International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
***Article Received**: 02 March 2023 **Revised**: 26 April 2023 **Accepted**: 12 May 2023*

E-commerce Product Price Monitoring and Comparison using Sentiment Analysis

Nitin Sakhare[1,*], Devika Verma[1], Vikas Kolekar[1], Avinash Shelke[1], Akhilesh Dixit[1], Nikhil Meshram[1]
[1,*]nitin.sakhare@viit.ac.in
[1]Department of Computer Engineering,
[1]Vishwakarma Institute of Information Technology, Pune

Abstract*:* The increasing prevalence of e-commerce has empowered consumers with vast choices and opportunities for online shopping. This research paper focuses on two essential aspects of online shopping: price comparison and sentiment analysis of product reviews. The paper presents a methodology for scraping product prices from multiple e-commerce websites and conducting sentiment analysis on the corresponding product reviews. The findings of this research have significant implications for both consumers and e-commerce businesses. Consumers can leverage price comparison data to identify the most cost-effective platforms for their desired products, while sentiment analysis enables them to assess the overall satisfaction levels of other customers. E-commerce businesses can utilize these insights to optimize pricing strategies, identify areas for improvement, and enhance customer experiences. Performance analysis of Support Vector Machine, Logistic Regression, VADER Lexicon and SentiWordNet Lexicon is also done.

Keywords: Price comparison, Web scraping, Support Vector Machine, Logistic Regression, VADER Lexicon and SentiWordNet Lexicon Sentiment analysis, Natural language processing (NLP).

I. Introduction

With the rise of online business, e-commerce has become a massive market for consumers to purchase goods conveniently from anywhere, using various smart devices. Consequently, online buyers are increasingly involved in the e-commerce industry, with numerous e-commerce websites offering a vast array of products to choose from. However, with the abundance of e-commerce websites, it can be overwhelming for consumers to search and compare prices for a single product across multiple platforms [1]. Our proposed solution aims to address this issue by providing a system for users to search for the best deals on their desired products from multiple e-commerce websites. By utilizing web scraping techniques, our system extracts data from e-commerce web pages to provide users with comprehensive product information, enabling them to make informed purchase decisions quickly. This saves users time, effort and helps them save money by finding the best deals available [2]. Consumer-generated content, such as product reviews, holds significant influence over buyer's purchasing decisions since they are often motivated by the experiences and recommendations of other shoppers. Therefore, it is crucial to develop systematic techniques for understanding the information conveyed in user-generated content. The most used method for gaining insights into customer's text reviews is sentiment analysis, which determines whether a review is positive or negative [3]. Furthermore, the sheer size of user-generated content repositories and their fast growth make it labour-intensive to manually monitor and extract sentiment from such content. Hence, automatic classification of textual content is the only practical approach for effective data classification and insights. When comparing a particular product on different e-commerce websites, such as Amazon and Flipkart, there are several factors to consider. It is important to compare the prices of the product on each website, as they may vary depending on the seller, promotions, discounts available, and other factors. This can be done by searching for the product on different websites and comparing the prices listed. Compare the shipping and delivery options available on each website. This includes the estimated delivery time, shipping cost, and whether the website offers free shipping. Some websites may also offer same-day or next-day delivery options, which can be a deciding factor for some buyers. Consider the availability of customer reviews and ratings for the product on each website. Reviews and ratings can give valuable insights into the quality of the product, and can also help you identify any potential issues or drawbacks [4]. Overall, comparing a particular product on different e-commerce websites like Amazon and Flipkart involves considering several different factors, including price, customer reviews and ratings, and user experience. By carefully considering all these factors, an informed decision about where to buy the product that best meets the needs and preferences. Web scraping is the process of extracting data from websites automatically using a program or a script as

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
***Article Received**: 02 March 2023 **Revised**: 26 April 2023 **Accepted**: 12 May 2023*

shown in Fig. 1. E-commerce websites are a common target for web scraping because they contain a large amount of data that can be useful for various purposes, such as price comparison, market research, and content analysis [5]. To scrape product details from an E-commerce website:

