International Journal of Innovation Studies

IMPROVING SENTIMENT ANALYSIS IN PRODUCT REVIEWS THROUGH PREPROCESSING AND CLASSIFICATION TECHNIQUES

S. Sheeladevi
Research Scholar, Department of Computer Science, Dr. M.G.R Educational and Research Institute, Maduravoyal, Chennai, Tamilnadu, India, E-Mail: sheelababubca12@gmail.com[1]

Dr. Viji Vinod
Professor & Head, Department of Computer Applications, Dr. M.G.R Educational and Research Institute, Maduravoyal, Chennai, Tamilnadu, India
E-Mail: vijivinod.mca@drmgrdu.ac.in[2]

Abstract:- *In the digital marketplace, product reviews are vital in forming consumer decisions and brand reputation. Because there is so much user-generated information on the internet, it is now crucial for businesses looking to understand customer sentiment better and enhance their offerings to analyze and categorize these evaluations effectively. This study proposes a Support Vector Machine (SVM) based model for classifying product reviews evaluated on the Amazon Product Reviews dataset from Kaggle. To improve classification accuracy, the model uses an extensive preprocessing pipeline that includes text normalization, tokenization, padding, word embeddings, and aspect extraction techniques. These approaches are used to discover important elements of the reviews. These elements are mixed with the review text to create enriched feature vectors, which provide the SVM classifier with better input. According to experimental results, this model outperforms conventional techniques regarding accuracy, precision, and recall, greatly improving classification performance. This hybrid strategy, which combines SVM and aspect extraction, presents a viable way to analyze sentiment in product reviews.*

Keywords*:- Product Review Classification, Support Vector Machine, Aspect Extraction, Text Preprocessing, Word Embeddings, Tokenization, Precision, Recall, Accuracy, Amazon Product Reviews, Natural Language Processing.*

1. INTRODUCTION

Product reviews play a vital role in shaping consumer decisions and brand reputations in today's digital era. With the exponential growth of e-commerce platforms, consumers are increasingly relying on online reviews to assess the quality and credibility of products before making purchases. This massive pool of user-generated content offers valuable insights into customer satisfaction, preferences, and areas for improvement. However, analyzing large volumes of reviews manually is impractical, making automated sentiment analysis a crucial tool for extracting meaningful information. By leveraging natural language processing (NLP) and machine learning techniques, sentiment analysis can classify customer opinions as positive, negative, or neutral, enabling businesses to understand customer sentiment effectively and make informed decisions.

The COVID-19 pandemic has led to a substantial increase in online purchasing, with many governments ordering their residents to stay at home as a result. Online shops frequently ask for textual evaluations and/or ratings from customers regarding their goods and services (Zhang J et al., 2021). Sentiment analysis offers a quick and simple way to classify reviews on online platforms, such as social media, which helps both buyers and sellers gain valuable insights into what customers are saying about the goods and services (Balakrishnan V et al., 2021).
Customers assess goods and services using several criteria. Reviews and advertising can all be used as evaluation tools. One of the most important variables influencing the sales of goods and services is reviews. Reviews boost trust between customers and companies in the e-commerce sector and help reduce the risk of being duped (Upendra Singh et al., 2022). NLP enters the scene here. To make sense of the text and enable machine learning algorithms to interpret and categorize the messages, it converts human language from text to arrays of numbers (D. Dharrao et al., 2023)

Classifying product reviews provides significant advantages for businesses and consumers alike. For businesses, it enables a deeper understanding of customer sentiments and preferences, helping to identify areas of satisfaction and dissatisfaction. By analyzing reviews, companies can uncover trends in consumer feedback, such as frequently mentioned product features or recurring issues, allowing them to make targeted improvements. Additionally, review classification supports the development of better marketing strategies by highlighting aspects of the product that resonate most with customers, thereby enhancing customer satisfaction and brand loyalty.

For consumers, product review classification simplifies the decision-making process. With the sheer volume of reviews available online, sorting through them can be overwhelming. Sentiment classification tools aggregate reviews and provide a summary of overall opinions, enabling customers to assess product quality quickly. This reduces the time spent evaluating options and fosters more informed purchasing decisions. Moreover, review classification can highlight common user concerns or praise, ensuring that potential buyers have a balanced view of the product before committing to a purchase. This dual benefit makes product review classification an essential tool for improving transparency and efficiency in e-commerce ecosystems.

