Sentiment Analysis on Customer Reviews of E-commerce Site

THESIS

SUBMITTED BY
SOUNDARYA MAMADAPUR

APPONIX ACADEMY

CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER-1 INTRODUCTION

The widespread use of the Internet and e-commerce is transforming global society and its way of life . When purchasing a product in the past, advertisements and recommendations from friends were important sources of information. There were only a few suggestions for comparing comparable goods from other brands. These days, with the growth of the e-commerce industry, they are providing a wider range of products. In the form of product reviews, the e-commerce websites also ask their clients to share their experiences with the items they purchased.

Online reviews, which have supplanted traditional "word-of-mouth" due to the explosive rise of electronic commerce, have a significant impact on customer purchasing behavior and product sales. Consumers use reviews as a platform to create trust and make educated purchases by evaluating the experiences of past buyers. From the perspective of the manufacturer, valuable internet reviews are essential for identifying customer needs when developing new products or making improvements to existing ones. Manufacturers can meet consumer demands in the target market by gathering pertinent internet reviews. Additionally, manufacturers receive knowledge of the competitive industry and current trends that affect their marketing choices. Reviewers have a variety of alternatives when it comes to publishing their evaluations on retail websites like Flipkart.com. For example, the user can rate the product using open-ended customer-authored comments or numerical stars, which typically range from 1 to 5. Online reviews are thought to boost a website's trustworthiness, draw customers, improve the hit ratio, and extend visitors' stay on the page. The user reviews on e-commerce sites are the sole reason for the growth of discovery platforms. Trustworthy customer reviews increase the base of customers and foster a sense of confidence among inexperienced users. Reviews, whether good or bad, benefit manufacturers and customers. E-commerce has advanced significantly thanks to big data commerce. It has made it possible for both major industries and consumers to make more intelligent decisions. One such paradigm that can be used to make more successful selections is the online reviews seen on e- commerce behemoths like Amazon and Flipkart. They benefit not just the customers but also the companies who make the products.

Internet evaluations have the power to give consumers information about a product's quality, functionality, and suggestions, giving prospective customers a comprehensive understanding of the item. One such untapped potential is the ability of online reviews to help manufacturers understand customer requirements through the analysis of beneficial reviews. Reviews, both good and bad, are important in determining client needs and obtaining feedback about the product more quickly.

The advantages of using e-commerce websites to do business online, such as loyalty, faster delivery, simple setup, time savings, cost-effectiveness, and flexibility, are depicted in figure 1.

Figure 1.1 Advantages of Using E-commerce

India has seen a huge increase in online purchasing. Millions of Indians now shop online because of increased Internet availability, which has made it a highly convenient method to do business. Some of them even exclusively rely on internet purchasing for their everyday need. Online shopping has several advantages, including You can select from thousands of goods. Your house receives delivery of the merchandise. A component of electronic commerce, online shopping, sometimes known as electronic shopping, enables customers to directly purchase goods or services from sellers online. In 1979, Michael Aldrich created the concept of internet commerce. With more people becoming internet literate in India, the potential for online marketing is growing. Other names for online shopping include virtual stores, e-web stores, e-shops, e-stores, Internet shops, web shops, and online stores. Internet shopping, also known as business-to-consumer (B2C) internet shopping, is analogous to physically purchasing goods or services from a brick-and-mortar store or shopping mall.

In India, the most well-known e-commerce platforms include eBay, Amazon.com, Snapdeal, Alibaba.com, and Myntra.com. A customer review is an evaluation of a good or service written by a person who has either bought, used, or otherwise dealt with the good or service. On websites for online shopping and electronic commerce, customer reviews are one type of client feedback. There are other specialized review websites, some of which use user reviews in place of or in addition to those written by experts. Other users may rate the reviews themselves in terms of accuracy or usefulness.

Determining the emotions expressed in a text is the aim of sentiment analysis, which has become one of the most active study fields at the interface of computer science and linguistics [5]. A favorable negative, or neutral assessment expressed through textual material, such as an online movie, book, or product review, is referred to as sentiment. It is now a standard feature of most social networking sites and is used by political analysts, marketers, and businesspeople. Sentiment analysis interprets the words used and assesses how language is employed to represent human emotions. Social networking sites have seen a surge in usage recently due to their popularity, drawing big numbers of users and producing vast amounts of data. However, to analyze such vast, intricate, and dynamic social media data, suitable approaches are needed. Sentiment analysis also emerges as a way to interpret human attitudes and sentiments to manage this enormous amount of data and draw conclusions from it. It makes a major contribution to linked concerns and the decision-making process. When making decisions about things like buying a product, investing, or investigating new concepts, for instance, customers are always looking to gain insight from the experiences of others.

These days, a lot of reviews can be found on social media platforms that can be used for sentiment analysis to identify the polarity of reviews [6]. This allows users to make decisions more quickly by identifying if a review is favorable or negative without having to read the full context. Numerous industries, including social media, marketing, customer service, clinical medicine, and healthcare, frequently employ sentiment analysis. The use of persuasive technology in mental health is one of the key domains for sentiment analysis in health. With so many uses, sentiment analysis is still a rapidly expanding field of study. Numerous researches have examined and applied factors that impact the accuracy of sentiment analysis processing.

In addition to presenting a comparative study of the most current works in this field, this work will explore the sentiment analysis literature to look at all the connected topics. It presents the framework and social networking background and shows how sentiment analysis fits into it by outlining the various steps and associated problems based on earlier studies.

Reviews on online services or goods for e-commerce can be found in abundance on the internet. Before making a payment to a service provider, customers always like to check evaluations. But in the fast-paced world of today, it is almost impossible to read every review. Every review has the potential to offer fresh insights into a product or its features. Thus, there's a chance you won't see any significant customer reviews. The polarity of the review, or whether it is positive, negative, or neutral, must be determined. Sentiment analysis will help us determine the reviews' polarity. Visuals can outperform written format by up to 80%. Thus, visualizing every review will facilitate the consumer's decision-making process. The consumer will make decisions more quickly because they can examine all reviews at a look.

1.1 History of customer reviews

Epinions was one of the first review websites, having been founded in 1999 [6]. Established in 1999, Epinions.com was a broad consumer review website. Shopping.com, which was then known as DealTime, purchased Epinions in 2003. eBay then purchased Shopping.com in 2005. Visitors could browse both recent and historical evaluations about a range of products at Epinions to assist them in making a decision. All community features including the ability to submit new reviews were deactivated on March 25, 2014. Customer comment boxes and customer service help lines were two ways that consumers could evaluate goods and services before the Internet. These techniques are being used today, even though the number of online review sites has increased dramatically.

Funny customer reviews can be found on several well-known e-commerce sites, including Amazon. These are frequently snarky or ironically glowing evaluations of kitschy or unimportant things. Methylated spirits presented in the manner of a wine evaluation is another example. A product might become viral online and garner a lot of fake reviews, which would increase sales.

ONLINE SHOPPING PLAYERS IN INDIA

AMAZON: Amazon was begun by Jeff Bezos in 1994 and the base camp is in Seattle [7]. The first name of the organization was Cadabra.com yet was changed to Amazon since it seemed like a "body". The store is named after the world's second-longest stream. Amazon is a web-based store that sells books, films, games, DVDs, music Discs, PC programming, and so forth things. It is the biggest internet-based retailer at present.

FLIPKART: Flipkart Pvt ltd is an Indian electronic commerce company based in Bengaluru, India Founded by Sachin Bansal and Binny Bansal senior and junior at the Indian Institute of Technology Delhi, and colleagues at Amazon in 2007 and the company's initial stage focused on book sales before expanding into other products such consumer electronics fashion n lifestyle and home care products. Flipkart Private Limited is an Indian e-commerce company established in 2007. It started with a primary focus on online book sales and soon, expanded to lifestyle products, electronics, home essentials, and groceries. Today, Flipkart is the biggest online Indian marketplace competing with the world leader Amazon.

SNAP DEAL: Snap Dael is an Indian online business organization situated in New Delhi, India. The organization was begun by Kunal Bahl, (a Wharton graduate as a feature of the double certificate M&T Designing and Business program at Penn), and Rohit Bansal, (a former student of IIT Delhi) in February 2010. Snap deal at present has 275,000 venders, north of 30 million items, and a range of 6,000 towns and urban communities the nation over.

Myntra: Myntra is an Indian design web-based business organization central command in Bangalore, India. The organization was established in 2007 by Mukesh Bansal alongside Ashutosh Lawania and Vineet Saxena. Myntra sold on-request customized gift things. It predominantly worked on the B2B (business-to-plan of action) during its underlying years.

E-Bay: E-bay is a worldwide web-based business enterprise, working with online C2C and B2B deals. An organization settled in San Jose, California. E-Bay was established by Pierre Omidyar in 1995. Today it has a multibillion-dollar business with tasks in around 30 nations. The organization oversees eBay.com, a web-based sale and shopping site in which individuals and organizations trade a wide assortment of labor and products around the world.

The data used in this research is a combination of product reviews from various e-commerce sites collected from a highly reputed source named Kaggle. Based on the collected ratings, we conducted a sentiment analysis. This research separates and provides a simple breakdown of different product views. Positive, negative, or neutral sentiments can be classified. The study identified the most frequently used words and word pairs as the highlights of the conversations. The study results will suggest that the presence of the most critical negative and positive words in the data set can help to understand the psychological state of the general public.

CHAPTER-2 LITERATURE SURVEY

Sentiment Analysis of reviews was fetched from Amazon. The Naïve Bayes algorithm showed comparatively better results than the Logistic Regression and SentiWordNet technique. The performance of these algorithms was measured by using quality metrics like Recall, Precision, and F-measure. A sentiment dictionary using external textual data was created and different classification models were compared along with a hybrid model. A hybrid model was made by combining a Support Vector Machine (SVM) and Gradient Boosting Decision Tree (GBDT) based on the Stacking approach, which is an ensemble learning approach usually used to combine different learning algorithms to get better performance. The baseline model made for verification compared between SVM, GBDT, and the hybrid model, in which the hybrid model was found to be the most effective.

