

# **OPINION MINING ON BRAND AIMIT USING SUPPORT VECTOR MACHINE**

Neha<sup>1</sup>, Neha K.S<sup>2</sup> and Prof Santhosh Rebello<sup>3</sup>

**Abstract-** Opinion mining is a broad area that focuses on extracting information about people's opinion on a particular organisation, product or services. This area concentrates on performing analysis on this extracted data to help organisations or a particular business to bring in new ideas or to bring about changes in their process. There are a number of techniques to perform this analysis. One such technique is machine learning technique. In this survey, we apply one of the machine learning technique i.e. Support Vector Machine which is an efficient technique for mining of extracted information. This survey is carried out to gather opinion from Students of a College to analyse the results to help the institute bring changes or new ideas through the opinion of students.

**Keywords** –feedback, Opinion mining, sentiment analysis, Machine learning, SVM.

## **I. INTRODUCTION**

What other people think” has always been an important piece of information for most of us during the decision-making process. Opinions are central to almost all human activities and are key influencers of our behaviours. Opinions and its related concepts such as sentiments, evaluations, attitudes, and emotions are the subjects of study of sentiment analysis and opinion mining. Sentiment analysis, also called opinion mining, is the field of study that analyses people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organisations, individuals, issues, events, topics, and their attributes. It represents a large problem space. While in industry, the term sentiment analysis is more commonly used, but in academia, both sentiment analysis and opinion mining are frequently employed. This field has become a very active research area. There are several reasons for this. First, it has a wide arrange of applications, almost in every domain. The industry surrounding sentiment analysis has also flourished due to the proliferation of commercial applications. This provides a strong motivation for research. Second, it offers many challenging research problems, which had never been studied before. Third, for the first time in human history, we now have a huge volume of opinionated data in the social media on the Web. Without this data, a lot of research would not have been possible.

Sentiment analysis is a NLP problem. It touches every aspect of NLP. However, it is also useful to realise that sentiment analysis is a highly restricted NLP problem because the system does not need to fully understand the semantics of each sentence or document but only needs to understand some aspects of it, i.e., positive or negative sentiments and their target entities or topics. There was little research before the year 2000 in either NLP or in linguistics. Part of the reason is that before then there was little opinion text available in digital forms. Since the year 2000, the field has grown rapidly to become one of the most active research areas in NLP. It is also widely researched in data mining, Web mining, and information retrieval.

This is a survey carried out to gather opinion about individuals from an institute which is done using one of the machine learning algorithms called Support vector machine. This survey is done to get feedback or opinion about an organisation from an individual to know where exactly changes or new ideas can be obtained, regarding this particular organisations.

---

<sup>1</sup> *Department of Information Technology AIMIT, Mangalore, Karnataka India*

<sup>2</sup> *Department of Information Technology AIMIT, Mangalore, Karnataka India*

<sup>3</sup> *AIMIT, Mangalore, Karnataka India*

## II. RELATED WORK

[1] proposed that sentiment analysis which is also called opinion mining is the field of study that analyses people's opinion. However, they are now under the umbrella of sentiment analysis or opinion mining. While in industry, the term sentiment analysis is more commonly used, but in academia, both sentiment analysis and opinion mining are frequently employed. They basically represent the same field of study. [2] Here it studies the problem about how sentiment analysis can be applied to extract the opinion of an individual with the explosive growth of social media on the web. It states that sentiment analysis is a popular research problem and highly challenging as a NLP research problem. It is also highly challenging as a NLP research topic. [3] Since the year of 2000, this field is rapidly growing in the field of Natural language processing. It is also widely researched in data mining, web mining and information retrieval.

JayashriKhairnar, MayuraKinikar, [2] proposed Machine Learning as to optimise the performance of a system for developing an algorithm by using different data sets. It provides a solution by learning the model from the data sets and classifying the unseen data. Data with higher dimensions makes the tasks complex thus Feature Selection is used to map the input data which reduces the dimensionality which helps in making the remaining tasks easier. Machine learning has been now an efficient technique in opinion mining with various algorithms being implemented.

