International Journal of Engineering Research and Modern Education (IJERME)
ISSN (Online): 2455 - 4200
(www.rdmodernresearch.com) Volume I, Issue I, 2016

AUTOMATIC SENTIMENT ANALYSIS OF USER REVIEWS

B. Kasthuri* & A. Anitha**

* PG Scholar, Department of Master of Computer Applications, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamilnadu

** Assistant Professor, Department of Master of Computer Applications, Dhanalakshmi Srinivasan Engineering College, Perambalur, Tamilnadu

Abstract:

Data mining is the turning raw data into useful information. The Main use Of data mining is to fetch the exact data what we are exactly need and there is a need to extract useful information from the data and to interpret the data. They are some techniques which is used for data mining (Association rule learning, Clustering, Classification, Anomaly detection, Regression and Summarization). This application is useful for Film owner and normal user to see the review which was given by the peoples and also to check whether the given reviews are positive or neutral or negative reviews. In Existing system, they use BAG OF WORDS to find whether the given review is positive or negative but they dint find the neutral review. In Proposed system, Same BAG OF WORDS is used, by using this BOW they find the dual sentiment analysis 3 which find all positive and negative and also neutral reviews which was given by the people. So the film owner can identify the level of the film by seeing the sentiment graph which was generated automatically based on the reviews which was given by the users. And also the user can see the rating for that particular film.

Index Terms: Natural Language Processing, Machine Learning, Sentiment Analysis & Opinion Mining

1. Introduction:

Sentiment analysis refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information in source materials. Sentiment analysis is widely applied to reviews and social media for a variety of applications, ranging from marketing to customer service. Analysers used for polarity identification. Analysers are of two types manual (domain oriented) and automatic (generalized oriented) we used domain oriented in are methodology. In manual analyser predefined data set exit which similar/ related term have to feed and result occurs and other hand automatic analyser .consist huge data set and also capable to handle multiple language at a time. Sentiment analysis is used to classify polarity and the sentiment analyser is used to define polarity opinion expressed is (+) tive, (-) tive or (=) neutral.

Sentiments:

Sentiment is a sincere and refined sensibility, a tendency to be influenced by emotion rather than reason or fact: to appeal to sentiment. Sentimentality implies affected, excessive, sometimes mawkish sentiment: weak sentimentality. The study of emotions in text can be conducted from two points of view. Firstly, one can investigate how emotions influence a writer of a text in choosing certain words and/or other linguistic elements. Secondly, one can investigate how a reader interprets the emotion in a text, and what linguistic clues are used to infer the emotion of the writer.

Sentiment Analysis:

Sentiment Analysis is process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the

writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral. "Sentiment Analysis is the task of identifying positive and negative opinions, emotions, and evaluations". Sentiment Analysis has many names. It's often referred to as subjectivity analysis, Opinion mining, and appraisal extraction, with some connections to affective computing (computer recognition and expression of emotion).

Dual Sentiment Analysis:

A model called dual sentiment analysis (DSA), to address this problem for sentiment classification. We first propose a novel data expansion technique by creating a sentiment reversed review for each training and test review. On this basis, 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.

Data Flow Diagram:

