

Systematic Study of Sentiment Analysis for Customer Reviews

¹Gopika Vijayan, ²Liz George

¹PG Scholar, ²Asst.Professor of St. Joseph's College Of Engineering And Technology Palai,

¹Department of Computer Applications,

¹St. Joseph's College Of Engineering And Technology, Palai, India

Abstract: Sentiment analysis, often known as opinion mining, is a technique for natural language processing (NLP) that determines the emotional undertone of a body of text. This is a typical strategy used by businesses to gauge consumer opinion on a certain good, service, or concept. As more people used online consumer reviews to assist them in making decisions, they became more popular. Hence, we may classify customer reviews as either positive or negative using sentiment analysis. This particular paper surveys different research papers that uses Sentiment Analysis.

Index Terms: Sentiment Analysis, Machine Learning, Opinion Mining, Natural Language Processing, Consumer Reviews.

I. INTRODUCTION

Natural Language Processing is utilized in sentiment analysis to ascertain whether an individual's feelings are pleased, negative, or neutral.

It helps to identify the genuine emotion behind the text or interaction. A machine learning tool is sentiment analysis. This analysis helps the business to understand the success and failure of their product and also the needs of the customer. So, by understanding the sentiments of customer the company can develop the product based on customer's needs. Sentiment analysis comes in three types. They are Knowledge-Based, Statistical and Hybrid. In keyword-based sentiment analysis, the text is categorized based on words that indicate emotion. In statistical sentiment analysis, numerous machine learning techniques and deep learning techniques are utilized for classification. Also, Hybrid is a fusion of both knowledge-based and statistical, Hybrid classification provides more accurate results. There is no need for someone to consult ten papers in order to acquire an overview of this topic; they can receive such overview by consulting this paper.

II. LITERATURE REVIEW

Sentiment Analysis of Customer Product Reviews Using Machine Learning ^[1]: In this four Lakh Mobile Phone Reviews taken from Amazon website is considered as data. Utilizing sentiment analysis, reviews are segregated into favorable and unfavorable categories. The three algorithms used in this paper are SVM, Naive Bayes, and Decision Tree. Naive Bayes accuracy is 66.95, SVM accuracy is 81.77, and Decision Tree accuracy is 74.75. SVM is the most accurate. A Proposed System for Understanding The Consumer Opinion of a Product Using Sentiment Analysis ^[2]: Reviews from Amazon customers are used as data. Random Forest, SVM, and Logistic Regression are the algorithms used. In contrast to random forest, logistic regression and SVM have a better F1 score of 0.84, making them the best. Twitter Sentiment Analysis Using Ensemble Techniques ^[3]: Twitter sentiment analysis is the main focus of this study. In order to comprehend how consumers are reacting to the new product, tweets are taken into consideration as data. In this paper, many ensemble techniques are used. Both Random Forest + XGBoost and Random Forest + AdaBoost fall under this category. Random Forest + XG Boost accuracy is 93.2, whereas Random Forest + Ada Boost accuracy is 95.7. The best accuracy is thought to be achieved with Random Forest + Ada Boost. Sentiment Analysis for Customer Reviews Using Hybrid Approach ^[4]: Both the customer's textual and audio content was taken into consideration for this study. As a result, this hybrid approach's accuracy will be higher because it takes both speech and text into account. This paper uses SVM as its algorithm. Here, the first attempt at speech emotion classification has an accuracy of 57.1. Second, the text emotion classification has been completed with a 76 percent accuracy rate. Next, a hybrid strategy that takes into account both text and audio is used to achieve an accuracy of 90%. Among them, hybrid approach accuracy is thought to be the best. Sentiment Analysis of Customer Feedback and Reviews for Airline Services ^[5]: The dataset of airline reviews used in this study was obtained from Kaggle. Here, they employed the BERT text classification approaches accompanied by a multitude of machine learning techniques, including AdaBoost, Decision Tree, Logistic Regression, SVM, KNN, and Random Forest. Random Forest has the best accuracy of these machine learning techniques, with an accuracy of 77. Google developed the BERT algorithm, which only requires a small amount of training data. After BERT classification, they obtained an accuracy of 83. The accuracy of BERT classification is found to be the best when they compare the results to the Random Forest. This leads us to the conclusion that for sentiment analysis, BERT architecture is superior to machine learning techniques. Sentiment Analysis of Product Reviews ^[6]: The intention of this paper is to appraise the comments based on customers. The comments can be positive or negative. The comments will be in sentence format, divided into words, adjectives, and adverbs, and processed using the Parts of Speech system (POST). They created a sample website as a means of gathering data, and from that website they gathered data in the form of feedback. The obtained accuracy will be 90.47. Customer Sentiment Analysis Through Social Media Feedback: A Case Study On Telecommunication Company ^[7]:

Tweets are used as data in this paper. The tweets are from 15 March to 20 April 2022. The algorithms used are Naïve Bayes, SVM and Random Forest. SVM, Random Forest, and Naïve Bayes have accuracy values of 0.8, 0.6, and 0.7, respectively. The best accuracy comes from SVM. Amazon Product Reviews Using Word2vec ^[8]: Here, customer ratings and reviews of products on Amazon are taken into account. It uses the Word2vec model. First, reviews are converted into vector representations, and then

machine learning algorithms are applied to that. Random Forest and LSTM are two machine learning algorithms used. The accuracy obtained with Word2vec and LSTM is 94.63, and the accuracy obtained with Word2vec and Random Forest is 92.62. Out of these, the accuracy that is attained when using LSTM is considered to be the best. Improved Sentiment Analysis to the Movie Reviews using Naive Bayes Classifier ^[9]: Imdb.dataset.csv is the name of the dataset, which was taken from the website kaggle.com. Four key methodologies namely data collection, data extraction, feature extraction, and classification evaluation are utilized in the proposed work. In this paper, the Bag of Words and Naive Bayes models are mainly utilized. The Naive Bayes classifier has the highest accuracy, which is 0.728. Sentiment Analysis of Online Reviews of Customers towards a Health Food Drink (Hfd) Brand ^[10]: This paper's goal is to examine customer reviews of the Pediasure product on Amazon. Here, 210 reviews are analyzed using NVivo software. This analysis found that the percentages of positive, negative, and neutral emotions were 40 percent, 16 percent, and 44 percent, respectively. The neutral has the best percentage of all.

III. APPLICATIONS

- A) *Decision-making assistance*: Making decisions is a crucial aspect of daily life. We can make decisions like "which phone to buy" and "which dress to buy" using opinions gleaned from reviews.
- B) *Commercial Application*: Every business wants to meet the needs of its customers in today's competitive market by developing novel, creative products. Today, surveys of people are a crucial component of businesses' efforts to get feedback from customers and implement desired product improvements. One illustration is the product search on Google.
- C) *Trend analysis and predictions*: By monitoring public sentiment, sentiment analysis enables one to foresee market trends. By analyzing reviews, it also helps the producers make predictions about the success and failure of the films.

IV. METHODOLOGY

- A) *Naïve Bayes*: A supervised learning algorithm for resolving classification problems that is based on the Bayes theorem. Spam filtering, Sentimental analysis, and article classification are a few examples of the Naive Bayes algorithm.
- B) *Decision Tree*: A supervised learning algorithm. Decision trees can be utilised for classification as well as regression problems.
- C) *KNN*: Among the easiest algorithms, based on the supervised learning method, is K-Nearest Neighbor. Since the algorithm is non-parametric, no assumptions about the underlying data are made.
- D) *Random Forest*: It is one of the most common Machine learning algorithm that handles both classification and regression problems.
- E) *Support Vector Machine*: It's also known as SVM. In machine learning, classification problems are addressed by it. The data is transformed using a method known as the kernel trick, and based on these transformations, an ideal boundary between the potential outputs is discovered.
- F) *XGBoost*: It is a distributed gradient boosting library that has been optimized for fast and scalable machine learning model training.
- G) *AdaBoost*: An ensemble method in machine learning is called AdaBoost, also recognized as Adaptive Boosting.
- H) *Logistic Regression*: It is utilized to predict the propensity that a binary (y/n) event will occur.
- I) *BERT Algorithm*: BERT is another name for Bidirectional Encoder Representations from Transformers. It's an NLP model created by Google that only needs a small amount of training data.
- J) *Parts of Speech Tagging*: It means altering a sentence into different forms. - Word list, list of tuples.
- K) *LSTM*: It is also known as long short-term memory networks, and deep learning utilizes it the most.
- L) *NVivo 12*: Unstructured or semi-structured data can be obtained, organized, evaluated, and envisioned using this qualitative data analysis tool.