- Identify the target website and the product pages to be scrapped.
- Inspect the HTML code of the product pages to identify the elements i.e., class, span, div etc, that contain the product details that need to be extracted, such as the product name, description, price, image, reviews, and ratings.
- Create a web scraping program or a script using a programming language such as Python or JavaScript. Libraries such as Beautiful Soup, Scrapy, or Puppeteer can be used to simplify the scraping process. The system uses Python and the BeautifulSoup library for scraping.
- Run the web scraping program or script and collect the product details from the target website. Issues such as anti-scraping measures, page navigation, and data formatting need to be handled.
- Store the scraped data in a structured format such as CSV, JSON, or a database. Data visualization tools can be used to explore and analyze the scraped data.

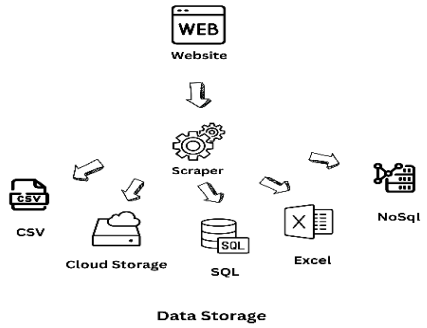

Fig. 1. Web scraping process

Web scraping can be a legally sensitive area, especially when scraping data from websites that have terms of use or copyright restrictions. Before scraping any website, the website's terms of use and its rules and policies must be considered. Additionally, some websites may block or restrict scraping activities, so it is important to use web scraping responsibly and ethically.

Examples:

Amazon is a multinational company world's largest online marketplace that offers a wide range of products, including electronics, books, clothing, home goods, and much more. The orientation of the product details does not affect the class name which helps in standardizing the same scraping attributes for all product genres.

Flipkart is one of the leading e-commerce companies, that offers a range of products. The orientation of product details changes the class value respectively. So, a change in the orientation of the product leads to a change in class value accordingly as shown in (Fig. 2, Fig. 3, and Fig. 4.).

- 'class': '_1fQZEK'

Fig. 2. Vertical list orientation of products

- 'class': 's1Q9rs'

Fig. 3. Horizontal list orientation of products (square)

- 'class': '_2UzuFa'

Fig. 4. Horizontal list orientation of products (rectangle)

Amazon product reviews are a valuable resource for both customers and sellers on Amazon's e-commerce platform. Product reviews are written by customers who have purchased and used a particular product, and they provide valuable information and feedback about the product's performance, quality, and other features. The reviews are typically displayed in reverse chronological

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
***Article Received**: 02 March 2023 **Revised**: 26 April 2023 **Accepted**: 12 May 2023*

order, with the most recent reviews appearing first. Each review includes a star rating, written text, and sometimes photos or videos uploaded by the customer [6]. Customers can use Amazon product reviews to make informed purchasing decisions by reading about the experiences of other customers who have used the product. Product reviews can also help sellers to improve their products by identifying areas for improvement and responding to customer feedback. Amazon product reviews are a valuable source of information for both customers and sellers, and they play an important role in the functioning of Amazon's e-commerce platform. A Rate-based rule tracks the rate of requests for each originating IP address and triggers the rule action on IPs with rates that go over the limit that you set. The rule is used by Amazon to put a temporary block on requests from an IP address that is sending excessive requests [7]. It does not allow multiple requests for web scraping because of which only one page with 10 reviews can be scraped for the sentiment analysis. Free proxy servers can be used for scraping the reviews but the free proxies are well-known and already blocked by the website. Flipkart is a popular e-commerce platform in India where customers can purchase a wide range of products online. One of the features of Flipkart is the ability for customers to leave product reviews. These reviews can provide valuable insights into the quality and usability of products for potential buyers. Flipkart product reviews are typically written by customers who have already purchased and used the product. They can include ratings on a scale of 1 to 5 stars and a written description of the user's experience with the product. Reviews can cover a wide range of information, including product quality, delivery experience, customer service, and more [8]. Relatively, it was easy to scrape the Flipkart website since it allows multiple requests from the same IP address. So, all customer reviews can be scraped but restricted to recent 1000 reviews because the time required to scrape the data is very high and the earliest review does not define the current status of the product.