Businesses can track the success of their products, identify possible problems, and react instantly to client input when reviews are classified accurately. Better prediction results can be achieved by developing more accurate classification models by honing in on input feature extraction and identifying salient features from the reviews. To increase the recall, accuracy, and precision of these classifications, this article investigates the use of ML models to categorize product evaluations according to consumer feelings. Through the use of thorough preprocessing procedures and the extraction of significant elements from review texts, we improve the caliber of input supplied to the classification model, which in turn results in more efficient sentiment analysis. The performance of this method is assessed using the Amazon Product Reviews dataset.

2. LITERATURE REVIEW

These days, advances in data analytics have made it possible to identify underlying trends using efficient computer models. End consumers provide a vast amount of online product reviews on various social media and e-commerce platforms, which greatly aid developers in their product creation process by providing invaluable insights. A brand-new cluster-based categorization algorithm for online product reviews was presented by P. Vijayaragavan et al. in 2020. Many processes make up the model that is being shown. Initially, a classification model based on SVM is used to categorize the product reviews. Next, a confusion matrix is created to account for the probability that each customer will buy the product. Next, the supplied data is clustered into two groups using the K-means clustering approach. Sentiment analysis is the method used in the next stage of obtaining the features. Lastly, fuzzy-based soft set theory is used to assess the customer's likelihood of successfully making a transaction. The iPod dataset is used for the experimental validation of the proposed model. The results of the simulation demonstrated the better qualities of the model that were presented in several ways.

Client fulfillment is the most crucial factor to take into account when starting a successful business. Only when customers are satisfied with a product or service and are willing to suggest it to others can a business expand. To better understand their consumers' perspectives, businesses also gather reviews and ratings from their clientele. Another enormous company that cherishes the opinions of its clients is Amazon. A thorough manual scan of every review for every product on the website would be extremely challenging, if not impossible. Natural language processing, or NLP, enters the frame here. To make sense of the text and enable machine learning algorithms to process and categorize the messages, it converts human language from text to arrays of numbers. Amazon reviews of different products and categories are processed by D. Dharrao et al., 2023, and the results are ranked from worst to best in five categories. The reviews were processed using supervised ML methods, including SVM, NBC(Naïve Bayes classifier), and DT(Decision Tree) classifier. Accuracy score bar charts have also been made to analyze and compare these methods.

Client instruction can be facilitated by the use of online product reviews, as evidenced by their effectiveness as a source of information. Online shoppers can often score things both numerically and textually, leaving comments about what they think they may have bought. A deeper awareness of market needs could help designers boost the likelihood that their products will succeed in the market. This information-rich data source could help with more than just employing online product review platforms as patient decision support systems. However, manually analyzing such data is becoming more difficult due to the growing amount and complexity of products on the market. In Abhinav Singh et al.'s 2017 study, the goal is to: i) develop an ML approach that distinguishes between three direct product characteristics and two indirect product characteristics ii) identify the ML algorithm that produces the best and most generalizable results in accomplishing the goal. iii) to measure the relationship between product ratings and attributes that are either direct or indirect. To support the validity of the suggested approach, a case study containing review data for products mined from e-commerce websites is shown. A multilayered validation strategy was introduced to investigate the generalisability of the suggested approach. The ML model that was developed as a result achieved classification accuracies of 82.44% within products, 80.84% across products, 79.03% across product types, and 80.64% across product domains. By quantifying the relationship between product reviews and qualities, the scientific contributions of this work have the potential to change how buyers and product designers integrate reviews into their methods for making decisions.

According to Prathyakshini et al. 2022, sentiment analysis is widely employed in a variety of industries, such as product analysis to determine the wants of the consumer. The notion of sentiment analysis is widely used in other fields, such as stock analysis, government, and healthcare. Genuine product reviews are posted by customers on internet purchasing portals. Additionally, it will be simple for customers to see the product reviews and ratings before deciding which ones to purchase. Sentiment classification is an analysis technique used to categorize user reviews into positive and negative viewpoints. It assists in locating the product's problems so that they can be fixed. However, product reviews encourage buyers to purchase the goods based on their evaluation, which also allows business owners to make adjustments. There are occasions when it's challenging to classify the text found in product reviews. Classification algorithms are a useful tool for accomplishing this. Several features are available, of which TF-IDF and N-Gram were chosen. Based on the results, it is clear that RF(Random Forest) outperforms the other approaches in the TF-IDF and N-Gram models for the product types of electronics, health and beauty, and accessories and apparel.