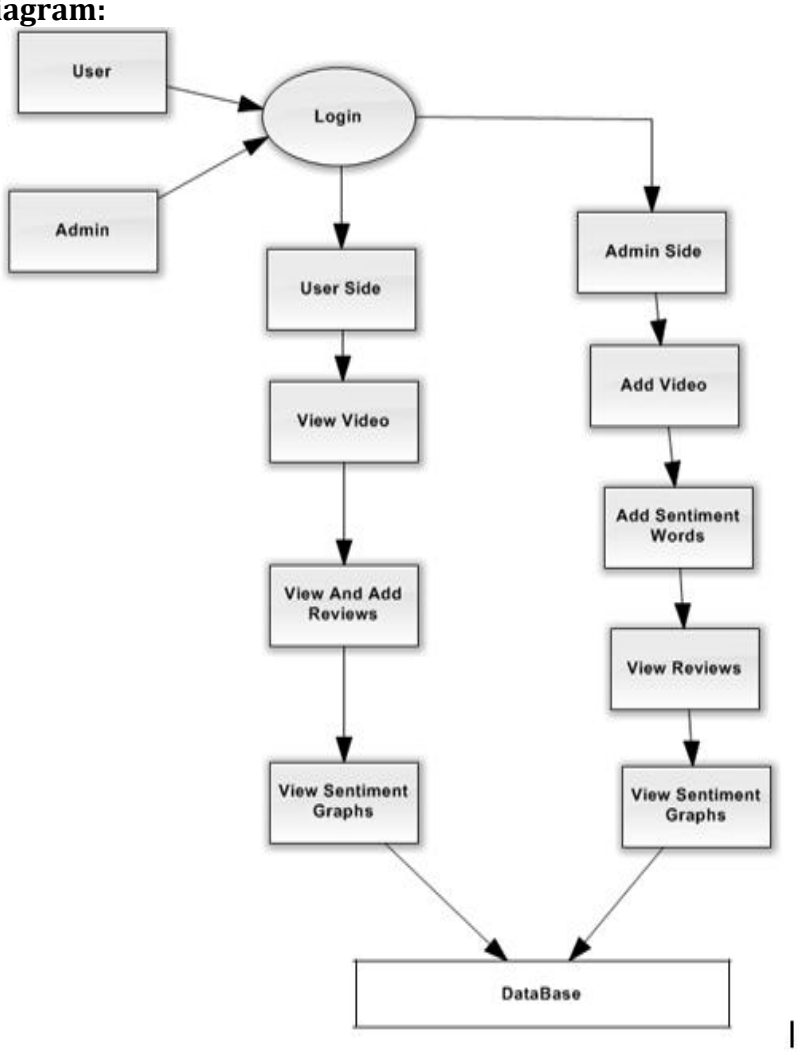


Figure 1: Data Flow Diagram

The reminder of this paper is organized as follows. Section II, describes the Related Works. Section III, describes the Proposed Work. Section IV, describes the Experimental Evaluation and Results. Section V summarizes the Conclusion and Future Enhancement.

2. Related Works:

The work of sentiment analysis and polarity shift, and then review the technique of data expansion.

Sentiment Analysis and Polarity Shift:

According to the levels of granularity, tasks in sentiment analysis can be divided into four categorizations: document- level, sentence-level, phrase-level, and aspect-level sentiment analysis. First, we strengthen the DSA algorithm by adding a selective data expansion procedure. Second, we extend the DSA framework from sentiment polarity classification to positive-negative-neutral sentiment classification. Third, we propose a corpus-based method to construct a pseudo-antonym dictionary that could remove DSA's dependency on an external antonym dictionary..

Data Expansion Technique:

The field of natural language processing and text mining, Agirre and Martinez proposed expanding the amount of labeled data through a Web search using monospermous synonyms or unique expressions in definitions from Word Net for the task of word sense disambiguation. Fujita and Fujino proposed a method that provides reliable training data using example sentences from an external dictionary. To the best of our knowledge, the data expansion technique proposed here is the first work that conducts data expansion in sentiment analysis. Different from the above mentioned techniques, the original and reversed reviews are constructed in a one-to-one correspondence.

3. Proposed Work:

A model called dual sentiment analysis (DSA), to address this problem for sentiment classification and also they use Bag of Words. By using these two we can find ,whether the reviews which was given by the user is positive or negative. And also in this proposed system we can find the neutral word which is very difficult to find.

Advantages:

- ✓ Output is more efficient
- ✓ It Can Analyse negative, positive and neutral comment
- ✓ It is easy to find neutral comments.

