

Actionable Business Intelligence for E-Commerce Using Data Mining Tool

¹Nelson Anand M, ²Karthick S

¹M.Tech Student, ²Assistant Professor,

Software Engineering, SRM University, Chennai, India

¹mjanand4455@yahoo.com, ²karthik.sa@ktr.srmuniv.ac.in

Abstract— A traditional Enterprise BI solution has the capability to extract internal organizational data and transform it into useful information. Ignores the importance of data available in social media. In recent times, social networks have become virtual business goldmines. Those organizations best able to differentiate themselves have an ability to incorporate social media analytics into their customer processes, to monetize their investments and integrate insight into their customer data. That would build relationship to new customers and stronger connections with existing ones. It is characterize the e-commerce framework of fetching data from social networks by using (Qlik view) tool. It is proposed that the implementation of principal elements of an e-commerce. It is becoming a realization that BI can render better decision-making when corporate data is analyzed in integration with social media & real time data.

Index Terms—Social media data, goldmines, monetize, e-commerce, BI, corporate data

I. INTRODUCTION

As e-commerce has grown rapidly, people are able to buy and sell products online easily, and then various information sources are available for customer's purchase references. One of the important information which is available from many ecommerce web sites is product reviews. These Product reviews contain customers' feelings or opinions which is very important for managers or business to find the orientation of customer feeling. However, as the quantities of product reviews are often large, it is difficult for them to read all of them and identity orientation. Nowadays, e-commerce is growing fast, so product reviews have grown rapidly on the web. The large number of reviews makes it difficult for manufacturers or businesses to automatically classify them into different semantic orientations (positive, negative, and neutral). Semantic orientation or sentiment classification is a method for automatic classification of product reviews into three classes: positive, negative and neutral, thus helping managers classify them. This classification method is usually used to classify customer's review in a whole to determine its class.

II. RELATED WORKS

Semantic orientation or sentiment classification is a method for automatic classification of product reviews into three classes: positive, negative and neutral, thus helping managers classify them. This classification method is usually used to classify customer's review in a whole to determine its class. One of the existing methods in opinion mining is holistic lexicon based approach; although this approach is efficient, it is not able to classify reviews granularity. The aim of this paper is to increase the accuracy of that approach by propose a fuzzy logic model to identify customers' semantic orientations in product reviews at a smaller granularity level opinion words.

Recently many researchers have studied the problem called opinion mining and sentiment analysis. One of the tasks in this area is: sentiment classification. Sentiment classification is a way to classify each review as positive, negative, or neutral. Representative works on classification at the document level include. Sentence level subjectivity classification is studied in which determines whether a sentence is a subjective sentence (but may not express a positive or negative opinion) or a factual one. Sentence level sentiment or opinion classification is studied A review sentence can contain multiple features, and the orientations of opinions expressed on the features can also be different(e.g. "the picture quality of this camera is good, but the battery life is short." picture quality" and "battery life" are features. The opinion on "picture quality" is positive, and the opinion on "battery life" is negative. Other related works at both the document and sentence levels include those in the above description

III. PROPOSED SYSTEM

The proposed method is based on the combinations of opinion words around each product feature in a review sentence. This methodology determines the strength of opinion orientation (positive, negative, moderate) on the product feature using the following statements

- 1) We extremely enjoy this camera.
- 2) We like this camera.
- 3) The picture quality is very good.
- 4) The picture quality is good

Although words that express positive and negative orientation are usually adjective and adverb, verb and noun can be used to express opinion. Consider that sentence: "very" as adverb, "good" as adjective, "enjoy" and "like" as verb. Researchers have compiled set of words and phrases for adjective, adverb and verb and noun respectively. Such lists are collectively called the opinion lexicon. Each set usually obtained through a bootstrapping process.

WAMP Server

WAMP Server is a Windows web development environment. It allows you to create web applications with Apache2, PHP and a My SQL database. Alongside, PHP, My Admin allows you to manage easily your databases. The data have been extracted from Social media through API using PHP language. The server refers to a set of free open source application. It have been combined with Microsoft windows which are commonly used in web server environments. By running local apache web server on a windows machine a web developer can test web pages in a web browser.

Technical Process

Traditionally, enterprise data analysis gives an insight to business performance. These days, Social Media Data (SMD) is very huge and the organizations have started listening to customers on social media. Those organizations best able to differentiate themselves have an ability to incorporate social media analytics into their customer processes, to monetize their investments and integrate insight into their customer data. This provides a foundation from which organizations can link their social media strategies to their revenue generation. But this must also involve cross-functional teams, such as IT and product teams all with the goal of driving revenue. Social media provides the means to do just that. And an integrated business analytics approach provides additional depth to capture, analyze and act on this insight. It is becoming a realization that BI can render better decision-making when corporate data is analyzed in integration with social media & real time data.

- **Business Problem**

These days, many companies have a dedicated team monitoring their facebook page. Mostly, they use their social media data distinct and separate from their corporate data. Consequently, analysis of their social media data cannot be considered as an indicator of their actual business positioning in the market. This effectively renders the social media analysis incorrect and is just being a redundant activity being performed by the company.

- **Business Solution**

Our proposed solution is to integrate the social media data with the corporate data. In other words, our proposal was to integrate interactive intelligence into business intelligence so that analysis of such data would give a relevant and practical insight into the performance of the business.

