

# Sentiment Analysis on Verbatim Responses for Understanding Students' Living and Learning Experience

Santhi P.<sup>1\*</sup> and Kabilamoorthy R.<sup>2</sup>

<sup>1</sup>Associate Professor, Department of CSE, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India. Email: [santhip.cse@mkce.ac.in](mailto:santhip.cse@mkce.ac.in)

<sup>2</sup>UG Student, Department of CSE, M.Kumarasamy College of Engineering, Karur, Tamil Nadu, India.

\*Corresponding Author

**Abstract:** Sentimental Analysis the computational study of sentiments, people's opinions and emotions expressed in written language. It is very popular because of the wide range of applications it possesses. Opinions are the key influences of our behaviours and decision-making. Opinion mining helps in taking the right decisions as the analyzed data reports give the right emotion on products, companies, services, public personalities, governments, etc. Students' opinions on their institutions, teachers, and trainers are collected nowadays. But the collected opinions are not useful unless it is analyzed and the measures according to it are taken. The conventional method is followed by analyzing it manually. Yet it could not be the consistent method for time and resource saving. Hence, we go for the computational method for opinion mining using Machine Learning techniques for extracting the emotions of the students. Machine Learning algorithms have come a long way, with Naive Bayes, Support Vector Machine and Maximum Entropy is the feature used in research. Sentiment classification by different categories involving the sentiments is the topic of research. This paper presents the survey on Sentimental Analysis on students' feedback to extract the emotions of the students with the text feedback.

**Keywords:** Sentimental Analysis (SA), Opinion Mining, Machine Learning (ML), Natural Language Processing (NLP), Student Feedback Analysis

## I. INTRODUCTION

Sentiment analysis is the process of using Natural Language Processing, text analysis and statistics to investigate the opinion of people. It provides the emotions and opinions of the people on a certain topic or an incident that is being held.

Modern world has a great influence of opinions over the decision-making. It is very important to know what people are thinking before making a decision. A marketer wants to have a clear idea about the opinion of their brands over the media. Considering a case where the feedbacks are collected from the students of an institution about a class, or an event, or any other facilities on the premises. It will help the institutions understand the mindset of the students and get their overall opinion about a class or any event that has been organized.

The feedback system helps the organization to provide a better experience as the faults or any other drawbacks on the previously organized events are collected and the shortcomings are reviewed. Hence, it will provide a chance to improve the quality of the events organized by an organization. There are various machine learning techniques to get emotions from words. Without Machine Learning the Lexicon-based Approach can be used for analyzing the text. Using Machine Learning, the sentimental Analysis can be done using Supervised Learning Techniques; it includes Linear Classifiers and Probabilistic Classifiers [8].

The Linear Classifiers has a lot of algorithms such as the Natural Language Processing (NLP) Algorithm which offers an advanced analysis of text using state tracking and recognition of languages. The Probabilistic Classifier includes the Naive Bayes algorithm that uses the metrics from sklearn libraries [10].

The visualization of the analyzed data can be done in various ways. The most effective way of visualization and representation is the way that is more understandable and gives the right insight to the user. PowerBI by Microsoft provides the tools to work on the analyzed data to create dashboards to view reports and insights of the processed data [9].

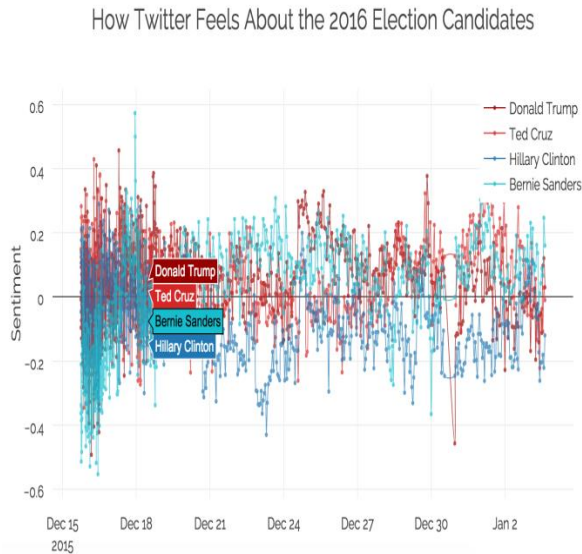


Fig.1: Ref: [towardsdatascience.com/samira-munir](https://towardsdatascience.com/samira-munir)

## II. LITERATURE SURVEY

### A. Natural Language Processing

Natural Language Processing has its origin from the early 1950s which was named as Symbolic NLP where the task is to automate the analysis and production of human language. Up to the 1980s the NLP systems were based on hand-written a rule which was named Statistical NLP. With the steady increase in computational power and the growth of the web, language processing relied on unsupervised and semi-supervised learning algorithms. The NLP algorithm in Machine Learning involves analyzing a set of examples like a book, a large corpus, or a collection of sentences into a statistical inference which is named Neural NLP at present [1]. Rather than coding a huge set of rules, NLP can count on machine learning to automatically check these rules. By using NLP, researchers can group and structure data to perform activities like object entity recognition, sentiment analysis, translation, text to speech, and many more. Mortal language is rarely exact or directly spoken. Understanding such language does not only about understand the words but also how they connect to form meaningful sentences [2]. In sentiment analysis, NLP automatically finds out the psychological tone of informal conversations. Such analysis can be used by companies for the product popularity process, brand reputations, understanding their customers and employees, etc. The Common NLP tasks are subdivided into categories for better understanding such as Text and Speech Processing, Morphological analysis, Syntactic analysis, Lexical semantics, Relational semantics, and Discourse analysis. The Higher-level NLP applications include dialogue management, grammatical error correction, machine translation, natural language generation, natural language understanding, and question answering [3].

