Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
CONFCALL - 2019 Conference Proceedings

Machine Learning based on Sentiment Analysis using Facebook

K. Leela Rani
Dept of CSE
Besant Theosophical College

Abstract:- **Semantics plays an important role in the accurate analysis of the context of sentiment expression. Sentiment Analysis is the process of determining whether a piece of writing is positive, negative or neutral. Consumers regularly face the trade-off in purchase decisions so nowadays if one wants to buy a consumer product one prefer user reviews and discussion in public forums on web about the product. Many consumers use reviews posted by other consumers before making their purchase decisions. People have a tendency to express their opinion on various entities. As a result, opinion mining has gained importance. Sentiment Analysis deals with evaluating whether this expressed opinion about the entity has a positive or a neutral or a negative orientation.**

INTRODUCTION

Company strategies, marketing campaigns, and product preferences. Many new and exciting social, geo political, and business-related research questions can be answered by analyzing the thousands, even millions, of comments and responses expressed in various blogs (such as the blogosphere), forums (such as Yahoo Forums), social media and social network sites (including YouTube, Face book, and Flikr), virtual Worlds (such as Second Life), and tweets (Twitter). Opinion mining, a sub discipline within data mining and computational linguistics, refers to the computational techniques for extracting, classifying, understanding, and assessing the opinions expressed in various online news sources, social media comments, and other user-generated content. Sentiment analysis is often used in opinion mining to identify sentiment, affect, subjectivity, and other emotional states in online text. For example, we might seek to answer these questions:
• What were the opinions of young US voters toward the Democratic and Republican presidential candidates during the most recent election?
• Since September 11, how do the international Jihadi forums introduce radical ideology and incite young members?
• What are the opinions and comments of investors, employees, and activists toward Wal-Mart in light of its cost-reduction efforts and global business practices?
• What was the most successful McDonald's promotional campaign conducted recently in China. and why did it succeed? Which McDonald's product is most preferred by young students in China and why? Much advanced research in this area has recently focused on several critical areas.1,2 In this installment of Trends & Controversies and the next, we review several contributions to this emerging field. The topics covered include how to extract opinion, sentiment, affect, and subjectivity expressed in text. For

example, resources online might include opinions about a product or the violent and racist statements expressed in political forums. Researchers have also been able to classify text segments based on sentiment, affect, and subjectivity by analyzing positive or negative sentiment expressed in sentences, the degree of violence expressed in forum messages, and so on.

Our work builds on previous studies focusing on the relationship betweenthe discussions held in firm-specific finance Web forums and public stock behavior. However, instead of assuming a shareholder view of participants in a finance Web forum as in previous research, and considering them to be uniformly representative of investors, we adopted a stakeholder perspective. This perspective more accurately represents the diversity of the constituency groups participating in the Web forum and closely aligns the analysis with the corporation's stakeholder theory. To address the broad questions posed in this research, and guided by the literature reviewed, we developed a framework for analysis with four major stages: stakeholder analysis, topical analysis, sentiment analysis, and stock modeling. During the stakeholder analysis stage, we identified the stakeholder groups participating in Web forum discussions. In the topical analysis stage, the major topics of discussion driving communication in the Web forum are determined. The sentiment analysis stage consists of assessing the opinions expressed by the Web forum participants in their discussions. Finally, in the stock modeling stage, we examine the relationships between various attributes of Web forum discussions and the firm's stock behavior.

EXISTING SYSTEM:

In earlier days the movie reviews are take manually in the form 1 star to 5 stars which are given by users or viewers, and the whole movie rating is given by average.
• And the other way to give rating is manually analysing the text which was written by viewers or users.
• And this is very hard to analyse all the reviews by human and time taking process.
• To overcome these problems automated machine learning approach is playing an important role.

PROPOSED SYSTEM:

Our work builds on previous studies focusing on the relationship between the discussions held in firm-

specific finance Web forums and public stock behavior. However, instead of assuming a shareholder view of participants in a finance Web forum as in previous research, and considering them to be uniformly representative of investors, we adopted a stakeholder perspective. This perspective more accurately represents the diversity of the constituency groups participating in the Web forum and closely aligns the analysis with the corporation's stakeholder theory.

To address the broad questions posed in this research, and guided by the literature reviewed, we developed a framework for analysis with four major stages: stakeholder analysis, topical analysis, sentiment analysis, and stock modeling. During the stakeholder analysis stage, we identified the stakeholder groups participating in Web forum discussions. In the topical analysis stage, the major topics of discussion driving communication in the Web forum are determined.

The sentiment analysis stage consists of assessing the opinions expressed by the Web forum participants in their discussions. Finally, in the stock modeling stage, we examine the relationships between various attributes of Web forum discussions and the firm's stock behavior.

MODULE DESCRIPTION:

1.POSTING OPINIONS:
In this module, we get the opinions from various people about business, e-commerce and products through online. The opinions may be of two types. Direct opinion and comparative opinion. Direct opinion is to post a comment about the components and attributes of products directly. Comparative opinion is to post a comment based on comparison of two or more products. The comments may be positive or negative.

OBJECT IDENTIFICATION:
In general, people can express opinions on any target entity like products, services, individuals, organizations, or events. In this project, the term object is used to denote the target entity that has been commented on. For each comment, we have to identify an object. Based on objects, we have to integrate and generate ratings for opinions.
The object is represented as "O". An opinionated document contains opinion on set of objects as $\{o_1, o_2, o_3 \ldots o_r\}$.