Almost all e-shopping activities are supported by online product reviews. The demand for automatic methods for prioritizing informative information is increasing due to the large amount of data and the wide range in quality of internet reviews. The reasoning underlying the selection of particular features is largely unexplored, despite a significant body of studies on review helpfulness prediction. Wider generalization is lacking in Jiahua Du et al., 20219 since they primarily focus on domain-and/or platform-dependent feature curation. Furthermore, the problem of reproducibility and result comparability arises from the frequent loss of data and source code. The most thorough feature identification, assessment, and choice is used in this study to fill in the gaps. To achieve this, 149 pertinent study papers are first used to identify the 30 most popular content-based traits, which are then categorized into five logical groups. Next, the most extensive publicly accessible Amazon 5-core dataset's six domains are used to do prediction using the attributes that were chosen. Features within each category, all features, and individual features are the three situations for feature selection that are taken into consideration. Empirical findings show that sentiment and other variables come in second and third in terms of predictive power for informative reviews, with semantics playing a major role. In today's common e-commerce market, feature combination patterns and selection criteria across domains are finally summarised to improve the customer experience. To improve result comparability and repeatability, the study's computational framework for usefulness prediction has been made public.

The issue of excessive information on internet review platforms has severely hindered the ability of many consumers to assess the caliber of goods or companies before making a purchase. A substantial amount of research has been done to forecast the value of online customer evaluations, with varying degrees of success being noted for different strategies. Furthermore, a lot of the current solutions limit generalization by utilizing handmade features and conventional ML techniques. The goal of Muhammad Bilal et al., 2022 is to improve the BERT basic model to suggest a more comprehensive strategy. The efficacy of BERT-based classifiers is then assessed by contrasting their performance with that of bag-of-words techniques. Experiments conducted on Yelp retail reviews demonstrate that optimized BERT-based classifiers perform better in identifying helpful and harmful reviews than bag-of-words methods. Furthermore, it is discovered that the BERT-based classifier's sequence length significantly affects the effectiveness of classification.

A thorough literature assessment of sentiment analysis techniques, applications, and difficulties was presented by Yanying Mao et al. in 2023. The purpose of the sentiment analysis task is explained, different methods are compared, the application domains of sentiment analysis are examined, the difficulties and constraints faced by researchers are highlighted, potential solutions are suggested, and future research directions are examined in this thorough literature review. The results of the study demonstrate the significance of sentiment analysis in people's lives and in their work, as well as the important role that AI technologies play in automatic text sentiment analysis. In addition to adding to the body of knowledge already available on sentiment analysis, this study offers references to help practitioners and academics select appropriate methodologies and best practices for sentiment analysis.

Effective competitive analysis, when incorporated into a business plan, enables companies to surpass their rivals and draw in devoted customers. Sentiment analysis can be used to investigate competitors, find market conditions, and gauge interest in particular issues as part of competitive research. Several fields have seen improvements in performance due to AI, most notably sentiment analysis. Sentiment analysis is the method of identifying emotions in text using AI. It can understand a statement's tone rather than only identifying whether a given word in a passage of text has a positive or negative meaning. Papers from 2012 to 2022 that address how competitive market research finds and contrasts key market metrics that aid in differentiating the products and services of rivals are reviewed by Hamed Taherdoost et al. in 2023. Sentiment analysis driven by AI can be used to find out what customers think of competitors in all business domains.

3. PROPOSED MODEL

The application of sentiment analysis in product reviews goes beyond simple categorization. It enables businesses to track customer feedback trends, identify recurring issues, and improve product features or services accordingly. Additionally, this technology benefits consumers by helping them quickly gauge the overall perception of a product, saving time and effort in decision-making. As the field advances, sentiment analysis continues to grow in accuracy and sophistication, providing deeper insights into consumer behavior and offering a competitive edge to businesses in highly dynamic markets. This capability underscores its importance in enhancing customer satisfaction and driving data-driven strategies for success.

3.1 Dataset Collection

The process starts with the **Amazon Product Reviews dataset**. This dataset contains customer reviews of various products on Amazon. It consists of textual data, which includes feedback, ratings, and opinions about the products. The initial step involves obtaining the Amazon Product Reviews dataset from Kaggle. This dataset has a substantial number of customer reviews accompanied by labels, such as positive or negative comments. The variety of product categories and consumer attitudes in this dataset makes it a perfect platform for testing and verifying the categorization algorithm. The review text is the main focus since it is the main source of information for the categorization process. The ML model is trained and assessed using the sentiment analysis of these texts as the basis for labeling.
The goal of this workflow is to process this data for sentiment analysis or other forms of product review classification. Fig.1 represents a process for analyzing the Amazon Product Reviews dataset using machine learning, particularly for classification tasks. Here's a detailed breakdown of each stage in the workflow:

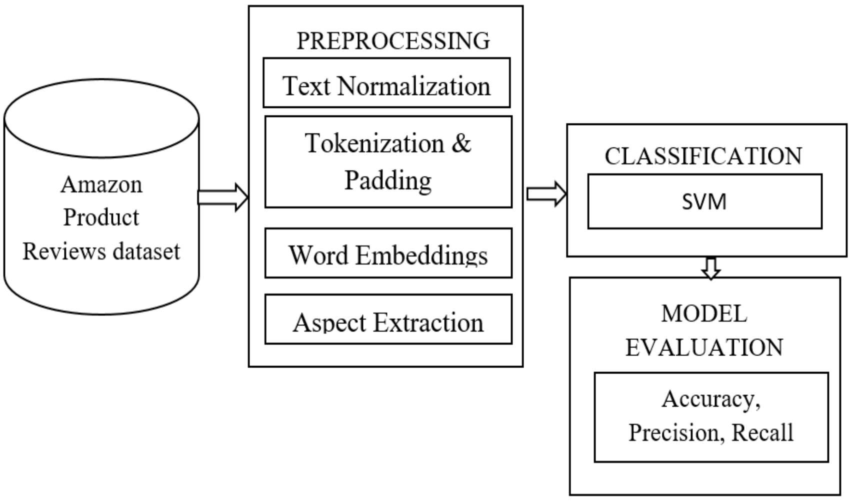

Fig 1 Architecture of Proposed Product Review Classification

3.2 Preprocessing

The next stage is preprocessing, where the raw text data is transformed into a format that can be used for machine learning models. Preprocessing plays a pivotal role in the accuracy and efficiency of product review classification using sentiment analysis. Raw text data from reviews often contains noise, such as typos, irrelevant information, slang, and special characters, which can hinder the performance of machine learning models. Preprocessing involves cleaning and organizing this data to ensure that it is suitable for analysis. Techniques such as tokenization, stop-word removal, stemming, and lemmatization help in reducing redundancy and standardizing the text. This process enhances the representation of the data, allowing algorithms to focus on meaningful patterns rather than irrelevant features.
Moreover, preprocessing reduces computational complexity by minimizing the dimensionality of the data and improving the quality of input features. For example, converting text into a structured format, such as bag-of-words or word embeddings, allows models to interpret the semantic and syntactic relationships effectively. Additionally, handling domain-specific challenges, like extracting relevant aspects from reviews or normalizing sentiment-laden words, ensures that the classification process is both robust and context-aware. Preprocessing

is a crucial step that lays the foundation for accurate sentiment classification, directly impacting the reliability of insights derived from product reviews.
Preprocessing is a critical step in ensuring the accuracy and reliability of product review classification. This stage includes several sub-processes:

Text Normalization:
The following stage after obtaining the dataset is preprocessing, which begins with text normalization. To ensure uniformity and consistency in the input data, all review text must be converted to lowercase. Special letters, numerals, and punctuation are eliminated to reduce noise that could impede the model's ability to learn. Furthermore, lemmatization or stemming is used to reduce words to their most basic form, guaranteeing that distinct grammatical variations of a word are handled equally. The text is further refined by eliminating common but uninformative words like "the," "and," or "is," leaving just the most pertinent terms for analysis.

$$Normalized\ Text = RemoveStopWords\left(Lemmatize\left(RemoveSpecialChars\left(Lowercase(T)\right)\right)\right) - - - (1)$$

The original text is T.
The function $Lowercase(T)$ changes every character in T to lowercase.
$RemoveSpecialChars\left((T)\right)$Eliminates all special characters, numerals and punctuation from T.
$Lemmatize(RemoveSpecialChars(T))$returns a term to its most basic form.
Common terms that are not helpful for analysis are eliminated using $RemoveStopWords(T)$

Tokenization & Padding:
Tokenization is the process of breaking the text into smaller chunks, typically words or subwords, so the model can process them more effectively. After tokenization, **padding** is applied to ensure all input sequences are of equal length, which is essential for models that require fixed-length input, such as neural networks. The text is tokenized—that is, it is divided into discrete words or tokens—after normalization. The process converts each review into a series of tokens, which the model uses to do word-level text analysis. Padding is used to guarantee that every token sequence has a consistent length because the reviews differ in duration. Longer sequences are shortened to a predetermined maximum length, while shorter sequences are padded with zeros. By ensuring that the data is in the right format, this phase helps the classification model process each input in a uniform manner.

$$Tokens(T) = [w_1, w_2, w_3, \ldots\ldots w_n] - - - (2)$$

where$w_1, w_2, w_3, \ldots\ldots w_n$ are the individual tokens (words) in the text, and T is the initial text.

$$Padded\ Sequence(T) = \begin{cases} [w_1, w_2, \ldots\ldots\ldots\ldots w_n] & if\ n = maxlen \\ [w_1, w_2, \ldots\ldots, w_n, 0, 0, \ldots] & if\ n < maxlen \\ [w_1, w_2, \ldots\ldots\ldots, w_{maxlen}] & if\ n > maxlen \end{cases} - - - (3)$$

Where:
T represents the tokenized sequence, while n denotes the sequence's length.