### B. Naive Bayes

The Naive Bayes Algorithm is based on Bayes' Theorem in which every feature that is been identified is independent of each other. It is a probabilistic algorithm used to classify various objects. The main advantage is that it doesn't require much training data and deals with both discrete and continuous data. It is faster than sophisticated algorithms and is used for real-time predictions. It is named naïve because; it passes over the prior distribution of features. [4] In Sentiment Analysis, the algorithm is implemented by performing 6 steps such as definition, pre-processing, compilation, installation, evaluation, and prediction. The NB classifier is used for defining the model and cannot be used in regression. The Noises such as punctuation marks, stop-words, and URLs are removed, decoding data and converting them to lowercase is done in the pre-processing phase. The Compilation is done by importing the predefined classes and modules [5]. In the compilation phase, the vectorizer is converted to the transformer and then to classifiers which are easier to deal with. The Model can be instantiated by fitting the dataset into the multinomialNB and then split into train and test sub-datasets.

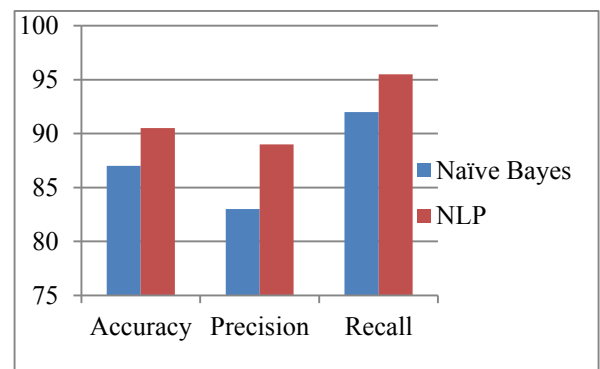


Fig.2: Naive Bayes Vs NLP

The two kinds of training datasets that is been used, Bernoulli Naive Bayes sets binary values to each feature and Multinomial Naive Bayes are used for labeling in the case of discrete data which is represented in the terms of their occurrences. The Evaluation is performed by setting up the metrics from sklearn library and calculating the confusion matrix for better understanding [6]. After fine-tuning the model, the accuracy is observed to be above 60%.

### C. Power BI

Power Bi is a tool for business services offered by Microsoft. It provides a simple interface that helps the end-users to create reports and dashboards on their own. The data representation is made easy by Power Bi as it has a user-friendly environment to work on the data which is simple enough to be understood by any user without any specialization in that [7].

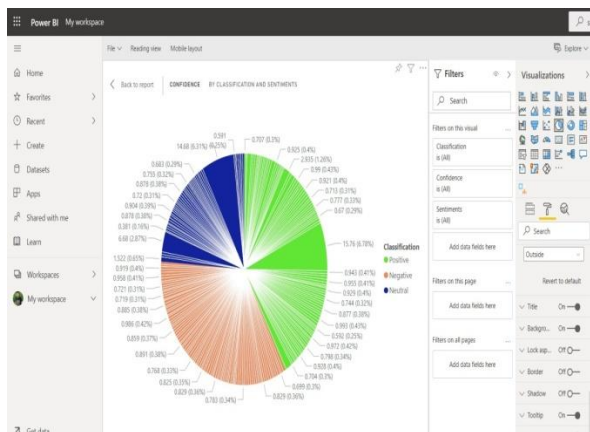


Fig.3: Power BI Implementation

### III. CONCLUSION

By summarizing, Machine Learning provides a promising method to analyze the data that has been collected from the students to get their emotions over a particular topic. This may lead to understanding the mentality of the students and thereby improving their overall living and learning experience in an institution. It is better to take measures according to the need of people. Hence, it helps in developing the quality of learning and living inside the campus of an institution.

### REFERENCES

- [1] Jason Brownlee, “Develop Deep Learning Models for Natural Language in Python”, in *Machine Learning Mastery, Edition: v1.1*, 2017.
- [2] R. Socher, A. Perelygin, “Recursive deep models for semantic compositionality over a sentiment tree bank”, *Empirical Methods of Natural Language Processing*, October 2013.
- [3] Xin Chen, Mihaela and Krishna P.C, “Mining Social Media Data for Understanding Students Learning Experiences”, *IEEE Transactions on Learning Technologies*, 2014.
- [4] Manjunath B Talawar and Prasad.M. R, “Analyzing Social-Media Data in Educational Sectors Using Data Mining Techniques”, *International Journal of Computer Science and Information Technologies*, Vol. 7 (4), 2016.
- [5] Eman M.G. Younis, “Sentiment Analysis and Text Mining for Social Media Microblogs using Open Source Tools: An Empirical Study”, *International Journal of Computer Applications*, 2015.
- [6] Aaquib Multani, “Sentiment Analysis for Understanding Students’ Learning Experiences: A survey Paper,

*International Journal of Scientific Research & Engineering Trends*, Volume 5, Mar-Apr-2019.

- [7] Shanshan Yi and Xiaofeng Liu, “Machine Learning based customer sentiment analysis for recommending shoppers based on reviews”, *Springer Nature in Complex & Intelligent Systems Complex & Intelligent Systems*, Volume 6, pp 621-634, 2020.
- [8] Jianqiang Z, Xiaolin G, Xuejun Z, “Deep convolution neural networks for twitter sentiment analysis”, *IEEE Access* 6:23253–23260, 2018.
- [9] Tan S et al, “Interpreting the public sentiment variations on twitter”, *IEEE transaction of knowledge Data Engineering*, 2014.
- [10] Quratulain Rajput, Sajjad Haider, Zarmeen Nasim, “Sentiment analysis of student feedback using machine learning and lexicon based approaches”. In *Proceedings of 5<sup>th</sup> ICRIIS*, 2017.