwad148@gmail.com

^{#1234}Department of Computer

JSPM's Imperial College of Engineering and Research,
Wagholi, Pune University, Pune.

ABSTRACT

BOW stands by Bag-of-words is the most popular way to model text in statistical machine learning approaches in sentiment analysis. However, the performance of BOW sometimes remains limited due to some fundamental deficiencies in handling the polarity shift problem. We propose a model called dual sentiment analysis (DSA), to address this problem for sentiment classification. Data mining technique is the process of turning raw data into useful information. The main use of data mining is to fetch the required data and extract useful information from the data and to interpret the data. Traditional system, BOWs model is used along with DSA in order to classify the reviews as positive, negative and neutral. However, the performance of Bag of Words sometimes remains limited due to some fundamental deficiencies in handling the polarity shift problem. Proposed system uses a dictionary based classification for accurately classifying the reviews as negative, positive and neutral. To enhance the accuracy in the classification of neutral reviews, Support Vector Machine algorithm is implemented. Product owner and the user both can identify the quality of the product based on the sentiment graph that is generated based on the reviews for each of the product video. A relative study of the sentiment graphs is performed in order to improve the efficiency of visual representation.

Keywords: Opinion Mining, Sentiment Analysis, Data Mining, natural language processing, machine learning.

ARTICLE INFO

Article History

Received: 12th December 2017

Received in revised form :

12th December 2017

Accepted: 14th December 2017

Published online :

14th December 2017

I. INTRODUCTION

Currently, with the growing amount of online reviews available on the Internet, sentiment analysis and opinion mining, as a special text mining task for determining the subjective attitude (i.e., sentiment) expressed by the text, is becoming a hotspot in the field of data mining and natural language processing. Sentiment classification is a basic task in sentiment analysis, with its object to classify the sentiment (e.g., positive or negative) of a given text. Analyzing these reviews will enhance both the end users and the account executive. Sentiment analysis has

earned its identification and is used in classifying the reviews. Sentiment analysis, also known as opinion mining, is the field of study that analyses people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. It represents a large problem space also there are more names and slightly different tasks, e.g., opinion mining, opinion extraction, sentiment analysis, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, review mining, etc. However, they are now all under the

umbrella of sentiment analysis or opinion mining. In academia both sentiment analysis and opinion mining are frequently employed. They basically represent the same field of study. The meaning of opinion itself is still very broad. Sentiment analysis and opinion mining mainly focuses on opinions which express or imply positive or negative sentiments. To do an analysis, classification plays a key role in opinion mining. A Classification Algorithm is a procedure for selecting a hypothesis from a set of alternatives that best fits a set of observations. Opinions are central to almost all human activities because they are key influencers of our behaviours. Whenever there is a need to make a decision, others' opinions are required. In the real world, businesses and organizations always want to find consumer or public opinions about their products and services. In previous, when an individual needed opinions, she/he asked friends and family. When a business or an organization required public or consumer opinions, it conducted surveys, opinion polls, and focus groups. Acquiring consumer and public opinions has long been a huge business itself for marketing, public relations, and political campaign companies.

Opinion summarization describes opinions of articles by telling sentiment polarities, correlated events and the degree and with opinion summarization, a customer can easily view how the existing customers feel about a product, and the product manufacturer can get the reason why different stands people like it or what they complain about. A seller's job can be quite complicated or it can be quite easy. The two contradictory terms define the selling experience, based on the fact as how seller interprets the consumer interests. Unless one is a psychic or knows how to get into others mind the actual demand of the customer's and the product can't be collaborated. Having a right product is important and equally important is to present it before the right customer.

II. LITERATURE SURVEY

A. Sentiment Analysis and Opinion Mining

This concept is the field of study that analyzes sentiments, peoples opinions, evaluations, attitudes, and emotions from written language.

This technique is most active research areas in natural language processing and is also widely studied in data mining, Web mining, and text mining. This research has scope outside of management sciences to the computer science and social sciences due to its importance to business and society as a whole. The gaining importance of sentiment analysis coincides with the develop of social media like as reviews, forum discussions, blogs, micro-blogs, Twitter, and social networks. we have a large opinionated data recorded in digital form for analysis. This systems are being applied in every business and social domain because opinions are central to almost all human activities and are key influencers of our behaviors. Our perceptions and beliefs of reality, and the choices we make, are largely conditioned on how others see and evaluate the world and when we need to make a decision we often seek out the opinions of others. This is true not only for individuals but also for organizations.