The predetermined maximum sequence length is $maxlen$. For shorter sequences, the padding is represented by the number 0.

Word Embeddings:
The next stage is to use word embeddings that have already been trained, like GloVe, to map the tokenized words into dense vectors. Word embeddings capture the semantic relationships between words by representing them in a continuous vector space. For instance, because of their similar meanings, the vector representations of words like "good" and "excellent" will be similar. The model can comprehend the reviews better because of these embeddings, which give words rich, meaningful representations. The model may make use of pre-existing word semantic knowledge by utilizing pre-trained embeddings, which improves the machine's accuracy in classifying reviews.
$Embedding(w_i) = [v_1, v_2, v_3.,,,, v_d] - - - (4)$
Where:

- w_i is the i-th word in the text.
- $[v_1, v_2, v_3.,,,, v_d]$ is the dense vector representation of the word w_i, and d is the embedding dimension

The average embedding of every word in the sequence can be computed to represent a whole sentence or review:

$$Sentence\,Embedding\ (S) = \frac{1}{n}\sum_{i=1}^{n} Embedding(w_i) - - - (5)$$

Where:
The sentence or review is denoted by S, and its word count is indicated by n. The embedding of every word w_i is called $Embedding(w_i)$.

Aspect Extraction:
Aspect extraction is used to pinpoint important details or subjects in the reviews. This stage aids the model in concentrating on particular features of a product that reviewers have emphasized, including "quality," "price," or "customer service." Using topic modeling approaches, such as LDA, which automatically finds and groups similar terms into separate aspects, is one popular method for aspect extraction. By adding content to the review and enhancing the input data with details about the product attributes that consumers are talking about, extracting aspects aids in the classifier's ability to distinguish between different sentiment aspects.
$Aspect\ (T) = [a_1, a_2, a_3 \dots\dots\ a_k] - - - (6)$
Where:
• The input text is T (the review).
• The retrieved aspects/topics are $[a_1, a_2, a_3 \dots\dots\ a_k]$.
• k is the total number of recognized factors.

Another key benefit of preprocessing is its role in reducing the computational burden of analyzing large datasets. By filtering out irrelevant information and structuring the data into formats such as term frequency or word embeddings, the complexity of the analysis is minimized. Preprocessing also ensures that domain-specific nuances are addressed, such as handling slang, abbreviations, or product-related keywords. These refinements contribute to creating a foundation for more precise sentiment classification, enabling better insights into customer opinions and improving decision-making processes for businesses.

3.3 Classification

SVM Model Training:

The SVM classifier is trained using preprocessed, tokenized, embedding, and aspect feature-enhanced reviews. The SVM model is a potent classification technique that operates by determining the best hyperplane in a high-dimensional space to divide various classes. The SVM model is better able to discern between the various feelings stated in the reviews by integrating the aspect-extracted features and word embeddings. The labeled review data is sent into the SVM throughout the training process so that it may discover the patterns and connections that distinguish between positive and negative comments.

Review classification using SVM involves determining the best hyperplane to divide reviews into different groups. Regarding a certain review denoted by feature vector x:

$$f(x) = w^T x + b - - - (7)$$

Where:

- w is the weight vector.
- x is the feature vector
- b is the bias term.
- $f(x)$ is the decision function.

The classification decision is made by:

$$Label = \begin{cases} +1 \ if \ f(x) > 0 \\ -1 \ if \ f(x) < 0 \end{cases} - - - (8)$$

Using SVM for product review classification offers several advantages due to the algorithm's robust performance in handling textual data. One of its key benefits is its ability to work effectively with high-dimensional datasets, which is common in natural language processing tasks where each word in the vocabulary may represent a feature. SVM excels at finding the optimal hyperplane that separates data into distinct classes, making it highly effective for binary or multi-class classification problems. Additionally, SVM is versatile, as it can use different kernel functions (linear, polynomial, or radial basis function) to capture complex relationships in the data, which is particularly useful when the reviews exhibit non-linear patterns.