The proposed method was based on a monolingual word alignment model (WAM) and for additionally best alignment quality, this approach used a Semi-supervised Word Alignment Model (SWAM) with supervision on alignment. The results showed that the true positive rate gradually increases in prediction by the proposed SWAM technique. Experimental results also showed that using the sentimental analysis the system provides the result with higher accuracy than the system without sentimental analysis.

The process of gathering online end-user reviews for the products or services was automated. These reviews were then analyzed based on the sentiments expressed about specific features. Online product reviews from Flipkart (a popular Indian e-commerce website) were the inputs to this project, which the system analyzed to generate results (for reviews) that have been sorted according to various geographic regions. These results were also shared with the manufacturers to look at the sentiments and judge the improvisation and deterioration of the Product.

The trustworthiness of an e-vendor was measured in multiple dimensions like the product, price, and shipping. This helped the customers in making a good purchase decision from the user with a high trust score. This was analyzed by using a multidimensional lexicon and sentiment compensation technique. The accuracy of the proposed method (SenseComp) was compared with manual, sentiment to dimension (S2D) and dimension to sentiment (D2S) methods.

The results of the experiment showed that the sentiment compensation technique increases the accuracy of SenseComp in all dimensions with an overall accuracy of 93.60% as compared to S2D and D2S methods. Several methods were used, including extraction, clustering, and classification for the sentiment analysis. Using the Flipkart product API, one can easily extract reviews, ratings, and other relevant information from the website. The CART and ROCK algorithms were used to classify reviews into positive and negative words from the comments, and finally determine which product has the highest percentage of positive reviews.

Sentiment analysis for Amazon and Flipkart products was done using Naive Bayes, Logistic Regression, Sentiment WordNet, Random Forest, and K-Nearest Neighbor techniques. Addresses and reviews were the concepts. It presented a detailed comparative study of such existing sentiment analysis algorithms and methodologies based on five key parameters. The study also proposed a Product Comment Summarizer and Analyzer (PCSA) system. It was an automatic and generic comment analyzer that could find the polarity of the sentiments and comments very effectively. It summarized the comments and classified them into the pre-defined positive, negative, or neutral classes. It results in evaluating their performance in terms of parameter rating, classifiers, and accuracy.

To gauge the sentiment, the researchers tried to figure out the opinion of a customer through a piece of text. They first took the review and checked if the review was related to the specific product with the help of a Decision tree. They used a Spam dictionary to identify the spam words in the reviews. In Text Mining they apply several algorithms and based on these algorithms, we got the specific results.

A research model was created to determine the impact of certain online purchase determinants on customer satisfaction in the Indian market. A conceptual model developed in the study was also referred to. Various factors like security, information availability, shipping, quality, pricing, and time were the determinants of customer satisfaction. Flipkart.com offers the best prices, good products, and a completely hassle-free shopping experience for our customers. The success of an e-commerce company in India depends upon its awareness, its brand image, its unique & fair policies, its customer relations, etc. Drawbacks included no proper return policy for the product, online payment systems, personal privacy, and personal customer services.

An experimental investigation and a suggested model using the Support Vector Machine (SVM) method have been released. The study examined various datasets of product evaluations to determine the polarity of the reviews—positive or negative—and terms associated with the products, such as good, poor, great, and super hit. The models' performance was evaluated to gauge the Support Vector Machine learning algorithm's correctness. In the end, the Support Vector Machine classification algorithm outperforms the others and achieves high accuracy. The study's conclusions show that a customer's intention to make an online purchase is highly correlated with their age, gender, education, level of security concern, level of technological comfort, and frequency of previous online purchases. Product type, frequency of purchases, and cost all have an impact on consumer purchasing behavior. It is also discovered that the online retailer's return, refund, and shipping policies influence their purchasing decisions. The research methodology improves knowledge of the variables influencing consumers' online purchasing behavior, aids in the profiling of typical Indian online shoppers, and could assist e- marketers in creating more targeted marketing campaigns to boost e-commerce sales.

To better understand the relationships between various aspects and perform sentiment analysis on a data set of smartphone reviews that is beneficial to both designers and consumers, researchers focused on doing sentiment analysis research. Based on statistical data, the top three brands in the market at the moment were Apple, BLU, and Samsung. The brand Samsung got the most favorable reviews. Further observations showed that more expensive products did not always elicit comprehensive evaluations, nor did detailed reviews always result in higher ratings. Nonetheless, higher ratings were noted for expensive products, indicating higher consumer satisfaction levels and superior product quality compared to lower-priced products. It was discovered that the top three brands—Apple, Samsung, and BLU—had a positive emotion orientation along with strong positive sentiments of surprise, delight, trust, and expectation. Data classification is as accurate as SVM accuracy following cross-validation was equivalent to 84.87%.

The research on sentiment analysis was conducted by dividing reviews of mobile phones into two categories: positive and negative sentiment. Three classification models were employed to categorize reviews after the data was balanced with an approximately equal ratio of positive to negative evaluations. A comparison of the three classifiers—Naïve Bayes, SVM, and Decision Tree—shows that SVM has the highest predictive accuracy.

After cross-validation of the accuracy results, SVM achieved the highest accuracy of the three models, at 81.75%. It was noted that the sentiment analysis result's % can be the same. In certain situations, it could be interpreted that the program favors a particular website if just one is suggested, which is unwarranted. Therefore, the decision of the user to purchase from a particular website is their own. When there is a significant discrepancy between the training and testing sets, it has been noted that the model may not be entirely accurate. The accuracy of the Training set was 90.4%, while the Testing set's accuracy was 86.3%, which is a respectable difference, according to the data.

Two sections of a specific user's transaction history were separated to provide recommendations to them in the test data set. The conditional probability model employed the first section, which was regarded as the purchase history, to generate suggestions. The second portion was the item that was purchased after the first and was regarded as the basis for the suggestion. The experiment was conducted under stringent conditions, with only one transaction in the second section. Let's say a user was situated close to a list of coupon bargains that were offered. To suggest the offers that have the best chance of being accepted, the experiment attempted to anticipate the appropriate category of the user's subsequent purchase. According to the study, each company had a distinct sentimental analysis based on the demographic of its target market, and each drew different conclusions from the study using a variety of emotional analytics characteristics derived from Twitter comments. However, a comparison would show that there are variations in millennial customers' behavior depending on the quality of the services they receive, or differences in emotional attachments. According to the Polarity test of Sentimental Analytics, the relationship between Amazon and Flipkart is such that there is a greater degree of trust and joy, but a lower degree of anger, sadness, fear, disgust, and surprise. Therefore, it was deduced that customers' opinions of Amazon's Big Billion Day Sale are more favorable than those of Flipkart's Big Billion Day Sale throughout September through October.

In the analysis, the internet reviews for the masks and sanitizers from Flipkart and Amazon were examined. The star rating, review length, and helpfulness are visual elements that convey content more quickly than words in reviews. However, no research has compared the differences in these aspects across the two online retailers.

It was demonstrated that the distribution of review attributes amongst the reviews from online shops differs significantly. The length, star rating, and helpfulness votes of reviews on the two websites differ. Reviews on Amazon and Flipkart generally differ from one another, and there are also disparities between the brands they sell. Online shoppers' ability to switch between stores is limited by this information asymmetry since the distributions of the review attributes are unknown.

The proposed system's results were personally checked and validated against a few standard evaluations found in Digit and GSMArena. Almost 120 distinct items, including computers, tablets, smartphones, and other electronic devices like smart TVs and wearables, are used to evaluate the suggested method. The suggested system's accuracy was 88.33%. Additionally, it was noted that the system might not consider user reviews that express opinions through the use of emoticons and unusual characters when calculating product scores. Multi-Attribute Analysis Using Fishbein Model revealed that Amazon had a significantly higher overall score than Flipkart. Because of three factors—product variety, customer service, and return and refund policies customers' perceptions of Amazon were far better than Flipkart's, giving Amazon a competitive edge. Consumers believe that the two most significant features are quality and return and refund services. While refunds and returns are among the most crucial features, customers' perceptions of Flipkart in this regard are inferior to those of Amazon. The study concentrated on the preferences, levels of satisfaction, and issues faced by Flipkart online shoppers and advertisers. The primary data for the study were used. The sample size for this investigation was 70, and the samples were chosen using the random sampling technique. The study's research findings helped determine the degree of customer satisfaction based on the goods and websites that customers used. For this study, a standardized questionnaire was used. In this study, it was attempted to ascertain the reasons behind people's selection of Flipkart as well as whether or not the participants are aware of the company's cutting-edge deals and services.

A classification solution was offered by Naive Bayes. To yield more precise results, training and preprocessing are used for data sets. The issues with earlier sentiment analysis research ought to be resolved by this investigation.

To overcome misclassification in the creation of predictive models used to analyze reviews and comments, a more thorough technique is required. The outcomes of a comprehensive hybrid sentiment analysis approach on data sets of various sizes were empirically investigated in this work. The hybrid ensemble method (HEM1) is the most reliable approach for the Balanced Data Models I, II, and III based on several accuracy criteria. The results showed that a compound combination of trigram, bigram, and unigram works well in almost all prediction algorithms. The results suggested that, despite SVMs' ability to handle any degree of data imbalance, data imbalance may have an impact on the use of SVMs in real-time class prediction applications. Extensive trials using real and benchmark device datasets have demonstrated the efficiency and superiority of the improved bagging process over numerous alternative approaches using various data sampling techniques. When using combined approaches, PCA is a powerful dimension-reduction strategy for both balanced and unbalanced datasets.

