

SENTIMENT ANALYSIS FOR MARKETING

Abstract Sentiment analysis, the automated extraction of expressions of positive or negative attitudes from text has received considerable attention from researchers during the past decade. In addition, the popularity of internet users has been growing fast parallel to emerging technologies; that actively use online review sites, social networks and personal blogs to express their opinions. They positive and negative attitudes about people, organizations, places, events, and ideas. The tools provided by natural language processing and machine learning along with other approaches to work with large volumes of text, makes it possible to begin extracting sentiments from social media.

I INTRODUCTION

People make judgments about the world around them when they are living in the society. They make positive and negative attitudes about people, products, places and events. These types of attitudes can be considered as sentiments. Sentiment analysis is the study of automated techniques for extracting sentiments from written languages. Growth of social media has resulted in an explosion of publicly available, user generated text on the World Wide Web. These data and information can potentially be utilized to provide real-time insights into the sentiments of people [1]. Blogs, online forums, comment sections on media sites and social networking sites such as Facebook and twitter all can be considered as social media. These social media can capture millions of peoples' views or word of mouth. Communication and the availability of these real time opinions from people around the world make a revolution in computational, linguistics and social network analysis. Social media is becoming an increasingly more important source of information for an enterprise. This paper reveals an approach which is implemented as a tool that can analyse sentiments on twitter social media addressing above issues and then develop an application to generate knowledge that can be useful for business environments using people's attitudes about their products and services.

II LITERATURE REVIEW

This section illustrates other similar work related to analysing sentiments. Most of these approaches analyse sentiments as positive and negative while some approaches are in research level and few more are commercially available. Adobe Social Analytic Adobe Social Analytics basically measures the impact of social media on businesses by understanding how conversations on social networks and online communities influence marketing performance. After capturing and understanding the conversations going on, it correlates the impact of those conversations with key business matrices such as revenue and brand value. Other than that it measures the interactions that businesses have with their customers in social media including how Facebook posts drive site visitors and purchase behaviours [3]. Adobe Social Analytics uses a natural language processing algorithm to implement sentiment analysis. Brand watch Sentiment Analysis

Brand watch is also a sentiment analysis tool developed by a team of PhD qualifiers in the United Kingdom, this is also commercially available currently. Through this tool they are trying to access whether a sentiment is positive, negative or neutral [4]

This is an on **Sentiment 140**

line tool for analysing sentiments of Twitter social network. This tool allows discovering the sentiment of a brand, product or topic on Twitter. This was created by three Computer Science graduate students at Stanford University and their main focus is analysing the languages English and Spanish. Sentiment 140 basically states whether the specified brand, product or topic is positive, negative or neutral [5].



Social Mention

Social Mention is a social media search and analysis platform which analyses user sentiments through social media. This is also an online tool that allows tracking what people are saying about a particular brand, product or topic in real time. This tool allows the user to define a time period in which to analyse user sentiments. The specialty in our system is, it does not only analyse the sentiments, and instead it uses the analysed sentiment scores to provide product profile, trend analysis and forecasting for the user.

III APPROACH

To analyse sentiments and then come to a conclusion through them, we need to have enough sentiments in the correct format. There are thousands and millions of sentiment data in the web, especially in social media sites that can be used to get valuable conclusions. But they are not in a correct format or not in a structured way to get maximum usage out of them. We need to convert them to a correct format and use them as we want. It is the first part in our approach, which is developing a crawler to crawl data from Twitter social media. The Crawler should be able to crawl user sentiments from twitter and at the same time get user details in order to do product profiling for customers as the later part of the whole approach. After having access to large sources of data which is in a structured manner through the crawler and using a database, the next step is analysing sentiments. Sentiments can be in different languages; in our project we cover the English language. Analysing sentiments is a way of processing natural languages, therefore this part is about natural language processing. For this we use Natural Language Toolkit, also known as NLTK which is a leading platform for building Python programs to work with human language data. There are different ways that we can use to analyse sentiment data using

IV DESIGN AND IMPLEMENTATION

There are four main modules naming, crawler, sentiment analysis tool, data mining module and dashboard in this project. The top level architectural diagram of the system as follows.

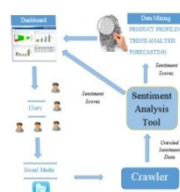


Fig 1: Top level architectural diagram of the system

1. Naïve Bayes Classifier

The Bayesian classification is used as a probabilistic learning method. In naive Bayes classifiers, every feature gets a say in determining which label should be assigned to a given input value. To choose a label for an input value, the naive Bayes classifier begins by calculating the prior probability of each label, which is determined by checking frequency of each label in the training set. The contribution from each feature is then combined with this prior probability, to arrive at a likelihood estimate for each label. The label whose likelihood estimate is the highest is then assigned to the input value.

Following equations are used in calculating label likelihoods as shown in the following figure, Fig 2.1

$$P(\text{label}|\text{features}) = P(\text{features}, \text{label})/P(\text{features})$$

$$= P(\text{label}) P(\text{features}|\text{label})/ P(\text{features})$$

$$= P(\text{label}) \prod_{i=1}^n P(f_i|\text{label}) / P(\text{features})$$

(Since features are independent)

of feature extraction methods because it is the major thing that affects to the accuracy of the classifiers. The evaluation results of the feature extraction methods as well as the classifiers with elected feature extraction method are as follows.