II. Methodology

2.1 Sentiment Analysis

Sentiment analysis is a natural language processing (NLP) technique used to determine the emotional tone or polarity of a piece of text. The goal of sentiment analysis is to classify the text as positive, negative, or neutral based on the language used in the text. The main goal of sentiment analysis is to understand and extract subjective information from text data. It is widely used in various applications: customer feedback and reviews, social media monitoring, market research, brand monitoring, political analysis, customer service and support. Sentiment analysis typically involves several phases as shown in Fig. 5 and steps, including text extraction, text pre-processing, feature extraction, and classification.

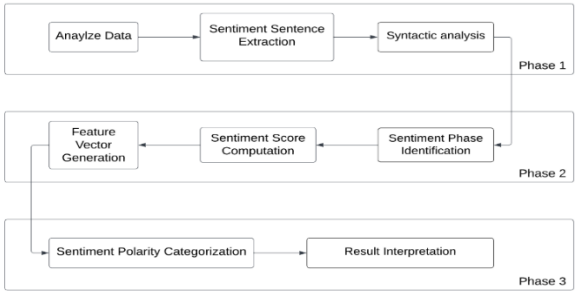

Fig. 5. Phases of Sentiment Analysis

1. Text Extraction

Text extraction includes the collection of reviews from e-commerce websites using real-time web scraping using an automated script or program to extract reviews from the product pages of e-commerce websites. The process involves the following steps:

- Identify the product pages on Amazon and Flipkart that contain the reviews for the products of interest.
- Use web scraping tools such as Beautiful Soup, Scrapy, or Selenium to extract the HTML content of the review pages.
- Parse the HTML content to extract the relevant information such as review text, review rating, reviewer name, and date of the review.
- Store the extracted information in a structured format such as a database or CSV file.

The scraping of all reviews can be done but we restricted it to 1000 reviews for the products having more than 1000 reviews because of an increase in computation time. The review title and review body section of reviews are scraped and stored in the database as shown in Table 1.

Table 1: Category of Reviews

Category	Data Type	Description
Review Title	String	Brief description of the review
Review Body	String	Detail description of the review

2. Text Pre-processing

Text pre-processing using NLP methods refers to the process of transforming raw text data into a format that can be easily analysed by machine learning models or other natural language processing tools. The process typically involves the following steps:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
Article Received: 02 March 2023 Revised: 26 April 2023 Accepted: 12 May 2023

- Tokenization: This involves breaking down the text into individual words or tokens. This can be done using tools such as NLTK or SpaCy.
- Stop word removal: Stop words are common words that do not carry much meaning, such as "the", "and", and "a". Removing stop words can help reduce the size of the dataset and improve the quality of the analysis.
- Stemming and Lemmatization: Stemming involves reducing words to their base or root form, while lemmatization involves converting words to their dictionary or base form. Both techniques can help normalize the text data and reduce the dimensionality of the dataset.
- Part-of-speech (POS) tagging: Labelling each word with its corresponding part-of-speech, such as noun, verb, adjective, or adverb. This can be done using tools such as NLTK or SpaCy.
- Named entity recognition (NER): Identifying and classifying named entities such as people, places, and organizations in the text. This can be done using tools such as NLTK or SpaCy.
- Spell checking and correction: Identifying and correcting spelling errors in the text. This can be done using tools such as PySpellChecker.
- Removal of Special Characters and Punctuation: Special characters, symbols, and punctuation marks are often removed from the text as they may not contribute to sentiment analysis or other text analysis tasks. Regular expressions or specific libraries can be employed to eliminate or replace such characters.
- Handling Abbreviations and Acronyms: Text may contain abbreviations or acronyms that need to be expanded for better understanding and analysis. Expansion can be done manually using lookup tables or using libraries that provide abbreviation expansion functionalities.
- Removal of Irrelevant or Noisy Elements: Depending on the specific analysis task, certain elements like URLs, numbers, special characters, or HTML tags may be irrelevant and can be removed from the text to reduce noise.
- Text normalization: Transforming the text into a standardized format, such as converting all text to lowercase or removing punctuation.