Three metrics were put forth: (i) the goodness of a product, which gauges its quality; (ii) the reliability of a rating, which measures its dependability; and (iii) the fairness of a user, which gauges how trustworthy the user is in rating the products. It makes sense for a user to be fair if they consistently give ratings that are somewhat close to the quality of the product. A definition of these metrics that is mutually recursive was developed. It also addresses cold start issues and takes user and product behavior into account. FairJudge, an iterative method was used to forecast the three measures' values. It was demonstrated that FairJudge has linear time complexity and is guaranteed to converge in a finite number of iterations. It was demonstrated that FairJudge greatly outperforms nine existing algorithms in predicting fair and unfair users by running five separate experiments on five distinct rating platforms. 80 people were accurately recognized (80% accuracy) out of the 100 most unfair users in the Flipkart network, according to a report sent to their review fraud inspector

CHAPTER-3 PROBLEM FORMULATION

3.1 Importance and Background

Since opinions have a significant impact on our behaviors, they are essential to all human actions. We must be aware of other people's opinions when we have to decide [30]. Businesses and groups always need to find out what the general public thinks of their products and services. Customers interact socially on a variety of online channels, such as web-based social networking sites like Facebook and Twitter. Buyer interaction occurs gradually through these web-based social networks. This sort of relationship presents an incredible opportunity to learn about advertising. People from diverse backgrounds, including color, ethnicity, sexual orientation, and class, use the Internet to exchange experiences and opinions on almost every aspect of their life.

In addition to sending emails, posting blogs, or making comments on business websites, a large number of people use unofficial organization sites to record their thoughts, communicate their emotions, and get insights into their daily lives. People write letters on almost anything, such as films, products, or social activities. These logs are shared throughout virtual communities, where buyers enlighten and influence one another. These logs offer advertisers deep insights into consumers' behavioral tendencies and a constant stream of information on customer sentiments and insights, as they arise, free from disruption or provocation.

However, because the information is dispersed, muddled, and split, the current growth of client-produced material on social media platforms is posing special challenges for gathering, analyzing, and interpreting written content. One information-mining technique that can get around these issues is opinion investigation, which efficiently and quickly separates and analyses web-based data. Despite the challenges posed by the amount and structure of the information, advertisers may continually ascertain the emotions and mental states of consumers using conclusion analysis. There are two reasons why this study is enthusiastic about using sentiment analysis as a tool to promote research tools. Organizations are strongly encouraged to use sentiment analysis to ascertain what consumers like and hate about their offerings and brand perception. Furthermore, the analysis of industry and organization data needs to support decision-making.

3.2 Levels of Sentiment Analysis

Sentiment analysis examines people's viewpoints, assessments, feelings, and attitudes toward certain persons, groups, goods, films, problems, occasions, etc. Sentiment analysis, which includes the process of identifying and extracting sentiment/opinion from the text and categorizing their sentiment, is a key study area within natural language processing (NLP).

3.2.1 Token level

The shortest word that may be constructed with English letters is the literal definition of the term token in English. This level, as its name implies, is the foundational level where Opinion Mining is carried out based only on every token or phrase. As an illustration "Even though the food quality was not very good, I enjoyed the restaurant's service," The first half of the statement builds towards a negative attitude or response, while the second portion expresses a good sentiment, hence the sentence cannot be fully positive or negative. In this instance, the machine learning system would identify this text as neutral, but as a human, you can see that the reaction is mostly favorable. In these situations, we require a token level of opinion mining, where we divide the sentence into manageable chunks, eliminate superfluous words, preprocess the text, and finally categorize the statement as positive, negative, or neutral.

3.2.2 Document Level

This is the final stage at which Opinion Mining may be used. It evaluates the papers based on the name. The entire paper has been examined at this level, and it has been classified as conveying either a good or negative viewpoint. Using this strategy, a single product review is evaluated to determine opinions on the same product. Opinions are voiced about a certain subject. This level won't function if a document has several product reviews since it isn't pertinent to papers with different types of product reviews.

The entire document is categorized as either good, negative, or neutral in this category. Examples of documents where we treat the entire thing as a single unit before attempting to categorize it are book reviews and in-depth descriptions of films. To process the data and evaluate the outcome, many machine learning algorithms are available.

3.2.3 Sentence Level

This level of work involves looking at the sentences and analyzing them to see if they include a neutral, positive, or negative viewpoint. This level divides sentences into objective and subjective categories in a manner akin to Subjectivity Classification. Sentences designated as Objective sentences and Subjective sentences, respectively, are found in both factual and subjective information.

This level is primarily concerned with sentences and serves as the main point of entry for the opinion-mining process. There are many distinct kinds of sentences in English, including declarative, imperative, simple, compound, complicated, interrogative, imperative, and exclamatory phrases. As an illustration, "He is a good boy." is a brief, concise phrase that expresses positivity without ambiguity. However, a statement such as "You are a good player, but I am very disappointed that you failed the exam," is an example of a difficult statement where the presence of many emotions prevents the model from accurately predicting the mood. Another example can be, "The client accepted our proposal and signed a contract with us, even though the presentation was not very good, he was happy." This is an illustration of a compound-complex sentence, which contains more than two sentences. A human can now correctly identify the sentiment by saying that, overall, the sentence is positive because the goal was achieved, but it can be difficult for a machine to classify these kinds of sentences as positive, negative, or neutral when it comes to a high degree of precision and accuracy.

3.2.4 Paragraph level

This level, which comes after the sentence level, calculates feelings by taking the entire paragraph into account. With social media comments, reviews, and feedback becoming more and more popular, these sources are useful for analyzing opinion mining at the paragraph level. These days, paragraph-level analysis can consider a post's captions as well.

3.2.5 Aspect Level

Opinion mining and summarization based on features are included in aspect-level sentiment analysis, also known as feature-level sentiment analysis. It is crucial to ascertain the precise reasons why people liked or disliked this level. This degree of sentiment analysis is more detailed.

The aspect level examines the viewpoint directly rather than through documentation or phrases. This level's output will include the entity, its aspect, the opinion holder's opinion, and time. The Samsung J7, for instance, offers the finest camera quality. At this point, the Samsung J7's camera is an element that conveys positivity. The actor, acting, actions, and specific scenes in a movie are some of the important elements under aspect-level sentiment analysis.

3.3 Sentiment Analysis Techniques

Sentiment analysis for social networking has seen the use of several methodologies. These methods may be divided into three primary groups: hybrid methods, lexical methods, and machine learning methods. Figure 1.1 shows the sentiment analysis techniques.

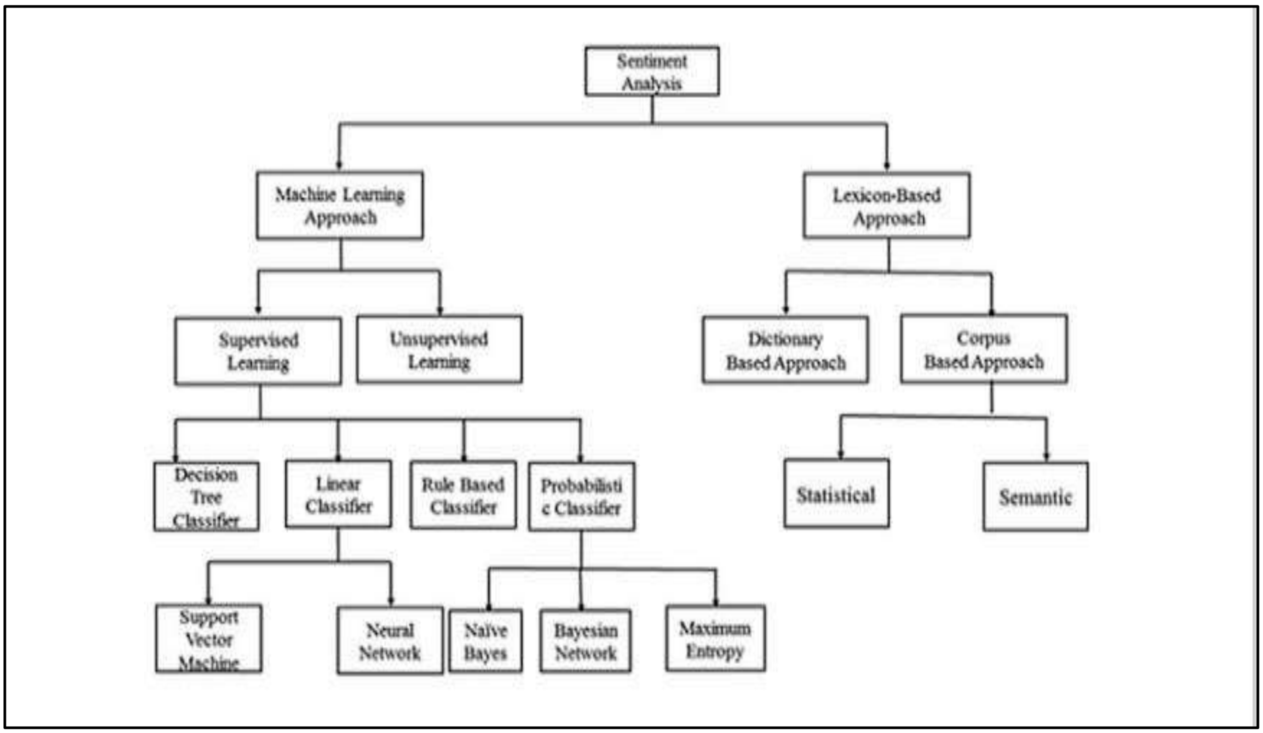


Figure 3.1: Sentiment analysis techniques [5]

3.3.1 Machine learning techniques

Three primary categories may be used to further categorize machine learning approaches: supervised, unsupervised, and deep learning techniques. These methods and the various algorithms used in sentiment analysis are presented in the next subsections.