Modules:

- ✓ Authentication Module
- ✓ Add Film module
- ✓ User View Reviews
- ✓ User Give Reviews
- ✓ Bag Of Words
- ✓ Dual Sentiment Analysis
- ✓ Sentiment Graph

Authentication Module:

In this application they are only two users, one the main site owner Admin and another one is a normal users. For Admin there is registration process because he is the site owner. But, for users there is a registration process, from this they will get a user id and password to get into this application. And Admin also have his particular user id and password to enter this application.

Add Film Module:

After completion of login process by the Admin, he will go for add film module. The main concept of this module is to make marketing for the upcoming films in film industry. For the purpose, Admin have to add the teaser or trailer video of the films which are going to release as soon as possible.

User View Reviews:

In this module, once the teaser or trailer was uploaded by a admin, it can be viewed by each and every user who registered in this application. Even though they can

view the reviews which was given my user who registered in this application. From these reviews the user can get some idea about the films.

User Give Reviews:

In this module, after seeing the film which was added by the admin the user can give their own comments about the film. In their comment itself a sentiment star will be generated for their comments. If their comment is positive means that star will be in green color, negative means the star will be in red color and if means that star will be in yellow color.

Bag of Words:

Bag of words is same as dictionary. In this application they are two bags of words, one bag of word is for positive word and another bag of word is for negative word. Once, the review was given by the users for that particular film. This both bag of words are used to check the sentiment about that revi.

Dual Sentiment Analysis:

The main theme of this application is coming in this module. This dual sentiment analysis will check whether the given review by the user is positive review or negative review and also it will find the neutral review. Once the comment is given by the user, Each and every word in the comment will be checked whether it is a positive or negative using bag of words.

Sentiment Graph:

This sentiment graph is to analyze, whether the given comments or positive or not. In this they are three graph one is positive graph, other one is negative graph and the last one is neutral graph. If the film has more positive comment means positive graph will be higher, if the film has more negative comment means negative graph will be higher and if the film has more neutral comment means neutral graph will be higher.

4. Experimental Analysis and Results:

Various life cycle processes such as requirement analysis, design phase, verification, testing and finally followed by the implementation phase result in a successful project management. System implementation is an important stage of theoretical design is turned into practical system. The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

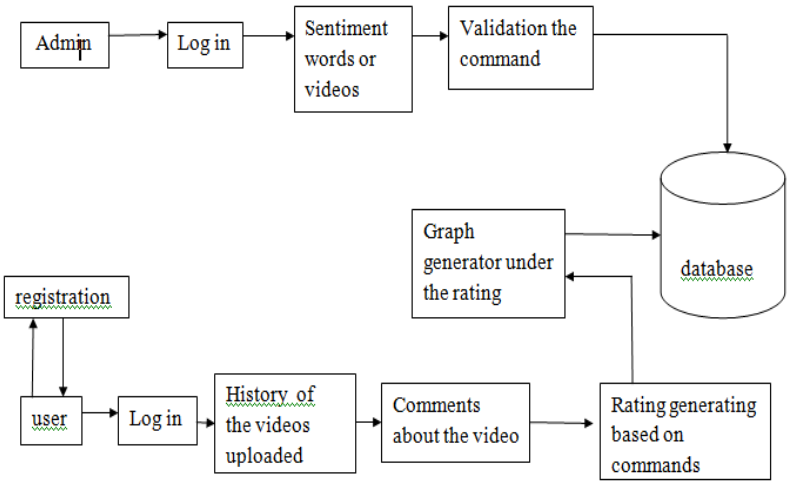
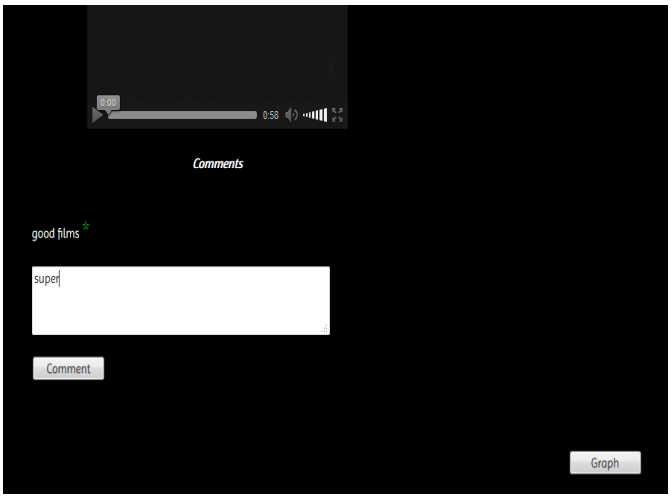