Bo Pang, Lillian Lee and ShivakumarVaithyanatham, [4] The different techniques under machine learning can be Naïve Bayes, Maximum entropy classification and support vector Machine. The techniques used in these three algorithms vary but all three can be effectively used in opinion mining. Naïve Bayes is used with classes having a problem in which the features are highly dependent. [5] Naive Bayes is considered when the input data is large and it is constructed using Bayes Theorem. Naive Bayes works well when the feature space is not very large but SVM is better for large feature space. Maximum entropy is a proven effective in NLP applications. [6] Both Naive Bayes and Maximum Entropy are much better as the feature space is been reduced. When Compared Maximum Entropy is better and performs well in the overall performance.

K.P Sonam, R. Loganathan, V. Ajay [7] Support vector machine is the best-known example for machine learning techniques. SVMs hold records in performance benchmark for hand written digit recognition, text categorization and information retrieval. [8] The working of SVM is by achieving a maximum margin hyperplane and separating each point from the input space into two separate classes and the hyperplane with the highest margin is chosen. SVM finds boundary to separate the cluster of data. [9] Computation is performed using mathematical formulas on the dataset to separate it into different classes. [10] The input is taken as input space and non-linear mapping of data into the higher dimension called feature space is done, using a kernel method.

## III. WORK AREA

This is a survey which was conducted in St Aloysius institute of management and information technology. A total of 179 student's opinion from the IT department, was gathered from this survey.

Table -1 Survey Results

Label	Question	Response		
1.	When you meet students who have taken a similar programme at other Colleges/Universities do you feel that your programme is?	Superior 104 0.581005587	Equal 66 0.368715084	Inferior 9 0.05027933
2.	After leaving AIMIT how will you talk about it?	Proudly 151 0.843575419	indifferently 27 0.150837989	Disparagingly 1 0.005586592

The above Table-1 was generated through Excel sheet out of which future analysis will be carried out using one of the Machine Learning Algorithms called Support Vector Machine. Next section we will see the implementation of SVM.

## IV. SUPPORT VECTOR MACHINE

SVM belongs to the class of Supervised Learning algorithms in which the learning machine is given a set of examples (input) with the associated labels (output values). [3] Support Vector Machine is stated as A New Avatar of

Kernel Methods. SVM formulations overcome some of the elementary kernel methods where it examined the entire database, which required RAM to store the entire data set and computation process gets slow.

SVMs construct a hyperplane that separates two classes and the algorithm tries to achieve maximum separation between the classes shown in Figure 1.

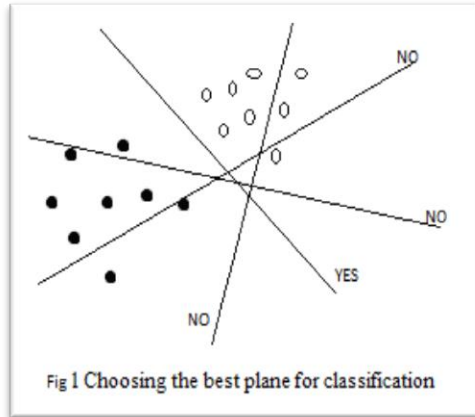


Figure 1. Choosing the best plane for classification

Separating the classes with a large margin minimises the expected generalisation error. The best classifier is one which achieves maximum separation margin between the classes. The two planes parallel to the classifier and which pass through one or more points in the data set are called bounding planes. The distance between these planes is called the ‘margin and SVM learning means finding a central hyperplane which maximises this margin. Shown in Figure 2.