3.3.2 Supervised Learning Techniques

To train the model, supervised learning techniques may be applied when labeled data is available. The two stages that are used in this context are model training and new case prediction. The labeled data set is fed into the classification algorithm during the training phase, and the algorithm generates a model as an output. The model is then given the test data to forecast the class of the newly discovered instances.

Sentiment analysis has made use of these approaches; for instance, some researchers created a training set using emoticons and hashtags, then utilized the emoticons as class labels to ascertain the polarity of social networking posts. An algorithm developed by Pang et al. determines the likelihood of each phrase in the training dataset, which is derived from movie reviews and used to determine whether or not a review is good. The algorithm divides the newly classified phrase into distinct word characteristics after classifying the new instance of data. The built model then computes the conditional probability of the combined features to forecast its class using the probabilities that were determined during the training phase. Furthermore, to improve the aspect-based sentiment analysis of evaluations from Arabic hotels, the Bayesian Network technique was used [37]. The dependence between the phrases is shown as an acyclic-directed graph with nodes representing the words as variables and edges showing the relationship between these variables. The Support Vector Machine (SVM) technique lowers the top bound of classification errors by using the hyperplane to provide the greatest separation between classes with the maximum margin of hyperplane.

With several weighting techniques, including termfrequency–inverse document frequency (TF-IDF), term occurrence, and Binary Occurrence, which use chi-square as a feature selection method, it may be utilized for dimensionality reduction and noise removal. The insight concealed in popular opinion has been interpreted by the use of Twitter data. It was employed to categorize Twitter tweets' feelings and identify whether they were favorable or unfavorable. Several dataset types have been subjected to the Decision Tree technique, which involves segmenting the training data into smaller pieces and using those segments to find hidden patterns that are then utilized for classification [39]. Furthermore, Reuters corpus documents are utilized as a training dataset in the Artificial Neural Network (ANN) for text categorization.

3.3.3 Unsupervised learning techniques

In certain domain areas, it might be simple to gather unlabeled data, which is why these strategies are employed. The categories of the sentences are determined by the keywords. One method used to conduct sentiment analysis on unlabeled data is clustering algorithms. These methods can assist in classifying the users' sentiment messages into three categories: neutral, negative, and positive.

To carry out clustering and provide a suitable number of clusters in a suitable amount of run-time, clustering algorithms were used. To do sentiment analysis, tweets were also categorized into positive and negative tweets using an unsupervised method based on spectral clustering. To conduct an aspect-based sentiment analysis, which focuses on obtaining and classifying broad viewpoints on the characteristics of a certain good or service, the clustering-based technique was created. A lot of aspect clustering methods are monolingual and need multilingual applications. This program does sentiment analysis in many languages and semantically classifies related characteristics.

3.3.4 Semi-supervised learning techniques

The hybrid kind of machine learning known as semi-supervised learning occurs when a portion of the information is fully unorganized but the remainder has a label attached to it. It reduces the issues with both supervised and unsupervised learning models and is frequently applied in situations when the dataset is enormous and has a matching label. One technique that may be used to create semi-supervised models in situations when there are massive quantities of data and associated labels is clustering.

3.3.5 Deep learning techniques

Deep learning techniques are multi-layer approaches that modify the neural network's hidden layers. The way these methods approach the feature extraction procedure sets them apart from other machine learning methods. While deep learning techniques learn and extract features automatically to obtain high accuracy in the learning process, supervised and unsupervised learning approaches extract features either manually or via a feature extraction method.

One of the greatest solutions available right now for a lot of sentiment analysis issues is deep learning. Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) are the three primary categories under which it falls. Other approaches that are based on integrated methodologies include the Belief Neural Network, Recursive Neural Network, and Hybrid Neural Networks.

The DNN is a multi-layered, intricate mathematical model that has various hidden layers that have several data processing capabilities. Three levels make up the DNN: an input layer that holds the input data, further hidden layers that house neurons, which are processing nodes, and an output layer that holds one or more neurons, the outputs of which are the network outputs. A classifier coupled with a word embedding model and a linear machine learning algorithm has been used to apply an improving DNN model to sentiment analysis for social networking.

Furthermore, a particular kind of neural network called a CNN is employed in computer vision, recommender systems, and natural language processing (NLP), among other fields. Inputs to the classification layer consist of the pooling or subsampling layers and the convolutional layers. It has been used to manage enormous volumes of unstructured data, which is a difficult procedure. CNN has been used to study feedforward neural networks with several hidden layers.

3.3.6 Classification techniques

Different tactics are employed to categorize unlabeled data through the development of classification systems in the field of machine learning. It's possible that classifiers need training data. Naive Bayes, Maximum Entropy, and Support Vector Machine are a few types of machine learning classifiers. Because they need training data, they fall under the category of supervised machine learning techniques. It is noteworthy to emphasize that future predictions will be simpler if a classifier is trained successfully.

Naïve Bayes

With strong (naive) independence assumptions between the features, this classification technique is based on Bayes' Theorem. A Naive Bayes classifier anticipates that there would be no correlation between the proximity of a particular characteristic (element) and the closeness of other elements in the class. For example, an organic fruit with a red color, a round form, and a breadth of around three inches may be classified as an apple. Because this natural fruit is likely an apple, a Naïve Bayes classifier would consider these attributes independent, regardless of whether these features depend on each other or the existence of other features.

In addition to being simple, the Naive Bayes algorithm has been shown to outperform even more advanced order schemes. Equation 1 is used as a computation technique, known as the Bayes hypothesis, and can be used to separate the likelihood P(a|b) from P(a), P(b), and P(b|a):

$$P(a/b)=[p(b/a)X(a)]/p(b) \qquad (1)$$

Where p (a/b) is the posterior probability of class a given predictor b.

p (b/a) is the likelihood that is the probability of predictor b given class a.

The prior probability of class a is denoted as p(a), and the prior probability of predictor p is denoted as p(b). The Naive Bayes is widely used in the task of classifying texts into multiple classes and was recently utilized for sentiment analysis classification.

Maximum Entropy

The Maximum Entropy (MaxEnt) classifier estimates the conditional distribution of a class marked a given a record b utilizing a type of exponential family with one weight for every constraint. The model with maximum entropy is the one in the parametric family P_{MaxEnt} (a/b) that maximize the likelihood.

Numerical methods such as iterative scaling and quasi-Newton optimization are usually employed to solve the optimization problem. The model is represented by the equation 2

$$P_{MaxEnt}(a/b)=\exp[\sum i\ ai\ fi(a.b)/\sum a \exp[\sum i\ ai\ fi\ (a,b)] \qquad (2)$$

where a is the class, and b is the predictor. The weight of the vector is denoted as $a_{i.}$

3.3.7 Lexicon Techniques

Lexicons are collections of tokens, each of which is given a predetermined score that denotes the text's neutral, positive, or negative character [50]. Tokens are awarded a score based on their polarity, which can be +1, 0, − 1 for positive, neutral, or negative, or based on the strength of the polarity, with a value range of [+1, − 1], where + 1 indicates extremely positive and − 1 indicates highly negative. In the Lexicon Approach, the summation of positive, negative, and neutral token ratings is done independently for a specific review or content.

The text is given an overall polarity in the final step, which is determined by the greatest value of each score. As a result, the document is first split up into tokens, which are made up of a single word. Next, the polarity of each token is determined and combined at the conclusion.

For sentiment analysis at the feature and sentence levels, the lexicon-based method is quite practicable. It might be referred to as an unsupervised approach because no training data is needed. Conversely, domain dependence is the fundamental drawback of this method as words can have several meanings and interpretations, making a term that is positive in one domain but negative in another. For example, the word "small" in the sentences "The TV screen is too small" and "This camera is extremely small" is negative because people generally prefer large screens, but in the second sentence it is positive because the small size of the camera will make it easier to carry. This problem can be solved by creating a new lexicon specific to a certain topic or by modifying an already-existing language. The lexicon-based method mainly uses lexemes, which are words or tokens. Tokens from the sentence are divided up and processed. These comments might be classified as positive or negative opinions.

Classification of Lexicon based approach is as follows:

● Corpus-based approach

It entered the picture to address the issues with the dictionary-based method. Compared to dictionary-based approaches, it is less effective since creating a big corpus encompassing English terms is a challenging effort.

Dictionary-based approach It is unable to locate opinions with a domain-specific focus. It is more efficient than the corpus-based method. It makes use of a dictionary that contains every synonym and antonym for every term.

3.3.8 Hybrid techniques

Several strategies have been incorporated in a large body of work to improve sentiment analysis. Some of these hybrid strategies will be introduced in the paragraphs that follow. The method shown was created to analyze feedback related to the same subject to extract relevant contextual information. The semantic orientation of a particular opinion was confirmed using semantic similarity measures.

By using language norms to carry out the semantic orientation of contextually dependent opinions, it is considered contextual opinions. It then gathered contextual data from other evaluations for the same product feature to evaluate opinions that rely on context. Peng and Shih studied an unsupervised learning technique that extracts each review's sentiment phrases using part-of-speech (POS) pattern rules.

They utilized each unknown sentiment word as a query to identify the top-N pertinent snippets for each sentiment phrase. The predicted sentiments of unknown emotion phrases were computed after compiling the lexicon's sentiment. Predictive sentiments were then found by analyzing the sentiments of nearby known sentiment phrases within the snippets. They limited their analysis to sentences with at least one identified sentiment term to extract opinions.

3.4 Applications

Sentiment analysis has a wide range of uses, such as examining consumer sentiment and assessing a patient's mental health based on postings made on social media. Moreover, sentiment analysis now has a wider variety of applications across almost all disciplines because of technology advancements like blockchain, IoT, cloud computing, and big data. A few significant domains and industries where Sentiment Analysis is applied are described below:

i. Business Analysis

In the field of business intelligence, sentiment analysis has several advantages. Businesses may also use sentiment analysis data to explore customer comments, enhance goods, and create creative marketing campaigns. In the field of business intelligence, sentiment analysis is most commonly used to examine how consumers feel about certain services or goods.