Another significant advantage of SVM is its resilience to overfitting, especially in cases where the number of features is much larger than the number of training samples. This makes it a great choice for small or moderate-sized datasets, which is often the case when working with product reviews in niche categories. Furthermore, SVM models are computationally efficient for smaller datasets and provide high accuracy, precision, and recall in classifying sentiments or product attributes. These characteristics make SVM a reliable and efficient option for businesses aiming to analyze customer feedback and gain actionable insights from product reviews.

4. RESULTS AND DISCUSSION

4.1 Dataset Description

The **Amazon Product Reviews dataset** contains comprehensive information about user interactions with various products on Amazon, with a size of 318.77 MB. It includes key attributes such as **userId**, which uniquely identifies each user, and **productId**, which uniquely identifies each product reviewed. The dataset also records the **Rating**, indicating the score each user assigned to a product, and the **timestamp**, which provides the exact time the rating was submitted. These attributes make the dataset ideal for analyzing user behavior, product popularity, and sentiment trends over time, enabling advanced review classification and recommendation system research.

4.2 Model Evaluation

Using a different testing dataset, the trained SVM model's performance is assessed using important metrics like accuracy, precision, recall, and F1-score. While precision and recall offer information on how successfully the model distinguishes between good and negative evaluations without creating an excessive number of false positives or false negatives, accuracy gauges the overall validity of the model's predictions. A fair metric that takes into account recall and precision is the F1 score. These metrics allow for the evaluation of the model's efficacy and the comparison of its results with those of other methods, showcasing the advantages of aspect extraction and improved preprocessing.

4.2.1 Accuracy Analysis

In product review classification, accuracy is a metric that evaluates how often the model correctly classifies reviews into their respective categories, such as positive, negative, or neutral. It represents the ratio of correctly predicted reviews to the total number of reviews analyzed. Essentially, accuracy indicates the model's overall ability to assign the correct label to the reviews.

The model's overall correctness in categorizing product reviews is measured by accuracy.

$$Accuracy = \frac{Number\ of\ Correctly\ Classified\ Reviews}{Total\ Number\ of\ Reviews} - - - (9)$$

Where:

- **Number of Correctly Classified Reviews** = True Positives (TP) + True Negatives (TN)
- **Total Number of Reviews** = TP + TN + False Positives (FP) + False Negatives (FN)

4.2.2 Precision Analysis

The precision metric quantifies the percentage of genuinely favorable reviews that are categorized as such.In product review classification, **precision** measures the accuracy of the model's positive predictions. It calculates the percentage of reviews that were correctly classified as positive (or any specific target class) out of all the reviews the model predicted as positive. This metric focuses on how many of the model's positive predictions are actually correct.

$$Precision = \frac{TP}{TP + FP} - - - (10)$$

True Positives (TP) = Number of positive reviews correctly classified as positive

False Positives (FP) = Number of negative reviews incorrectly classified as positive

4.2.3 Recall Analysis

In product review classification, **recall** measures the model's ability to correctly identify all the reviews belonging to a specific category, such as positive or negative. It represents the proportion of correctly predicted positive reviews compared to the total actual positive reviews in the dataset. Recall focuses on capturing all relevant instances, ensuring that the model doesn't miss many true positives.

The recall metric quantifies the percentage of real positive reviews that the model accurately detects.

$$Recall = \frac{TP}{TP + FN} - - - (11)$$

Where:

- **True Positives (TP)** = Number of positive reviews correctly classified as positive
- **False Negatives (FN)** = Number of positive reviews incorrectly classified as negative

The effectiveness of several machine learning models for classifying product reviews using SVM in conjunction with various preprocessing methods is shown in Table 1. Three primary criteria are used to assess the models: recall, accuracy, and precision. Let's examine the outcomes and provide justifications for every possible pairing.

Table 1 ML Models Vs Metrics

ML Model/Metrics	Accuracy	Precision	Recall
Text Normalization+SVM	89.50	0.87	0.85
Tokenization & Padding + SVM	90.25	0.88	0.87
Word Embeddings + SVM	91.42	0.89	0.87
Aspect Extraction + SVM	93.50	0.91	0.89