FEATURE EXTRACTION:
An object can have a set of components (or parts) and a set of attributes (or properties) which we collectively call the features of the object. For example, a cellular phone is an object. It has a set of components (such as battery and screen) and a set of attributes (such as voice quality and size), which are all called features (or aspects). An opinion can be expressed on any feature of the object and also on the object itself.

With these concepts in mind, we can define an object model, a model of an opinionated text, and the mining objective, which are collectively called the feature-

based sentiment analysis model. In the object model, an object "O" is represented with a finite set of features,
$F \square \square \{f_1, f_2, \ldots, f_n\}$
which includes the object itself as a special feature. Each feature $f_i \in \square F$ can be expressed with any one of a finite set of words or phrases
$W_i \square \square \{w_{i1}, w_{i2}, \ldots, w_{im}\}$
which are the feature's synonyms.

OPINION-ORIENTATION DETERMINATION:
The opinion holder is the person or organization that expresses the opinion. In the case of product reviews and blogs, opinion holders are usually the authors of the posts.An opinion on a feature f (or object o) is a positive or negative view or appraisal on f (or o) from an opinion holder. Positive and negative are called opinion orientations. From this opinion orientation we have to determine the type of opinion whether it is direct opinion or comparative opinion.

❖ **DIRECT OPINION:**
A direct opinion is a quintuple (o_j, f_{jk}, oo_{ijkl}, h_i, t_l),
Where o_j is an object,
f_{jk} is a feature of the object o_j,
oo_{ijkl} is the orientation of the opinion on feature f_{jk} of object o_j,
h_i is the opinion holder, and
t_l is the time when the opinion is expressed by h_i.
The opinion orientation oo_{ijkl} can be positive, negative, or neutral.

❖ **COMPARATIVE OPINION:**
A comparative opinion expresses a preference relation of two or more objects based their shared features. A comparative opinion is usually conveyed using the comparative or superlative form of an adjective or adverb, such as "Coke tastes better than Pepsi."

INTEGRATION:

Integrating these tasks is also complicated because we need to match the five pieces of information in the quintuple. That is, the opinion oo_{ijkl} must be given by opinion holder h_i on feature f_{jk} of object o_j at time t_l. To make matters worse, a sentence might not explicitly mention some pieces of information, but they are implied using pronouns, language conventions, and context. Then generate ratings based on above tasks. Thus we can clearly see how holders view the different features of each product.

CONCLUSION:

We developed to test the hypotheses regarding the relationship between the Web forum discussions and the Wal-Mart stock behavior directly followed those proposed in previous research. The models examined the correlation and developed both contemporaneous and predictive regression models using the variables presented in the

Special Issue - 2019

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
CONFCALL - 2019 Conference Proceedings

research design. In the contemporaneous Regressions, the stock behavior variables are regressed on the measures of the forum discussions occurring on the same day. In the predictive regressions, measures of the forum discussions occurring on a specific day are utilized to predict the stock behavior on the following trading day. We summarize selected results from predictive regression using overall forum data.

In this Trends & Controversies department and the next, we include three articles on opinion mining from distinguished experts in computer science and information systems. Each article presents a unique innovative research framework, computational methods, and selected results and examples. In this first issue, Bing Liu's article "Sentiment Analysis: A Multifaceted Problem" argues that sentiment analysis is not a single problem, but a combination of many facets or sub problems. Liu introduces some of these problems and suggests several technical challenges, including object identification, feature extraction and synonym grouping, opinion-orientation classification, and integration.

FUTURE WORK:

Building on what has been done so far, I believe that we just need to conduct more refined and in-depth investigations as well as build integrated systems that try to deal with all the problems together because their interactions can help solving each individual problem. I am optimistic that the problems will be solved satisfactorily in the next few years for widespread applications. It can also perform integration to a good extent based on several automated discovery functions. For real-life applications, a completely automated solution is nowhere in sight.

However, it is possible to devise effective semi-automated solutions.

The key is to fully understand the whole range of issues and pitfalls, cleverly manage them, and determine what portions can be done automatically and what portions need human assistance. In the continuum between the fully manual solution and fully automated solution, we can push more and more toward automation. Beyond what have been discussed so far, we also need to deal with the issue of opinion spam, which refers to writing fake or bogus reviews that try to deliberately mislead readers or automated systems by giving untruthful positive or negative opinions to promote a target object or damage the reputation of Another object. Detecting such spam is vital as we go forward because spam can make sentiment analysis useless. Finally, despite these difficulties and challenges, the field has made significant progress over the past few years.

This is evident from the large number of start-up companies that provide sentiment-analysis and opinion mining services. A real, substantial need exists in industry for such services. This practical need and the technical challenges will keep the field vibrant and lively for years to come.

REFERENCES

[1] User Interfaces in C#: Windows Forms and Custom Controls by Matthew MacDonald.
[2] Applied Microsoft®.NET Framework Programming (Pro-Developer) by Jeffrey Richter.
[3] Data Communications and Networking, by Behrouz A Forouzan.
[4] Computer Networking: A Top-Down Approach, by James F. Kurose.
[5] Operating System Concepts,by Abraham Silberschatz.
[6] B. Liu, "Sentiment Analysis and Subjectivity," Handbook of Natural LanguageProcessing, 2nd ed., N. Indurkhya and F.J. Damerau, eds., Chapman & Hall, 2010, pp. 627–666.