B. Sentiment Analysis Model for Polarity Classification Based On Movie Reviews Using Lexicon Based Technique

Bag-of-words (BOW) is now the most popular way to model text in statistical machine learning approaches in sentiment analysis. However, performance of Bag of Words sometimes remains limited due to few fundamental deficiencies in handling the polarity shift problem. The propose model called dual sentiment analysis (DSA) address the problem for sentiment classification. Firstly propose a novel data expansion technique by creating a sentimentreversed review for each training and test review. On this base, 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 and check the test reviews by considering two sides of one review. It extend the DSA framework from polarity (positive-negative) classification to 3-class (positive negative-neutral) classification, by taking the neutral reviews into consideration and by the dual sentiment analysis propose approach is to analysis sentiment as well as its automatic rating count. This can be calculated by using user review on the basis of positive, negative and neutral response. Then calculate all review and display the result analysis.

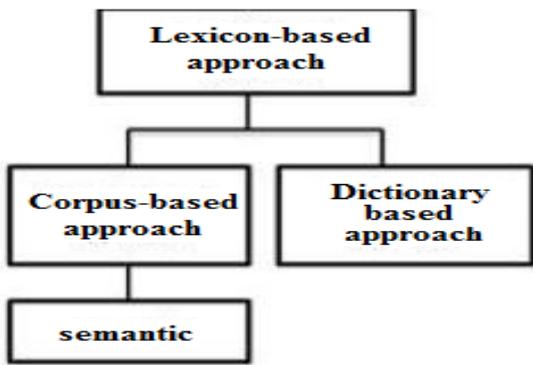


Figure 1: Lexicon based approach.

C. Sentiment Analysis : On Product Review

The Web technology has rapidly changed the way that people express their views and opinions about any product. Now if one wants to purchase a product, people are no longer limited to asking their friends and families because there are many product reviews on the Web which give opinions of existing users of the product.

Here we show the system which facilitates us information about such products and services in summarization form. Finding opinion sources and monitoring them on Web can still be a difficult task because there are large numbers of different sources, and each source may have a huge volume of text with opinions or sentiments. In major cases, opinions are hidden in conference posts and blogs. It is complicated for a human reader to find relevant sources, extract related sentences with suggestions, read them, summarize them, and manage into usable forms.

Thus, automated summarization systems are needed. Using this summarization we can identify the importance, quality, popularity of product and services. In this system we can make summarization for product but we can use this system anywhere, where text analysis is required. Sentiment analysis is also known as opinion mining, grows out of this need.

It is challenging natural language processing or text mining problem. Due to its tremendous value for practical applications, there has been an excessive growth of both research in academia and applications in industry.

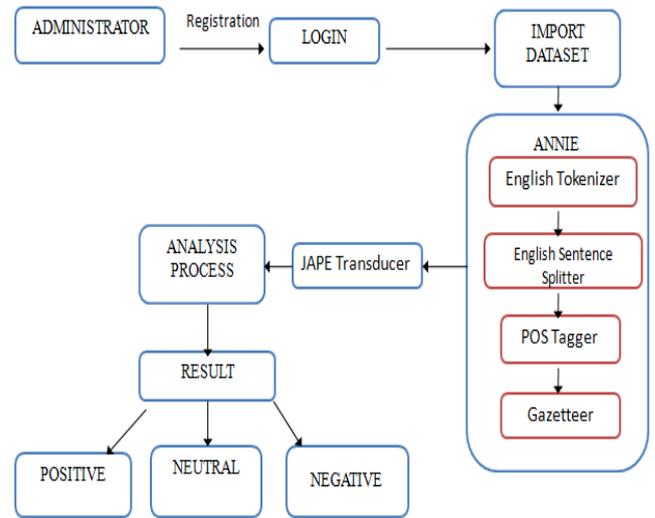


Figure 2: Block Diagram of Sentiment Analysis.