However, customers may also utilize these studies to assess products and make more informed selections, so they are not just useful for product manufacturers. There are several benefits to sentiment analysis in business intelligence. Businesses can utilize the findings of sentiment analysis, for instance, to improve products, look into customer feedback, or create new marketing strategies.

ii. Market research and competitor analysis

Aside from public opinion research and brand image monitoring, market research is arguably the most popular use of sentiment analysis. Sentiment analysis is used to compare marketing efforts and identify the company that is standing out from the competition. It may be used to build up a comprehensive image of a brand's and its rivals' consumer base.

Sentiment analysis has the potential to gather information from many sources, such as blogs, Facebook, and Twitter, provide measurable outcomes, and get over obstacles in business intelligence.

iii. Reputation Management

Utilizing sentiment research in a variety of marketplaces allows for reputation management and brand monitoring. Fashion labels, marketing firms, IT firms, hotel chains, media outlets, and other enterprises can all benefit from assessing consumer perceptions of their brand, good, or service. The use of a sentiment analysis tool enhances the diversity and intelligence of a brand's and its goods' representation. It helps companies to monitor how consumers view their brands and to pinpoint the exact information about their opinions. Observe patterns and modifications, and focus on influential people's presentations. Sentiment analysis may be used in conjunction with other automation techniques to automate both the alarm system and the media surveillance system. Monitor the brand's comments and reviews across a range of social media channels.

iv. Aspect analysis

Businesses may maximize their use of the vast quantities of data they generate by implementing aspect-based sentiment analysis. Businesses will be able to extract the most crucial elements of customer feedback and service thanks to the aspect-based strategy.

v. Voice of customers

Gather and evaluate all customer input via chat rooms, email, surveys, call centers, and the Internet. Sentiment analysis will make it possible to classify and arrange data to identify patterns, recurring problems, and concerns. A successful company operation requires both the identification of a suitable client group and the subsequent development of a value offer, both of which may be facilitated by sentiment analysis. However, it needs to keep an eye on its consumers' pulses to stay current and keep the product in demand.

vi. Social Media Monitoring

When anything negative begins to spread, sentiment analysis of social data will track customer sentiment in real-time, 24 hours a day, 7 days a week. This allows for quick responses and positive comments to boost one's reputation. Additionally, consistent and trustworthy client data is obtained, allowing decision-makers to monitor client development seasonally. People often provide some of the most honest opinions on businesses, goods, and services on social media because they make comments without being asked. They have to share their emotions with the world.

vii. Stock Market

Predicting stock prices is one use for sentiment analysis. It may be accomplished by forecasting stock price movements and examining all stock market news. Numerous sources, including blogs, Twitter, and news articles, can be used to gather data. Sentiment analysis at the sentence level may be performed on these messages, and the overall polarity of the texts of news about a certain firm can then be determined.

3.5 Challenges

There are several difficulties in assessing opinions from reviews, comments, etc. Reviews typically include unpredictable and conflicting facts. Individuals communicate their feelings in a variety of ways; occasionally, they use a lot of acronyms and shorthand. In reviews, they typically struggle with using good language. Using opinion words and phrases that are typically used to communicate opinions, we evaluate whether reviews are favorable or negative.

You can use these expressions of opinion and phrases in both good and negative contexts. As an illustration, good is for positive, and bad is for negative. Determining the review's positive and negative feelings depends on the surrounding environment.

Very few words may consistently be used to infer a positive or negative meaning from an utterance. Irony and repressed feelings are also present in reviews and comments. Because opinionated texts sometimes contain ambiguity and subjective words, assessing sentiments may be a difficult undertaking.

Words with the same meaning that appear more than once in a single phrase are known as ambiguous words. When irony and language are used to express meaning, ambiguity turns into a major issue. Consider this statement, for instance. The requirement for lexicons for other languages is one of the main problems with lexicon-based techniques.

Only a few common languages, such as English, Arabic, Chinese, and so on, have lexicons available; lexicons are not available for lesser-known languages. Additionally, the lexicons of the Chinese and Arabic languages are not comprehensive enough to contain all of the terms used in these languages.

i. Multiple language input

The information may be available in more than one language because it contains an assortment of user evaluations. However, the classifier speaks English mostly. As a result, training the algorithm for languages other than English becomes extremely challenging. Therefore, one of the main challenges in sentiment analysis is multiple language input.

ii. Fake inputs

Fake or fraudulent reviews deceive users or consumers by offering fictitious or unrelated good or negative feedback about a product. Mostly, this is done to make a product more or less popular. Therefore, spotting phony reviews is a difficult and nearly impossible process.

iii. Emoticons and sarcastic reviews

Emoticons are expressions conveyed through pictures. It is simpler for the consumer or customer to comprehend one's sentiments when emoticons are used to explain the product. However, the emoticons grow more challenging for a machine to comprehend. Training an algorithm with emoticons as input is a challenging task. Sarcastic evaluations are hard for the computer to understand. To provide a precise response, the model must be trained using an increasing amount of these kinds of data. Emoticons and caustic evaluations are therefore among the most difficult aspects of sentiment analysis.

3.6 Research gap

Sentiment analysis is one of the computer science subfields that is growing at a fast pace, it is challenging to keep up with all of the developments in the field. Despite significant advancements in the application of machine learning (ML) for sentiment analysis, several research gaps persist. Firstly, there is a notable variance in the effectiveness of these models across diverse populations, highlighting a need for more inclusive and representative datasets to train these algorithms. The said studies often rely on data from specific geographic locations or demographics, potentially limiting their generalizability. Secondly, many existing models prioritize accuracy over interpretability. Also, the research on sentiment analysis of product reviews mostly focuses on sentiment analysis methods and rarely involves feature extraction and large-scale data recognition. The findings reviewed imply that understanding the psychological condition of the general population can be aided by identifying the most important positive and negative terms in the data set.

3.7 Objectives

1. To study and analyze the existing machine learning and deep learning techniques for sentiment analysis.
2. To develop and apply the model for the flipkart.com product review dataset.
3. To visualize and analyze the results obtained from the proposed model.

3.8 Research design

The choice of descriptive and exploratory research was made with the expectation that it would provide marketers with a clear understanding of the millennial mindset. By using the Bayes algorithm to classify emotion, this method seeks to extract emotions from the dataset and categorize them into emotions by assigning scores based on the emotions connected with that tweet. The graph is plotted following the given score. Figure 3.1 shows the research design for the sentiment analysis model and figure 3.2 shows the sentiment polarity model design.

Figure 3.2: Research flow diagram.

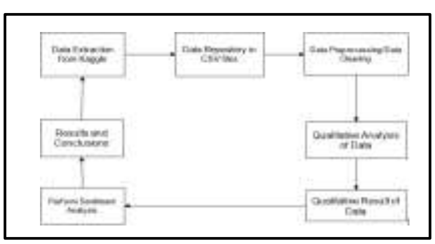
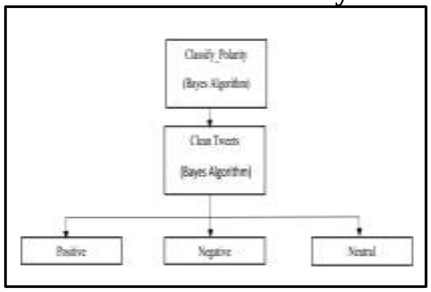


Figure 3.3: Sentiment Polarity Model



3.9 Research Methodology

The process of gathering data is the first step in any analysis. Social media sites now provide research and analysis with access to a vast array of data. Web scraping methods may be used to obtain real-time data from social media networks. Web scraping in Python is done with the Beautiful Soup package. Preprocessing this retrieved raw data results in clean data, which streamlines the process of additional analysis. To obtain more reliable findings, preprocessed data is used for feature extraction. Sentiment categorization is done by a variety of algorithms. In the entire procedure, this is the most crucial stage.

Data Collection: Gathering substantial volumes of data from pertinent sources is the first stage in the Opinion Mining process. Some of the previously established, well-known datasets, such as the data from Facebook posts, reviews, and comments on Twitter, may be utilized to extract the dataset for the same purpose. Another way to get data is to conduct a survey using a Google Form, analyze the results, and go through blogs, ratings, popular discussion forums, and other sources of similar information. Direct datasets are also accessible for download on the Kaggle platform, where massive numbers of datasets are available. A dataset based on reviews of products that users bought from flipkart.com during the period of the last 6 months was obtained from a highly trusted source known as Kaggle. This dataset contains information about product names, product prices, ratings, reviews, and summaries. There are 104 different product types on flipkart.com, such as electronics items, men's, women's, and children's clothing, home decoration items, automated systems, and so on. It has 205053 rows and 5 columns. It contains five features as documented in table 3.1.

Table 3.1: Features Included in the Dataset

Feature	**Description**
Product Name	Name of the Product
Product Price	Price of the Product
Rating	User Rating between 1 to 5
Review	User reviews are provided for every Product
Summary	A detailed description including specifications, features, and product dimensions

Data Preprocessing: Pre-processing is the first action taken following data cleansing, during which data is separated into lists, converted to lowercase, and all punctuations are eliminated. It is very advised to put all of the data through this process after it has been gathered using forms, APIs, or other platforms. To ensure that the outcomes of pre-processing are unaffected, this stage entails removing any undesired or noise-filled data.