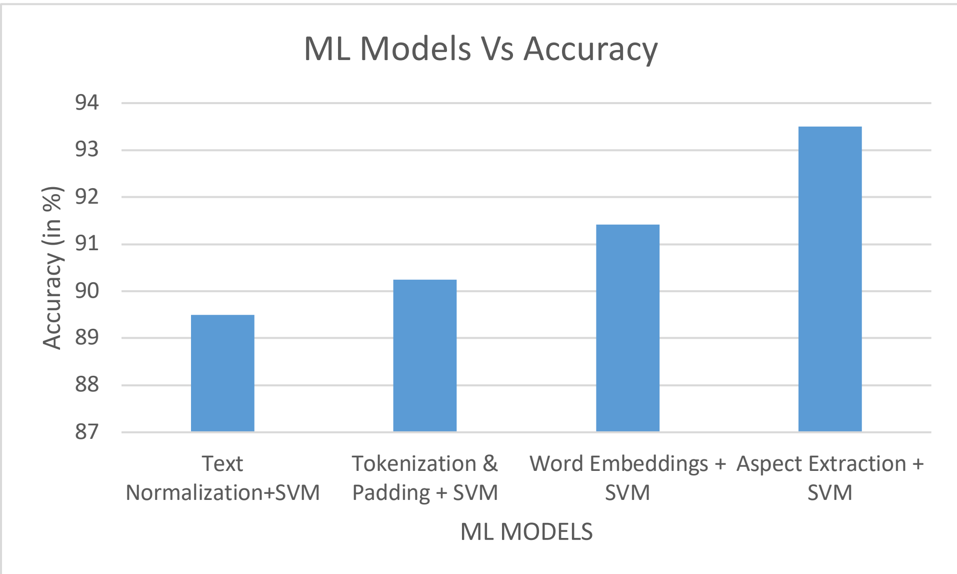

Fig 2 Comparison of Accuracy with ML Models

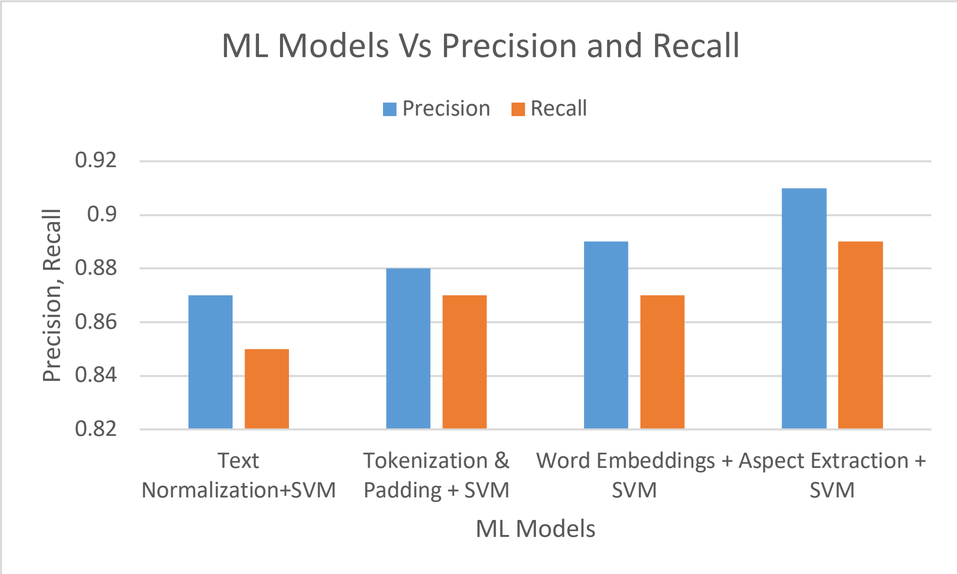

Fig 3 Comparison of Precision and Recall with ML Models

When text normalization (such as changing text to lowercase or eliminating stopwords) is combined with SVM, the results are reasonably good. 89.5% of reviews are correctly classified, and precision and recall are balanced well. There is potential for improvement, particularly in the recall rate, which suggests that some good evaluations may have been mistakenly labeled as negative.

Accuracy, precision, and recall are marginally increased by adding tokenization and padding, which divide the text into discrete tokens and guarantee consistent sequence lengths. In comparison to the previous model, this one performs better since it can better capture the text's structure, as evidenced by its 90.25% accuracy and higher recall.

The model learns more about the relationships between words and becomes more accurate and precise by using word embeddings, which are vector representations of a word's semantic meaning. Using embedding vectors improves the model's classification performance to 91.42%, particularly in differentiating between comparable phrases (e.g., "good" and "great").

The best results are obtained when SVM is paired with aspect extraction, which identifies important features like "price" or "battery". The model can focus on particular features and enhance its classification abilities by isolating important components of the review. This allows the model to achieve the highest accuracy (93.5%) and the best balance between precision and recall (91% precision, 89% recall). This suggests that aspect extraction lowers classification errors by enabling the model to concentrate on crucial product-specific information. These findings demonstrate how much better preprocessing techniques, such as aspect extraction and word embeddings, enhance the SVM classifier's performance.