Data Extraction from Social Media

Data extraction and data mining can help you take strategic business decisions that can shape your business goals. We provide data extraction services from a variety of sources such as Websites, Databases, Images and Documents. This enables you to have data quickly and in the right format. Social networks contain a huge quantity of information, and this information is very valuable because users talk and interact naturally with each other in them. There are also precious discussions and opinions from consumers about products, organizations and many different topics. Another important information that can be gathered from social networks are the relationships between users and communities. It is a great market advantage to be able to easily access this information and get to know the market and what customers think about a product, brand or topic.

PHP through API for Twitter can extract social media data that match the search criteria that you specify, making it much more efficient for social media data mining. You can define search criteria, search for topics, and extract social media data from all three social media networks. You can load the extracted social media data to a target and then use the data for text analytics and sentiment analysis. You can download the social media data mining demo file that includes the necessary sample mappings, workflows, and sessions. Using the demo file, you can extract the following social media data. Tweets that contain the topic and the corresponding user details.

Sentiment Analysis

Sentiment analysis (also known as opinion mining) refers to the use of natural language processing, text analysis and to identify and extract subjective information in source materials.

Generally speaking, sentiment analysis aims to determine the attitude of a speaker or a writer with respect to some topic or the overall contextual polarity of a document. Sentiment analysis and opinion mining is the field of study that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions from written language. It is one of the most active research areas in natural language processing and is also widely studied in data mining, Web mining, and text mining. In fact, this research has spread outside of computer science to the management sciences and social sciences due to its importance to business and society as a whole.

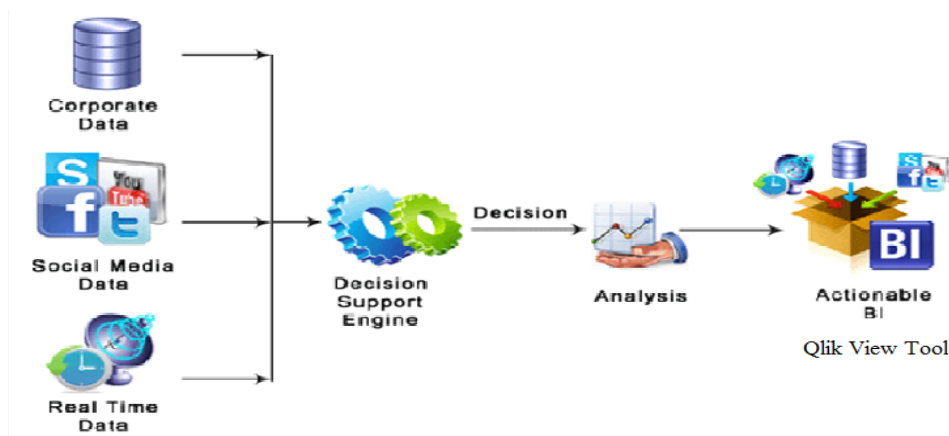


Fig 1

Data is extracted from social media sites through API using a PHP code. The relevant data is provided in the format in decoded with help of PHP Coding and then it is stored in the database. The data is scheduled to be loaded into a source table in SQL server 2008 on a daily basis. The data is then cleansed and moved to a staging area.

Real time data is extracted from appropriate websites. The data is populated to a cumulative staging area. The data is then transformed and loaded to the Target data model in SQL server. The sentiment engine analyses tweets/comments/messages and categorize them into positive, negative and neutral entities. The sentiment data is also loaded to the Target data model. The data from various sources are loaded into the cumulative staging area following the same steps above. The reports & dashboards are generated for analysis. Decisions made out of the above analysis will then be taken into consideration by Actionable business intelligence. Text mining is done on the comments for categorizing the messages into different threads viz., jobs, product release, product update, events etc.

The user response is compared with the corporate data by joining the social media table with the enterprise table using the user's email id. The product performance is then mapped to the corporate data by joining the poster name with the product name

IV. EXPERIMENTAL RESULTS

Data Extraction

The customers reviews are being extracted from the respective social media web sites then its converted into the meaningful format that would be stored in the database.

Sentiment Analysis

In this section the Sentiment analysis or semantic orientation will be done. These comments are classified as a positive or negative statements based on the reviewers feedbacks.

Dashboards

Finally the report will be submitted as the dashboards. When we integrate the social media data to corporate data using this it makes interactive intelligence into Business intelligence.

V. CONCLUSION

Our interconnected world and the rising tide of social media only makes it more compelling to integrate corporate data with data from social media and real-time. The fact is, social media puts opportunities and your reputation on the line every time someone logs on. And like it or not, a social network leaves users and friends in total control to influence your brand with a single networked comment favorably, or unfavorably. To successfully execute organization's goals and objectives, social media data would play an important role. Organization needs to draw a clear strategy to bring social media unstructured data into the current data warehousing. This framework helps in a holistic analysis of data including all types of data which in turn helps in better decision making. This paper proposed for identifying semantic orientation of opinions expressed by reviews on product feature. The method is able to classify reviews into five classes: very strong negative (or positive), strong negative (or positive), moderate negative (or positive), very weakly negative (or positive) and very weakly negative (or positive). As a feature work, we intent to implement our method using benchmark dataset from report our interval result.

REFERENCES

- [1] Andreevskaia and S. Bergler. Mining Word Net for Fuzzy Sentiment: Sentiment Tag Extraction from Word Net Glosses. In EACL'06, pp. 209-215,2006.
- [2] P. Beineke, T. Hastie, C. Manning, and S. Vaithyanathan. An Exploration of Sentiment Summarization. In Proc. Of the AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications, 2003.
- [3] S. Das, and M. Chen. Yahoo! for Amazon: Extracting market sentiment from stock message boards. APFA'OI, 200 1.
- [4] K. Dave, S. Lawrence, and D. Pennock. Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews. WWW'03,2003