D. Weakly Supervised Joint Sentiment-Topic Detection from Text

Opinion mining or sentiment analysis aims to use automated tools to detect subjective information such as opinions, attitudes, and feelings expressed in text. This technique proposes a novel probabilistic modelling framework called joint sentiment-topic (JST) based on latent Dirichlet allocation (LDA), which detects sentiment and topic simultaneously from text. Joint sentiment topic model called Reverse joint sentiment-topic, obtained by reversing the sequence of topic generation and sentiment in the modeling process, is also studied. Results presents that when sentiment priors are added then JST performs better than Reverse-JST. This is inspected by the practice results on data sets from five different domains. Topics and topic sentiment detected by JST are indeed informative and coherent.

III. PROPOSED ARCHITECTURE

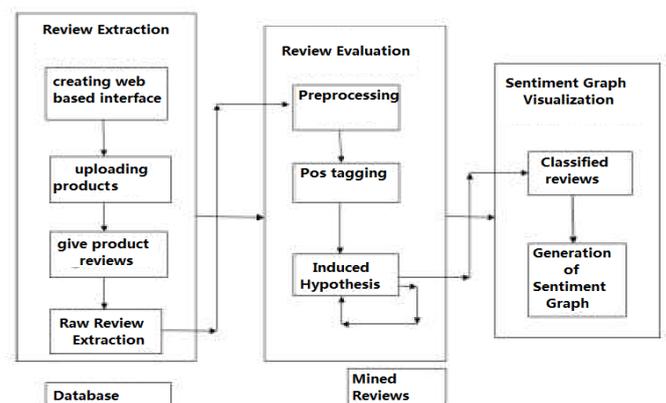


Figure 3: System Architecture.

IV. CONCLUSIONS

Sentiment analysis is essential for anyone who is going to make a decision. Sentiment analysis is helpful in different field for calculating, identifying and expressing sentiment. It is helpful for everyone when they want to buy a product and they can decide which product is best. Sentiment analysis is very important for Enterprises and helps them to know what customers think about their products. Therefore companies can take decisions about their products based on customer's feedback. Thus companies can modify their products features and introduce new products according to customer's opinion in a better and a faster way.

V. ACKNOWLEDGMENTS

I would like to take this opportunity to express my profound gratitude and deep regard to my guide Prof. Guide Name for his exemplary guidance, valuable feedback and constant encouragement throughout the duration of the project. His valuable suggestions were of huge help throughout my project work. His perceptive criticism kept me working to make this project in a much better way. Working under him was an extremely knowledgeable experience for me.

REFERENCES

- [1] Shivaprasad T K and Jyothi Shetty, "Sentiment Analysis of Product Reviews: A Review", International Conference on Inventive Communication and Computational Technologies (ICICCT 2017).
- [2] Chenghua Lin, Yulan He, Richard Everson and Stefan Ruger, "Weakly Supervised Joint Sentiment-Topic Detection from Text", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 24, NO. 6, JUNE 2012.
- [3] Rui Xia, Feng Xu, Chengqing Zong, Qianmu Li, Yong Qi, and Tao Li, "Dual Sentiment Analysis: Considering Two Sides of One Review", IEEE TRANSACTIONS ON JOURNAL NAME, MANUSCRIPT ID, Citation

information: DOI
10.1109/TKDE.2015.2407371.

- [4] Bing Liu, "Sentiment Analysis and Opinion Mining", <http://www.morganclaypool.com/toc/hlt/1/1>, Copyright © 2012 by Morgan & Claypool, DOI: 10.2200/S00416ED1V01Y201204HLT016, ISSN 1947-4040
- [5] Abbasi,A.H. Chen, and A. Salem, "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums," ACM Trans. Information Systems, vol. 26, no. 3, pp. 1-34, 2008.
- [6] Miss.Harshada Deshmukh & Prof.P.L.Ramteke, "Sentiment Analysis Model for Polarity Classification Based On Movie Reviews Using Lexicon Based Technique," Imperial Journal of Interdisciplinary Research (IJIR) Vol-3, Issue-5, 2017 ISSN: 2454-1362, <http://www.onlinejournal.in>.
- [7] Ugandhara Nikumbh, Priti Nikumbh, 3Bhagyashree Gaikwad, "Sentiment Analysis : On Product Review", International Journal for Research in Engineering Application & Management (IJREAM) ISSN : 2494-9150 Vol-02, Issue 02, MAY 2016.