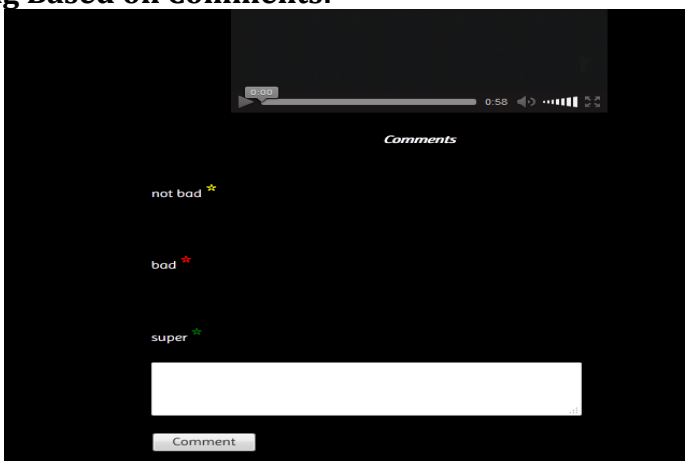
Figure 2: Sentiment Analysis Reviews Based on Commands

Each program is tested individually at the time of development using the data and has verified that this program linked together in the way specified in the programs specification, the computer system and its environment is tested to the satisfaction of the user. The system that has been developed is accepted and proved to be satisfactory for the user and so the system is going to be implemented very soon. A simple operating procedure is included so that the user can understand the different functions clearly and quickly. The final stage is to document the entire system which provides components and the operating procedures of the system.

View Videos and Give The Comments:



Ratting Generating Based on Comments:



View Graph by User:

5. Conclusion:

This project is mainly used to identify whether the comments given by the user is positive, negative or neutral. In this, Admin will add the film videos, so user can see that videos and they can give comments for those videos. By using bag of words and dual sentimental analyse 3 we can find the given comment is positive or not. And also we can see the sentimental graph for the comments.

6. Future Enhancement:

Furthermore extend the DSA algorithm to DSA3, which could deal with 3-class (positive-negative-neutral) sentiment classification. The dual training and dual prediction algorithm by taking the neutral reviews into consideration. The experimental results also prove the effectiveness of DSA3 in 3-class sentiment classification. Finally, to remove DSA's dependency on an external antonym dictionary, we propose a corpus-based method to construct a pseudo-antonym dictionary. We also plan to consider more complex polarity shift patterns such as transitional, subjunctive and sentiment-inconsistent sentences in creating reversed reviews.

7. References:

1. M. Hu and B. Liu, "Mining opinion features in customer reviews," Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 2004.
2. V. Ng, S. Dasgupta and S. Arifin, "Examining the Role of Linguistic Knowledge Sources in the Automatic Identification and Classification of Reviews," Proceedings of the International Conference on Computational Linguistics and Annual Meeting of the Association for Computational Linguistics (COLING/ACL), pp. 611-618, 2006.
3. J. Na, H. Sui, C. Khoo, S. Chan, and Y. Zhou, "Effectiveness of simple linguistic processing in automatic sentiment classification of product reviews," Proceedings of the Conference of the International Society for Knowledge Organization (ISKO), 2004
4. S. Fujita and A. Fujino, "Word sense disambiguation by combining labeled data expansion and semi-supervised learning method," Proceedings of the International Joint Conference on Natural Language Processing (IJCNLP), pp. 676-685, 2011.