Numerical data, URLs, special symbols, emails, and statements that are extremely repetitious and don't add anything to the analysis section are examples of unwanted data. This is the foundational phase since it guarantees an even greater level of accuracy and precision once the data has been completely cleansed and prepared for pre-processing. Numerous methods exist that are intended for the aim of cleansing data. These techniques, which are some of the checks necessary for the data cleaning on the complete dataset, include data range, data type, data limitations, and cross-examination of data. Three processes are completed as part of the preparation of the data: tokenization, elimination of stopping words, and use of the global constant to fill in the missing values. -Tokenization is the process of dissecting a string sequence into its constituent words, phrases, symbols, characters, and many other things. Tokens that form words, phrases, or even entire sentences can be created. Many characters, including punctuation, are eliminated during tokenization. The remaining characters are subsequently utilized in a variety of applications, including text mining and parsing.

-Eliminating stop words: In text mining, a stop word is any phrase element that doesn't contribute to any subdivision. Usually, these phrases are not used to improve the accuracy of the evaluation. There are several kinds of stopwords, depending on the language, the realm, etc. Nonetheless, there are a few stopwords in English.

-Filling the missing value with a global constant: The system looks for missing values in the dataset during this step. The procedure is then finished by substituting the correct constant for any missing values.

Data preprocessing includes proper fragmentation of data and cleaning of data. Here in research work, NLP preprocessing techniques like removal of stop words, chunking data, stemming data, etc. were used. Data preprocessing led us to robust data that had less noise. For data preprocessing, the use of the Natural Language Tool Kit (NLTK) library implemented in Python was considered. NLTK is a platform for natural language processing developed in Python. Table 3.2 shows the dataset after preprocessing.

Table 3.2: Dataset After Preprocessing

Sr No.	Summary	Sentiment
1	great cooler excellent air flow price amazing ...	Positive
2	best budget fit cooler nice cooling	Positive
3	quality good power air decent	Positive
4	bad product fan	Negative
5	ok ok product	Neutral

Feature Extraction:

It is the next action that follows the complete data pre-processing. It is not possible to train the model directly using all of the data; instead, specific point selection must be carried out to lessen the modal complexity and improve the model's delicacy. While building a modal with a normal machine learning algorithm, some factors need to be considered while determining the various characteristics to consider. The processes of feature extraction and feature selection are interdependent.

The first subtask that minimizes the collection of features to be considered for the model is feature extraction. This uses the original dataset to create a complete set of features. Feature selection aims to produce a dictionary of key-value pairs, where the features are the keys and the values are the values. Values are the rankings that are given to every single characteristic. Words are mathematically represented by identifying their sentiment characteristics and applying the word embedding approach. For feature extraction in this work, a hybrid strategy combining the TFIDF hybrid technique and the Skip N-Gram model is employed.

-Model of Skip N-Gram: A skip-gram is a generalization of n-grams from computational linguistics, especially language modeling, where the constituents (usually words) of the skip-gram are allowed to include gaps rather than necessarily being sequential in the text being examined. The data sparsity issue that n-gram analysis presents can be solved using a skip-gram. Skip-grams are more resilient against attacks than n-grams in terms of computer security. N-grams are sequentially occurring closed sequences of tokens $w_1 ... w_n$.

-TFIDF hybrid method: Words in texts can have their mathematical significance determined using the TF-IDF. The TF-IDF value is obtained by multiplying the TF and IDF numbers. The phrase "Term Frequency" (Term Frequency) describes the ratio of target terms to total terms in a text, as the name suggests. The initial IDF values are computed as a logarithm of the ratio of documents containing the targeted phrase to the total number of documents. At this point, it doesn't matter how many times the phrase appears in the document

Classification: The process of text classification is divided into two stages: The training stage and the Testing stage. In the training stage, a classification model is created using a testing dataset. In the testing stage, the accuracy of classification is evaluated using the classification module. We have used the Naïve Bayes classifier. The Mapper class assigned sentiment values and the reducer class evaluated polarity. The training job created the model. Combining jobs combined the model and test data. Classify job performed polarity determination. Customer response can be positive, negative, or neutral. For each review, subjectivity value (range 0 to 1) and polarity value (range -1 to +1) was calculated and the total value of sentiment was calculated as the summation of the product of subjectivity and polarity values of the individual reviews. Table 3.2 provides a snapshot of the sentiment calculation process.

Table 3.3. Sentiment Calculation

Review	Subjectivity (s)	Polarity (p)	Sentiment Score (s*p)
great cooler and excellent airflow and for this price it is so amazing and unbelievable just love it	0.812500	0.725000	0.5890625
best budget 2 fit cooler nice cooling	0.566667	0.666667	0.37777819
the quality is good but the power of the air is decent	0.633333	0.433333	0.27444409
very bad product it's only a fan	0.933333	-0.455000	-0.42466652
ok ok product	0.500000	0.500000	0.25

The text is positive if the polarity is more than 0, negative if the polarity is less than 0, and neutral if the polarity is equal to 0. The subjective range runs from 0.0 to 1.0. A more excellent score indicates that the text is more subjective. Table 3.4 shows the value counts for positive, negative and neutral reviews.

Table 3.4: Value_Count of reviews

Positive	166581
Negative	28232
Neutral	10239

Visualization: Visualization of the classification and the outcomes generated by the machine learning algorithms is strongly advised after developing a machine learning model. Any common dataset used to train the machine should be represented as a graph for sentinel analysis so that the continuous distribution of the data is visible. The results are displayed using graphs and charts. The word cloud was built using the frequency of occurrence of words. Figure 3.4 shows the visualization and figure 3.5 shows the Count Plot of the sentiments. Figure source: self-generated

Figure 3.4: Visualization.

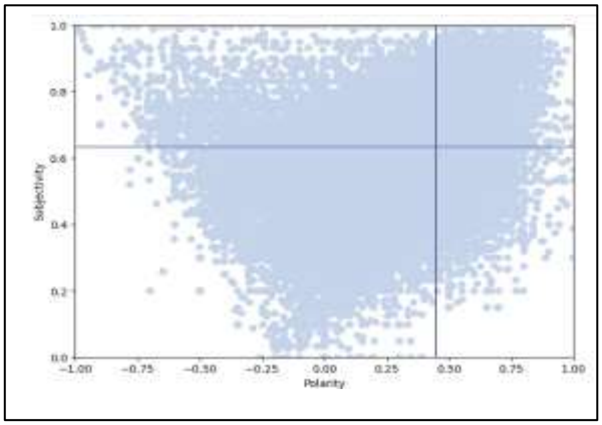
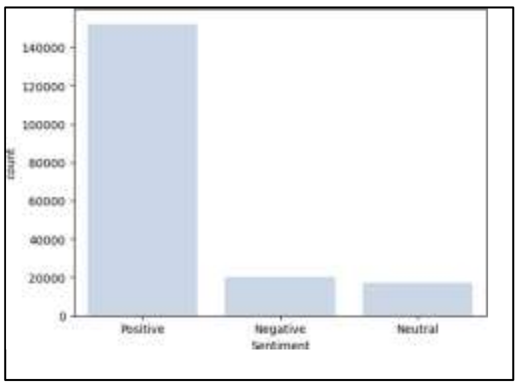


Figure 3.5: Count Plot of the sentiments.



Testing: Testing is necessary to confirm whether or not the developed model accurately predicts the intended result. Another name for it is the validation step. The fact that this process should function well when applied to large-scale applications makes it crucial. The final stage is testing, in which a user inputs a text into a machine at runtime, and the computer makes predictions about the statement: whether good, negative, or neutral.

-Algorithms for sentiment analysis

Evaluation criteria like accuracy, recall, precision, and f1-score are used to gauge how well algorithms work [60]. Sentiment analysis may be carried out using a variety of algorithms. They are Random Forest, Naïve Bayes, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN).

Naïve Bayes: An algorithm for supervised learning is the Naive Bayes classifier. Its foundation is the Bayes theorem. Given that it makes predictions based on the probabilities of the item, Naive Bayes is a probabilistic classifier. The Bayes theorem determines a hypothesis's probability based on prior information. It computes a posterior, or revised, probability, which is a conditional probability. The probability of any two random variables, A and B, is given by, following Bayes' rule.

P(A/B) = (P(B/A) P(A)) / (P(B))

Where, P(A/B) = Posterior probability,

P(A) = Prior probability,

P(B) = Marginal probability,

P(B/A) = Likelihood probability

Naive Bayes theorem predicts a class value for a given set of attributes. For each known class value,

- Naive Bayes calculates probabilities for each attribute, conditional on the class value.
- It uses the product rule to obtain a joint conditional probability for the attributes.
- It uses the Bayes rule to derive conditional probabilities for the class variable. When all class values are calculated, output the class with the highest probability.

-Evaluation of Algorithms: The algorithms covered in the preceding section are all assessed using metrics like F1 score, accuracy, precision, and recall. Confusion matrix measurement is used for these measures. Precision, recall, and F-measure, Sensitivity, Specificity, Accuracy are the performance metrics used to assess the categorization outcomes. The values of true positive (TP), false positive (FP), true negative (TN), and false-negative (FN) assigned classes on empirical are used to calculate those metrics [61]. As in other studies, this paper uses sensitivity, specificity, accuracy, precision, recall, and F1 score as metrics of model evaluation. Following are the calculation parameters.

True Positives (TP) - These are the correctly predicted positive values which mean that the value of the actual class is yes and the value of the predicted class is also yes.

True Negatives (TN) - These are the correctly predicted negative values which mean that the value of the actual class is no and the value of the predicted class is also no.

False Positives (FP) - When the actual class is no and the predicted class is yes.
False Negatives (FN) - When the actual class is yes but the predicted class is no.
Confusion Matrix: Confusion matrix is a tool for analyzing how well our classifier can recognize tuples of different classes. Figure 3.6 shows the classical confusion matrix.