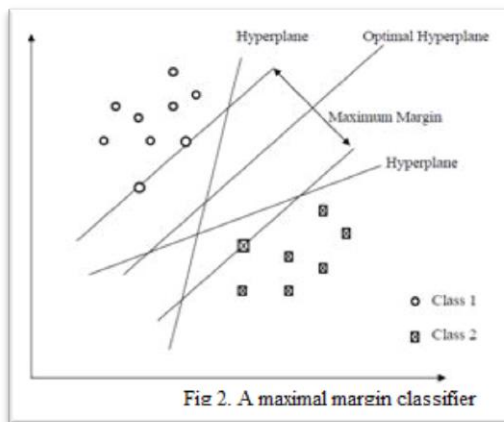


Figure 2. A maximal margin Classifier

**A. Application Procedure–**

Non-Linear mapping of original space of data points (input space) into some higher dimensional space called ‘feature space’, F.

From two-dimension input point (x1,x2) , a three-dimensional point (x12,x22,√2x1x2) are derived from non-linear mapping into feature space.

$$t1 = x12 \quad t2 = \sqrt{2}x1x2 \quad t3 = x22 \quad (1)$$

Finding a hyperplane with the maximum margin.

$$f(x) = w1t1 + w2t2 + w3t3 \\ =w1x12 + w2 \sqrt{2}x1x2 + w3x22 \quad (2)$$

## V.RESULTS

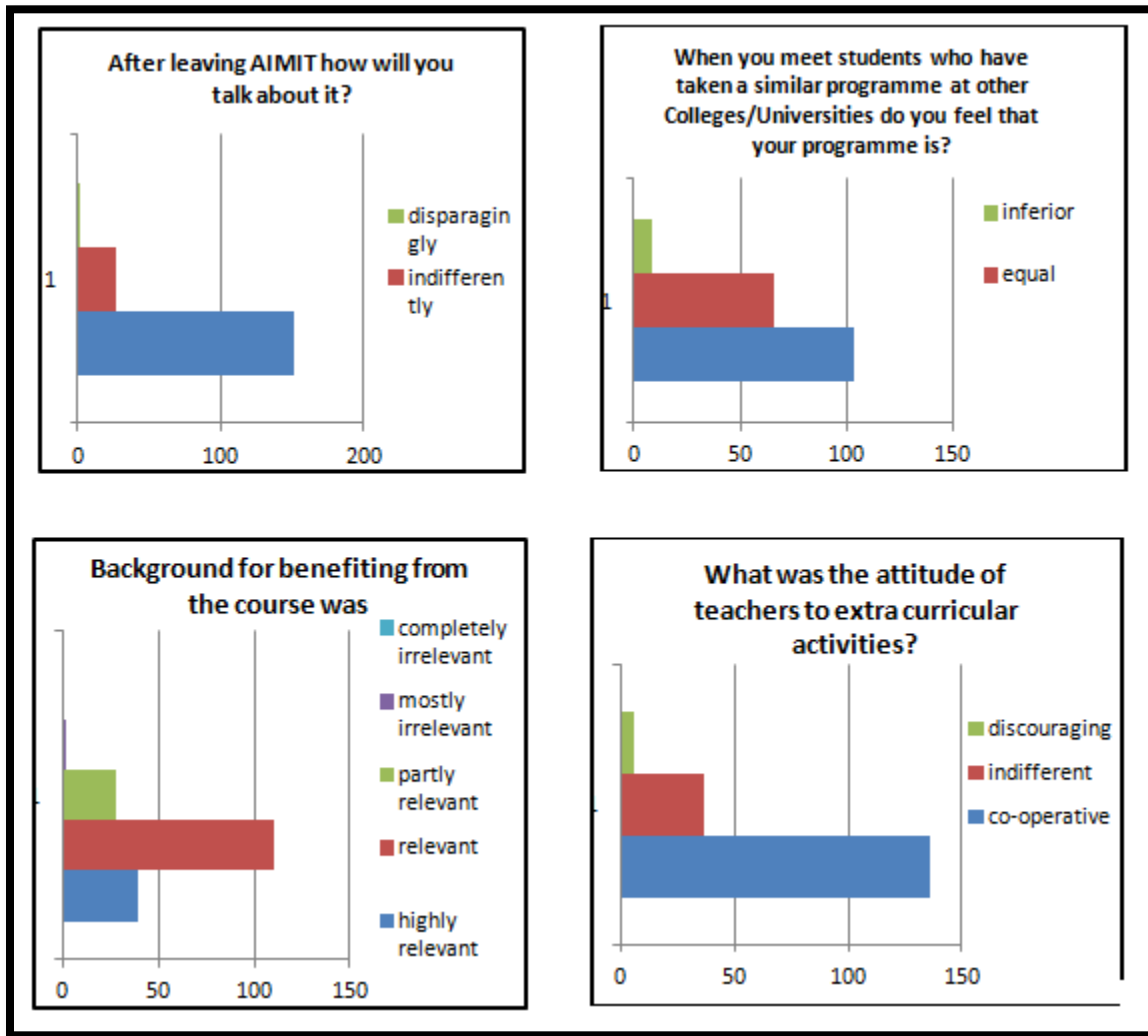


Figure 3. Survey Graphs

Figure 3 shows the results in the form of graphs which are generated using the excel sheet. We will further work on how to generate these graphs automatically in a programmatic way using Support Vector Machine algorithm.

**VI.FUTURE WORK**

In this survey, we have analysed the opinions of each individual and extracted graphs which give a manual description about these graphs. This survey is an ongoing process where we concentrate on generating automatic graphs for analysis through programming by using the machine learning technique i.e., Support vector machine which will help us to generate the results automatically. Doing this survey in a technical way will help us to generate results in an efficient way, which will be helpful for this institute to bring in changes or new ideas. It will also help the institute to know where exactly they stand and what exactly they are lagging behind in. This overall survey is carried out for analysis of opinion given by each individual to help the institute to know their strengths and weaknesses.

The Automated system can go through huge quantities of data and perform efficient analysis compared to manual work. Therefore, carrying out this survey in technical aspect is better compared to the manual aspect.