5. CONCLUSION AND FUTURE WORK

This work used SVM to classify product reviews utilizing a variety of text preprocessing approaches, including Text Normalisation, Tokenisation & Padding, Word embedding, and Aspect Extraction. The Amazon Product Reviews dataset from Kaggle was used for the tests, and accuracy, precision, and recall measures were used to assess each model's performance. The greatest results were obtained by combining Aspect Extraction with SVM, which yielded 93.50% accuracy, 91% precision, and 89% recall, respectively. This enhancement demonstrates how useful it is to pinpoint review characteristics that provide the model with more significant features to work with when making predictions, such as cost, value, or battery life. The findings demonstrate that aspect-based models are more effective at capturing insights unique to a product, which improves classification results.

There are other directions to pursue in this area in the future. First, adding transformer-based architectures (like BERT) or DL-based models (like LSTM) could increase the model's performance in classification and further improve its comprehension of context. Furthermore, adding sentiment analysis at the aspect level could aid in improving the classification by allowing the model to determine which aspects are perceived favorably or unfavorably in addition to determining if a review is good or negative. Furthermore, expanding the model's testing to include a wider range of datasets and adding multilingual support could make it more resilient and applicable to a wider range of e-commerce platforms. Finally, by using this approach for sentiment-driven recommendation engines or review summarisation, real-time systems could help customers receive more individualized and succinct product insights.

REFERENCES

[1] D. Dharrao, S. Deokate, A. M. Bongale &nd Urolagin(2023), "E-Commerce Product Review Classification based on Supervised Machine Learning Techniques," *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2023, pp. 1934-1939, doi: 10.1109/ICACCS57279.2023.10112717.

[2] Abhinav Singh & Conrad S. Tucker (2017), "A machine learning approach to product review disambiguation based on function, form and behavior classification", Elsevier, Science Direct, Decision Support Systems, Vol. 97, pp. 81-91

[3] Prathyakshini; Preethi Salian K; Puneeth B. R; Tanzila Nargis; Supriya Salian(2022), "A Novel Approach for Classification of Online Product Reviews using various Machine Learning Techniques", *022 6th International Conference on Electronics, Communication and Aerospace Technology*, Coimbatore, India, 2022, pp. 878-884, doi: 10.1109/ICECA55336.2022.10009447.

[4] https://www.kaggle.com/datasets/saurav9786/amazon-product-reviews

[5] P. Vijayaragavan R. Ponnusamy b, M. Aramudhan(2020), "An optimal support vector machine based classification model for sentimental analysis of online product reviews", Future Generation Computer Systems", Vol. 111, pp. 234-240

[6] Zhang J, Zhang A, Liu D& Bian Y(2021), " Customer preferences extraction for air purifiers based on fine-grained sentiment analysis of online reviews", Knowl-Based Syst. 2021 doi: 10.1016/j.knosys.2021.107259.

[7] Balakrishnan V, Lok PY& Rahim HA(2021), "A semi-supervised approach in detecting sentiment and emotion based on digital payment reviews", J Supercomput. 2021;77:3795–3810. doi: 10.1007/s11227-020-03412-w.

[8] Jiahua Du, Jia Rong, Sandra Michalska, Hua Wang& Yanchun Zhang(2019), “Feature selection for helpfulness prediction of online product reviews: An empirical study”, PLoS One. 2019; 14(12): e0226902. doi: 10.1371/journal.pone.0226902
[9] Muhammad Bilal& Abdulwahab Ali Almazroi(2022), “Effectiveness of Fine-tuned BERT Model in Classification of Helpful and Unhelpful Online Customer Reviews“,Electron Commer Res. 2022 Apr 29: 1–21, doi: 10.1007/s10660-022-09560-w
[10] Upendra Singh, Anant Saraswat, Hiteshwar Kumar Azad,Kumar Abhishek&S Shitharth(2022), “Towards improving e-commerce customer review analysis for sentiment detection“, Sci Rep. 2022; 12: 21983, doi: 10.1038/s41598-022-26432-3
[11] D. Dharrao, S. Deokate, A. M. Bongale &S. Urolagin(2023), "E-Commerce Product Review Classification based on Supervised Machine Learning Techniques," *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2023, pp. 1934-1939, doi: 10.1109/ICACCS57279.2023.10112717.
[12] Yanying Mao , Qun Liu & Yu Zhang(2024), “Sentiment analysis methods, applications, and challenges: A systematic literature review”, Journal of King Saud University - Computer and Information Sciences, Vol. 36, Issue 4, April 2024, 102048
[13] HamedTaherdoost&MitraMadanchian(2023), “Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research”, Computers , Vol. 12 , No.2 , 10.3390/computers12020037