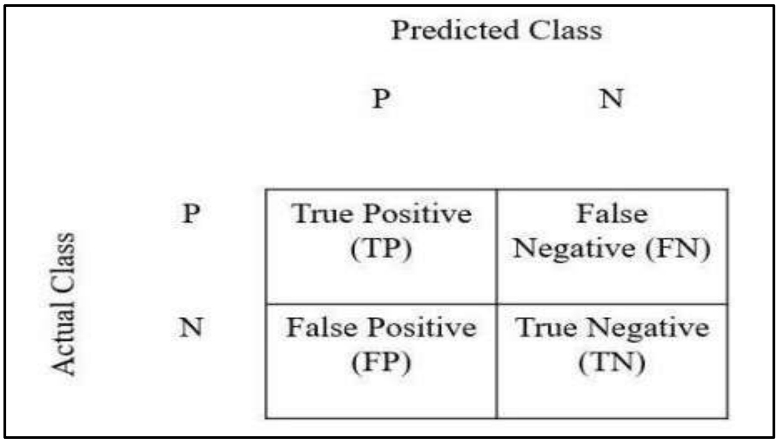


Figure 3.6: Confusion matrix [33]

Accuracy: Any classification algorithm's accuracy is measured by the proportion of test set tuples that the model properly classifies.

Accuracy = ((TP + TN)) / ((TP + TN + FP + FN))

TP = True Positive,
TN = True Negative,
FP = False Positive,
FN = False Negative

Precision: Precision refers to an algorithm's accuracy. It is the proportion of correctly identified positive tuples that are positive.

Precision = TP/ (TP + FP)
TP = True Positive,
FP = False Positive

Recall: The test of completeness is recall. It is the proportion of positively labeled tuples by the classifier.

Recall = TP/ (TP + FN)
TP = True Positive,
FN = False Negative

CHAPTER-4 RESULTS

The most often used terms that produced both positive and negative feelings may be seen in the sentiment analysis of the dataset of customer reviews for e-commerce products. To discriminate between bad and positive evaluations, the dataset was sorted by the sentiment columns. The positive word cloud is depicted in Figure 4.1 whereas the negative word cloud is depicted in Figure 4.2. Figure 4.3 lists the most frequently used terms. Table 4.1 shows the common positive word count, Table 4.2 shows the common negative word count and table 4.3 shows the common neutral word count. Figure source: self-generated.

Figure 4.1: Positive Word cloud

Table 4.1: Common Positive Word Count

Sr No.	Common Words	Count
1	Good	17442
2	Awesome	11294
3	Nice	9396
4	Worth	9037
5	Wonderful	9029
6	Recommended	6695
7	Great	5703
8	Brilliant	5648
9	Perfect	5615
10	Super	5608
11	Classy	5605

Figure 4.2: Negative Word cloud

Table 4.2: Common Negative Word Count

Sr No.	Common Words	Count
1	Waste	2199
2	Disappointed	1756
3	Rubbish	1154
4	Terrible	1148
5	Utterly	1131
6	Worthless	1119
7	Worst	1094
8	Hated	1088
9	Don’t	1079
10	Scam	50

Table 4.3: Common Neutral Word Count

Sr No.	Common Words	Count
1	Good	1288
2	Okay	570
3	Decent	482
4	Better	259
5	Utterly	1131
6	Valueformoney	239

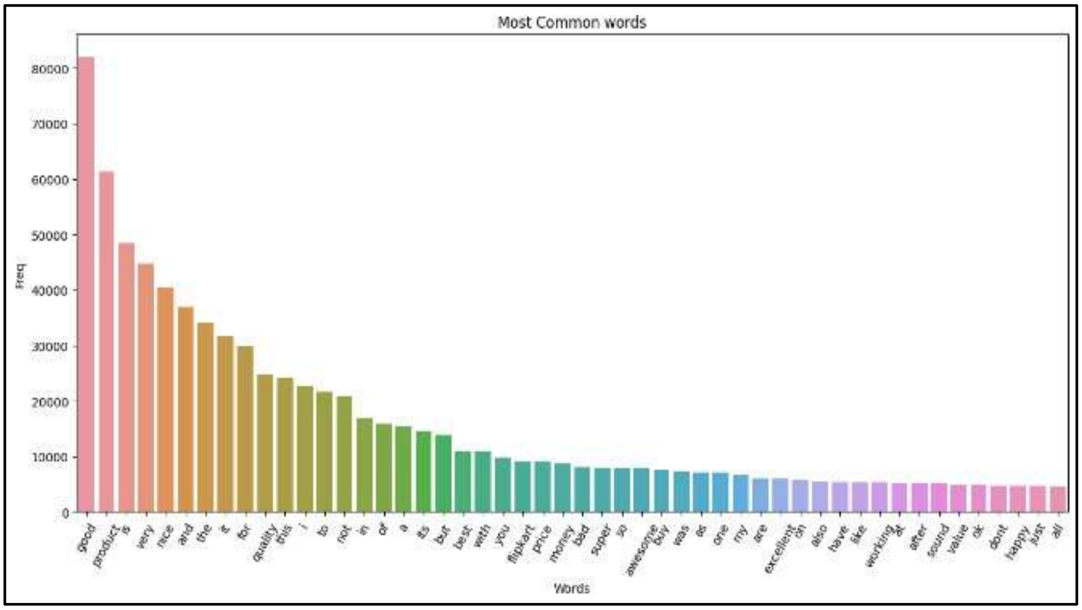


Figure 4.3: Most Common Words

One metric used in the model evaluation is cross-validation, which shows how well the system can predict the future. The effectiveness of the models has been confirmed by the application of K-fold cross-validation. The dataset is split into k subsets for K-fold cross-validation, and each subset is repeated k times. A training sample of k subsets and a testing sample of k-1 subsets are used for each iteration. The model's cross-validation is displayed in Table 4.4.

```
from sklearn.naive_bayes import ComplementNB
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import cross_val_score
from sklearn import metrics
from math import *
cnb = ComplementNB()
cnb.fit(X_train, y_train)
cross_cnb = cross_val_score(cnb, X, y,n_jobs = -1)
print("Cross Validation score = ",cross_cnb)
print ("Train accuracy ={:.4f}%".format(cnb.score(X_train,y_train)*100))
print ("Test accuracy ={:.4f}%".format(cnb.score(X_test,y_test)*100))
train_acc_cnb=cnb.score(X_train,y_train)
test_acc_cnb=cnb.score(X_test,y_test)

Cross Validation score =  [0.8617 0.8642 0.8597 0.8556 0.8621]
Train accuracy =93.0475%
Test accuracy =86.2400%
```

Figure 4.4: Cross Validation Results

Table 4.4: Cross Validation for Naïve-Bayes

Runs	1	2	3	6	5
Average	86.17	86.42	85.97	85.56	86.21

The confusion matrix is used to present the rates of TP, FP, TN, and FN of the sample. Based on these rates, the evaluation metrics (specificity, accuracy, recall, precision, and F1-score) were calculated to evaluate the Naïve Bayes model using unseen data to predict the sentiment of customers. Figure 4.6 shows the visualization of the confusion matrix.

```
# Print the Confusion Matrix and slice it into four pieces
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, predicted)
print('Confusion matrix\n\n', cm)
print('\nTrue Positives(TP) = ', cm[0,0])
print('\nTrue Negatives(TN) = ', cm[1,1])
print('\nFalse Positives(FP) = ', cm[0,1])
print('\nFalse Negatives(FN) = ', cm[1,0])
```

```
Confusion matrix

 [[4329  648]
 [ 728 4295]]

True Positives(TP) =  4329

True Negatives(TN) =  4295

False Positives(FP) =  648

False Negatives(FN) =  728
```

Figure 4.5: Confusion Matrix

```
# visualize confusion matrix with seaborn heatmap
import seaborn as sns
cm_matrix = pd.DataFrame(data=cm, columns=['Actual Positive:1', 'Actual Negative:0'],
                                 index=['Predict Positive:1', 'Predict Negative:0'])

sns.heatmap(cm_matrix, annot=True, fmt='d', cmap='YlGnBu')

<Axes: >
```

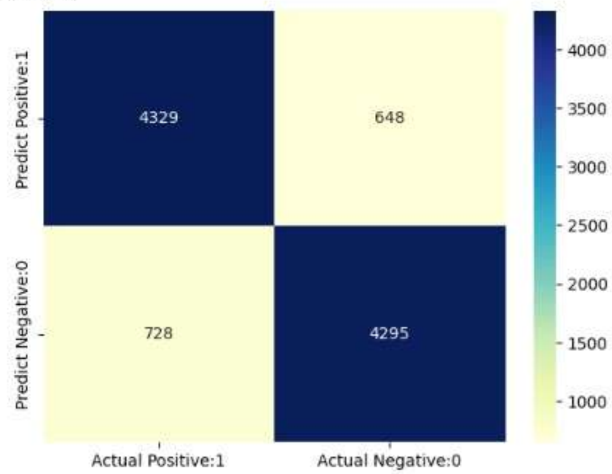


Figure 4.6: Visualization of Confusion Matrix

Precision, Recall, and F1-measure are computed from the confusion matrix. The result of each of the parameters of the Data set is shown in figure 4.7

```
[62]
from sklearn import metrics
predicted = CNB.predict(X_test)
accuracy_score = metrics.accuracy_score(predicted, y_test)

print('ComplementNB model accuracy is',str('{:04.2f}'.format(accuracy_score*100))+'%')
print('------------------------------------------------')
print('Confusion Matrix:')
print(pd.DataFrame(confusion_matrix(y_test, predicted)))
print('------------------------------------------------')
print('Classification Report:')
print(classification_report(y_test, predicted))

ComplementNB model accuracy is 86.24%
------------------------------------------------
Confusion Matrix:
      0     1
0  4329   648
1   728  4295
------------------------------------------------
Classification Report:
              precision    recall  f1-score   support

           0       0.86      0.87      0.86      4977
           1       0.87      0.86      0.86      5023

    accuracy                           0.86     10000
   macro avg       0.86      0.86      0.86     10000
weighted avg       0.86      0.86      0.86     10000
```

Figure 4.7: Classification Report

CHAPTER-5 CONCLUSION

Sentiment analysis is one of the fastest-developing areas in computer science, making it difficult to track every advancement in the discipline [2]. Sentiment analysis has become a valuable tool for the generation and evaluation of different types of data, helping the decision-making processes that lead to the improvement of businesses and companies. Social networking creates a large amount of data that requires processing and analysis to obtain relevant insights. Based on the research, it can be said that younger generations are becoming more and more accustomed to purchasing online. E-commerce websites are seeing an increase in purchases from educated and higher income groups. People are reluctant to shop online because they are worried about their security. People are often reluctant to adapt as a result of the technological sophistication involved in making internet purchases. Businesses engaged in online shopping ought to concentrate on fostering reliable connections between suppliers and consumers.