VICONCLUSION

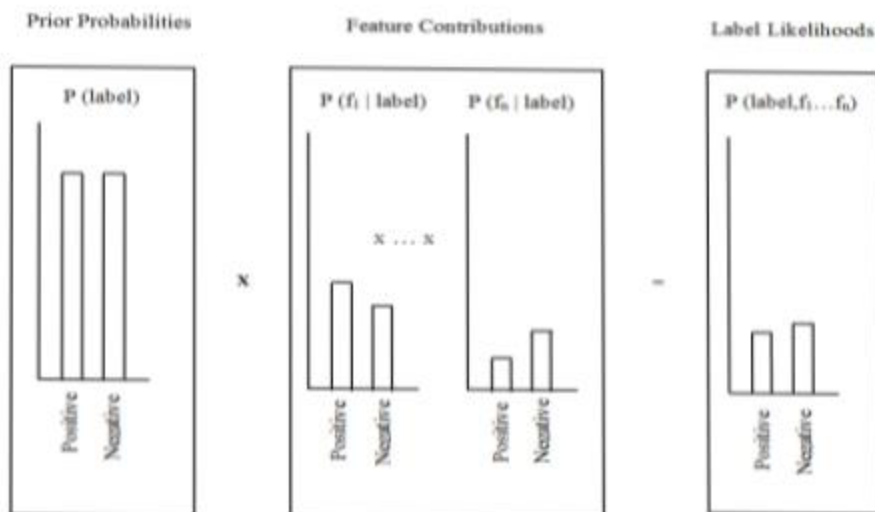


Fig 3: Calculating label likelihoods with Naive Bayes

In this classification method, the naïve Bayes assumption or independence assumption which is independence of the

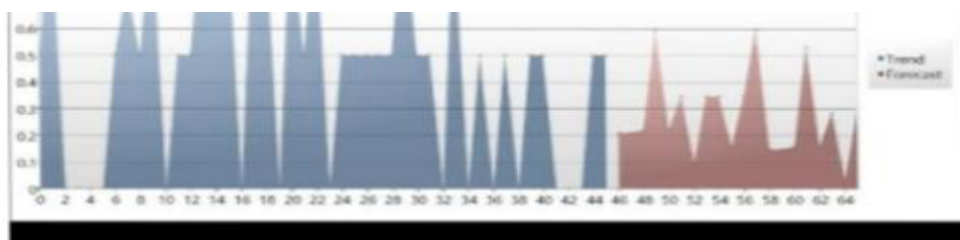


Fig 5: Trend analysis and forecasting



It is a very important fact to analyse how people think in different comes about different things. This becomes more important when it comes to the business world because

V EVALUATION

Here we give priority to evaluate the classifiers in order to select the most probable classifier to use in the project Sentiment Analysis for Social Media. This evaluation was an effective tool to guide us to make improvements to the selected model also. To evaluate these two classifiers, we needed a large labelled text corpus. Since it could find a very large corpus, it trained the classifiers on large set of training data and tested on increasing amount of test data. Prior to state the evaluation results of the classifiers, it needs to present evaluation results business is dependent on their customers and they always try to make products or services in order to fulfil customer requirements. So knowing what they want, what they think and talk about existing products, services and brands is more useful for businesses to make decisions such as identifying competitors and analysing trends. Both because people express their ideas on social media and it can access those data, it has been enabled in some way to do the above mentioned things by using those data. The project. Sentiment Analysis for Social Media does that. From the view at the top level of the project, we get data from social media sites to extract sentiments out of them and keep record of those sentiments with the information of the users who stated those sentiments in order to be used later. Finally it does data mining with the extracted sentiments so that it can be used in product profiling, trend analysis and forecasting. Implementation machine learning techniques, it could achieve more accurate results after building classifiers training on large labelled data sets but still there are some issues of processing natural language. Finally, using the sentiment scores for sentiments regarding particular product or service with the user's information, it could successfully profile the products, analyse trends and forecasting. So, as overall, the system is capable of saying that how a set of people of a particular age range, a particular area with a particular profession think about a particular product or service and how it will change it the future which are most useful information when it comes to business world.

ACKNOWLEDGMENT

We would like to thank the staff members at the Faculty of Information Technology for providing valuable insights and comments and helping us in many ways.

REFERENCES

- [1] David Osama and Francesco Murad, "Research challenge on Opinion Mining and Sentiment Analysis"
- [2] Maura Conway, Lisa McKinney, Neil O'Hare, Alan F. Seaton, Adam Birmingham, "Combining Social Network Analysis and Sentiment to Explore the Potential for Online Radicalisation," Centre for Sensor Web Technologies and School of Law and Government.
- [3] Adobe® Social Analytics, powered by Omniture®.
- [4] Brand watch. [Online]. <http://www.brandwatch.com/>
- [5] Sentiment 140.[Online]. <http://www.sentiment140.com>
- [6] Maurizio S. Andrea E., "Determining the Semantic Orientation of Terms through," October 31–November 5 2005.





Edit with WPS Office