Table 1: Analysis

Journal Title	Used Algorithm	Year	Accuracy
Sentiment Analysis of Customer Product Reviews Using Machine Learning	Naïve Bayes Support Vector Machine Decision Tree	2017	66.95 81.77 74.75
A Proposed System for Understanding the Consumer Opinion of a Product Using Sentiment Analysis	Logistic regression And SVM	2023	0.84
Twitter Sentiment Analysis Using Ensemble Techniques	Random Forest + XGBoost Random Forest + AdaBoost	2022	93.2 95.7
Sentiment Analysis for Customer's Reviews using Hybrid Approach	Speech emotion Classification Text emotion classification Hybrid approach	2022	57.1 76.90

Sentiment Analysis of Customer Feedback and Reviews for Airline Services	Logistic Regression K Neighbour Support Vector Decision Tree Adaboost BERT Random Forest	2023	0.65 0.67 0.65 0.67 0.72 0.83 0.77
Sentiment Analysis on Online Product Reviews	Parts Of Speech Tagging	2017	90.47
Customer Sentiment Analysis Through Social Media Feedback: A Case Study On Telecommunication Company	Support Vector Machine Random Forest Naïve Bayes	2022	0.8 0.6 0.7
Amazon Product Reviews Using Word2vec	Random Forest Long Short-term Memory	2022	92.62 94.63
Improved Sentimental Analysis to the Movie Reviews using Naive Bayes Classifier	Naïve Bayes	2022	0.728
Sentiment Analysis Of Online Reviews Of Customers Towards A Health Food Drink (Hfd) Brand	auto code feature (sentence coding) in NVivo 12	2022	Positive :40% Negative : 16% Neutral : 44%

V. CONCLUSION

To conclude, the study of ten papers has been successfully completed. Sentiment analysis is becoming more popular nowadays as customers are expressing their feelings more openly through reviews. Also, through sentiment analysis we can identify the needs of customer and to analyse whether or not the product that companies are manufacturing will have a demand in the market. Research in sentiment analysis is still an open field. The two most frequently employed machine learning techniques in sentiment analysis are Support vector machines and Naive Bayes.

VI. ACKNOWLEDGMENT

First and foremost, I give all glory, honor and praise to God Almighty who gave me wisdom and enabled me to complete this work successfully.

I also want to thank my parents from the bottom of my heart for their encouragement and help with this work and with all of my other endeavours.

I am incredibly grateful to Dr. V. P. Devasia, Principal of SJ CET in Palai, for letting me use all of the facilities there as well as for his support. Words cannot adequately express how grateful I am.

Sincere gratitude goes out to Mr. Anish Augustine, HOD in charge, Department of MCA, SJ CET, Palai, who has served as a constant source of inspiration and whose invaluable support and assistance have made this work possible.

I owe a particular debt of gratitude to Ms. Liz George, Asst. Professor, Department of Computer Science and Applications, SJ CET, Palai, for all the necessary help and support that she has extended to me. Her valuable suggestions, corrections, and sincere efforts to accomplish this work even under a tight time schedule were crucial to the successful completion of this work.

I extend my sincere thanks to all of our teachers and non-teaching staff at SJ CET, Palai, for the knowledge they have imparted to me over the last three years.

Additionally, I would like to thank all of my friends for their encouragement, advice, and support.

REFERENCES

1. Zeenia Singla, Sukhchandan Randhawa, Sushma Jain, "Sentiment Analysis of Customer Product Reviews Using Machine Learning", 2017.
2. Eesha Martis, Rutuja Deo, Sejal Rastogi, Keshav Chhaparia, Ameyaa Biwalkar, "A Proposed System for Understanding the Consumer Opinion of a Product Using Sentiment Analysis", 2023.
3. Dileep Sai Ellanki, Navya Alam, Riya Dalal, Shrey Modi, "Twitter Sentiment Analysis Using Ensemble Techniques", 2022.
4. Sharma Rohith, "Sentiment Analysis for Customer's Reviews using Hybrid Approach", 2022.
5. Aksh Patel, Parita Oza, Smitha Agarwal, "Sentiment Analysis of Customer Feedback and Reviews for Airline Services ", 2023.
6. Raheesa Safrin, K.R.Sharmila, T.S.Shri Subangi, E.A.Vimal, "Sentiment Analysis on Online Product Reviews ", 2017.
7. Siti Nur Syamimi Mat Zain, Nor Azuana Ramli, Rose Adzreen Adnan, "Customer Sentiment Analysis through Social Media Feedback: A Case Study on Telecommunication Company ", 2022.
8. Shivani Joshi, Jiten Prasad, Divakar Sinde, "Amazon Product Reviews Using Word2vec ", 2022.
9. Gowri S, Surendran R, Divya Bharathi M ,Jabez J , "Improved Sentimental Analysis to the Movie Reviews using Naive Bayes Classifier ", 2022.
10. Dr. Payal Sanan, Tarun Kanade, Dr. Radhakrishna Batule, "Sentiment Analysis Of Online Reviews Of Customers Towards A Health Food Drink (Hfd) Brand ", 2022.