With the help of this research, we aim to better understand how the various attributes relate to one another and do sentiment analysis on the dataset of product reviews, which will be worthwhile from both the standpoints of consumers and designers of products. The evaluations can be calculated to give the designer the most relevant data, leading him to improve the product or introduce a new one that best meets the needs of customers. The experimental results showed that our model was satisfactory in all the measurement metrics. This paper presents an experimental study along with a proposed model through Naïve-Bayes algorithm on dataset of Product reviews to measure the polarity of reviews whether positive or negative and words related to the products. This extensibility in the research will be of great benefit to the industry and will make requirement gathering much less time-consuming, shrinking down the expenses incurred using surveys, questionnaires, interviews, market research and trends.

CHAPTER-6 Future Scope

Future work can be concentrated on mining reviews from multiple websites and multiple products etc. The same work can be extended to incorporate many more classification algorithms which will help us to decide or to choose the best classifier for opinion mining and sentiment analysis. With the designer's concerns in mind, this research can be expanded to mine client requirements. To give the designer, the most valuable information possible to improve the product or introduce a new product to the market while satisfying the greatest number of consumer criteria, the helpfulness of reviews can be calculated. The industry will greatly benefit from this research's flexibility, which will save the time and costs associated with obtaining requirements through surveys, questionnaires, interviews, market research, and trends.

References

[1] J. P. Singh, N. P. Rana, and W. Alkhowaiter, “Sentiment analysis of products’ reviews containing English and Hindi texts,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015. doi: 10.1007/978-3-319-25013-7_33.

[2] Z. Singla, S. Randhawa, and S. Jain, “Statistical and sentiment analysis of consumer product reviews,” in *8th International Conference on Computing, Communications and Networking Technologies, ICCCNT 2017*, 2017. doi: 10.1109/ICCCNT.2017.8203960.

[3] S. Elzeheiry, W. A. Gab-Allah, N. Mekky, and M. Elmogy, “Sentiment Analysis for E-commerce Product Reviews: Current Trends and Future Directions,” May 2023, doi: 10.20944/PREPRINTS202305.1649.V1.

[4] Dr. R Jayanthi, “A study on customer’s review about online shoppers in India,” National Journal of Multidisciplinary Research and Development ISSN: 2455-9040. Accessed: Feb. 19, 2024. [Online]. Available: https://www.multidisciplinaryjournal.org/assets/archives/2018/vol3issue1/3-1-119-497.pdf

[5] S. Al-Otaibi and A. Al-Rasheed, “A Review and Comparative Analysis of Sentiment Analysis Techniques,” *Informatica (Slovenia)*, vol. 46, no. 6. Slovene Society Informatika, pp. 33–44, 2022. doi: 10.31449/inf.v46i6.3991.

[6] J. Guerreiro and P. Rita, “How to predict explicit recommendations in online reviews using text mining and sentiment analysis,” *Journal of Hospitality and Tourism Management*, vol. 43, 2020, doi: 10.1016/j.jhtm.2019.07.001.

[7] S. Raghav and J. Sondhi, “Online Product Reviews on Consumer Behaviour With Special India”.

[8] K. Thriveni Kumari, “New Research in E-commerce Business Management,” 2023.

[9] K. L. Santhosh Kumar, J. Desai, and J. Majumdar, “Opinion mining and sentiment analysis on online customer review,” in *2016 IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2016*, 2017. doi: 10.1109/ICCIC.2016.7919584.

[10] Y. Kai, Y. Cai, H. Dongping, J. Li, Z. Zhou, and X. Lei, “An effective hybrid model for opinion mining and sentiment analysis,” in *2017 IEEE International Conference on Big Data and Smart Computing, BigComp 2017*, 2017. doi: 10.1109/BIGCOMP.2017.7881759.

[11] S. A. Sadhana, L. Sairamesh, S. Sabena, S. Ganapathy, and A. Kannan, “Mining target opinions from online reviews using semi-supervised word alignment model,” in *Proceedings - 2017 2nd International Conference on Recent Trends and Challenges in Computational Models, ICRTCCM 2017*, 2017. doi: 10.1109/ICRTCCM.2017.66.

[12] P. K. Singh, A. Sachdeva, D. Mahajan, N. Pande, and A. Sharma, "An approach towards feature specific opinion mining and sentimental analysis across e-commerce websites," in *Proceedings of the 5th International Conference on Confluence 2014: The Next Generation Information Technology Summit*, 2014. doi: 10.1109/CONFLUENCE.2014.6949312.

[13] P. Porntrakoon and C. Moemeng, "Thai sentiment analysis for consumer's review in multiple dimensions using sentiment compensation technique (SenSecomp)," in *ECTI-CON 2018 - 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, 2018. doi: 10.1109/ECTICon.2018.8619892.

[14] J. N. L, S. K. Kumar L, and J. Majumdar, "Sentiment Analysis of Customer Reviews on Laptop Products for Flipkart," *International Research Journal of Engineering and Technology*, 2018.

[15] A. Dadhich and B. Thankachan, "Social & Juristic challenges of AI for Opinion Mining Approaches on Amazon & Flipkart Product Reviews Using Machine Learning Algorithms," *SN Comput Sci*, vol. 2, no. 3, 2021, doi: 10.1007/s42979-021-00554-3.

[16] A. Sinha, N. Arora, S. Singh, M. Cheema, and A. Nazir, "Fake Product Review Monitoring Using Opinion Mining," *International Journal of Advances in Engineering and Management (IJAEM*, vol. 2, 2008.

[17] D. M. A. Marathe, "'Literature Review On Customer Satisfaction Towards Online Shopping - With Special Reference To Flipkart.,'" *Dogo Rangsang Research Journal*, vol. 10, no. 06, 2020.

[18] E. Tyagi and A. K. Sharma, "Sentiment Analysis of Product Reviews using Support Vector Machine Learning Algorithm," *Indian J Sci Technol*, vol. 10, no. 35, 2017, doi: 10.17485/ijst/2017/v10i35/118965.

[19] U. Kanchan, N. Kumar, and A. Gupta, "A STUDY OF ONLINE PURCHASE BEHAVIOUR OF CUSTOMERS IN INDIA," 2015, doi: 10.21917/ijms.2015.0019.

[20] Z. Singla, S. Randhawa, and S. Jain, "Sentiment analysis of customer product reviews using machine learning," in *Proceedings of 2017 International Conference on Intelligent Computing and Control, I2C2 2017*, 2018. doi: 10.1109/I2C2.2017.8321910.

[21] G. K. Soor, A. Morje, R. Dalal, and D. Vora, "Product Recommendation System based on User Trustworthiness & Sentiment Analysis," *ITM Web of Conferences*, vol. 32, 2020, doi: 10.1051/itmconf/20203203030.

[22] T. Keerthana, T. Bhavani, N. Suma Priya, V. Sai Prathyusha, and K. Santhi Sri, "Flipkart product recommendation system," *Journal of engoneering science*, vol. 11, no. 4, 2020.

[23] P. Rakshit, P. K. Srivastava, M. Afjal, and S. K. Srivastava, "Sentimental Analytics on Indian Big Billion Day of Flip Kart and Amazon," *SN Comput Sci*, vol. 2, no. 3, 2021, doi: 10.1007/s42979-020-00441-3.

[24] S. Vijayakumar, G. Vidyashankar, R. Venkatesakumar, S. Madhavan, and S. Riasudeen, "Online

Review Characteristics and Information Asymmetry Is it easy to switch between Online Shopping sites? A Case Study of Reviews from Amazon and Flipkart," *SDMIMD Journal of Management | Print* |, vol. 12, 2021, doi: 10.18311/sdmimd/2021/26704.

[25] K. Wase *et al.*, "Sentiment analysis of product review," *International Journal of Innovations in Engineering and Science*, vol. 3, no. 5, pp. 2456–3463, 2018, Accessed: Nov. 30, 2023. [Online]. Available: www.ijies.net

[26] P. Kumari, "Comparing Customer Attitude towards Amazon and Flipkart in Patna," *International Research Journal on Advanced Science Hub*, vol. 2, no. Special Issue ICSTM 12S, 2020, doi: 10.47392/irjash.2020.252.

[27] S. Janakiballav Mishra, S. Sangita Ray, and R. Kantha, "Consumer Satisfaction Level of E-Commerce with Special Reference to Flipkart in Bhubaneswar," 2021, Accessed: Feb. 25, 2024. [Online]. Available: www.ijaresm.com

[28] A. Rohanroy, K. Rathor, B. Jannawat, and S. Sethi, "REVIEWING FLIPKART PRODUCT COMMENTS USING METHODS BASED ON SENTIMENT ANALYSIS", Accessed: Feb. 25, 2024. [Online]. Available: www.irjmets.com

[29] S. Kumar, B. Hooi, M. Kumar, C. Faloutsos, and V. S. Subrahmanian, "FairJudge: Trustworthy User Prediction in Rating Platforms".

[30] A. Alsaeedi and M. Z. Khan, "A study on sentiment analysis techniques of Twitter data," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 2, 2019, doi: 10.14569/ijacsa.2019.0100248.

THANK YOU