## VII.CONCLUSION

Determining reviews or feedback is an important research topic. It is especially useful for products services and organisation that have a large number of opinions. The field of sentiment analysis or opinion mining performs surveys on the current state-of-the-art. Due to many challenging research problems and a wide variety of practical applications, the research in the field has been very active in recent years. It has spread from computer science to management science as opinions about products or organisations are closely related to profits or to bring in changes respectively. For applications, a completely automated and accurate 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 the fully automated solution, as time goes by we can push more and more towards automation. The existing techniques for dealing with analysis provides an efficient way to solve or deal problems in this area.

## REFERENCES

- [1] Bing Liu. , “Sentiment Analysis and Opinion Mining”, April 22, 2012 ([liub@cs.uic.edu](mailto:liub@cs.uic.edu)).
- [2] JayashriKhairnar, MayuraKinikar, ”Machine Learning Algorithms for opinion mining and sentiment classification, International journal of Scientific and Research Publications, June 2013, ISSN 2250-3153.
- [3] K.P Sonam, R. Loganathan, V. Ajay, “Machine Learning with SVM and other kernel Methods”.
- [4] Bo Pang, Lillian Lee and ShivakumarVaithyanatham “Thumbs Up? Sentiment classification using machine learning techniques”.
- [5] M. Daiyan, Dr S.K. Tiwari, A. Alam, “To Classify Opinion of Different Domain Using Machine Learning Techniques”. International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, ISO 9001:2008
- [6] Shalom. M. Weiss “Text Mining, Predictive Methods for Analysing Unstructured Information. ”.
- [7] Bo Pang, Lillian Lee2, “Opinion mining and sentiment analysis” , Foundations and Trends in Information Retrieval Vol. 2, No 1-2 (2008) 1–135.
- [8] S. ChandraKala1 and C. Sindhu, “OPINION MINING AND SENTIMENT CLASSIFICATION: A SURVEY”.
- [9] Preety, Sunny Dahiya, “SENTIMENT ANALYSIS USING SVM AND NAÏVE BAYES ALGORITHM”, International Journal of Computer Science and Mobile Computing, ISSN 2320–088X
- [10] Rohini S. Rahate PG, Emmanuel, “Feature Selection for Sentiment Analysis by using SVM” International Journal of Computer Applications (0975 – 8887).