3. Feature Extraction

Feature extraction involves selecting and representing relevant features from text data that can be used to determine the sentiment or emotional tone [9]. Some feature extraction approaches are:

- Bag-of-words (BoW): The bag-of-words approach represents each document as a collection of its constituent words, disregarding grammar and word order. The presence or absence of words in a document is used as a feature.
- Term Frequency-Inverse Document Frequency (TF-IDF): TF-IDF represents the importance of each word in a document by considering both its frequency in the document (term frequency) and its rarity across all documents (inverse document frequency). Words that occur frequently in a document but rarely in the entire dataset tend to have higher TF-IDF values and can be indicative of sentiment.
- N-grams: N-grams are contiguous sequences of n words in a text. By considering not just individual words but also sequences of words, N-grams capture contextual information.
- Word embeddings: Word embeddings are dense vector representations of words that capture semantic relationships between words. Pre-trained word embedding models such as Word2Vec, GloVe, or fastText can be used to transform words into fixed-length numerical vectors. These embeddings can capture the meaning and context of words, allowing sentiment analysis models to leverage semantic information.
- Part-of-speech (POS) tagging: POS tagging involves labeling words in a text with their respective grammatical categories, such as nouns, verbs, adjectives, or adverbs. The presence or frequency of specific POS tags can be used as features. For example, adjectives often carry sentiment information (e.g., "good," "bad"), so their presence can be indicative of sentiment.
- Sentiment lexicons: Sentiment lexicons are curated lists of words or phrases with pre-assigned sentiment scores. Lexicons such as AFINN, SentiWordNet, or VADER contain words associated with positive or negative sentiment. By matching words in the text against the lexicon, sentiment scores can be assigned to the document based on the presence or intensity of sentiment-bearing words.
- Syntax and syntactic patterns: Analyzing the syntactic structure of sentences can provide valuable information for sentiment analysis. Identifying syntactic patterns, such as negation (e.g., "not good") or intensifiers (e.g., "very good"), can help determine the sentiment orientation of the text.

After considering these approaches TF-IDF draws out to be the best feature extraction approach for the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
***Article Received**: 02 March 2023 **Revised**: 26 April 2023 **Accepted**: 12 May 2023*

machine learning-based sentiment analysis and Sentiment lexicons for the lexicon-based sentiment analysis.

Term Frequency Inverse Document Frequency (TF-IDF) vectorization is a technique used to convert text data into a numerical representation that can be used for machine learning algorithms. The basic idea behind TF-IDF vectorization is to assign weights to each word in a document based on how frequently it appears in the document and how important it is in the corpus.

The TF-IDF vectorization process involves the following steps:

- Tokenization: To break down the text into individual words or tokens.
- Term frequency (TF) calculation: Calculate the number of times each word appears in the document. This is known as the term frequency (TF) and can be calculated as shown in equation 1.

$$TF(word) = \frac{(Frequency\ of\ Word\ in\ the\ Doc)}{(Numbers\ of\ Word\ in\ the\ Doc)} \quad (1)$$

- Inverse document frequency (IDF) calculation: The IDF value of a word is calculated as the logarithm of the total number of documents in the corpus divided by the number of documents that contain the word as depicted equation 2. The IDF value reflects how important a word is in the corpus.

$$TF(word) = log(\frac{Freq.of\ Word\ in\ the\ Doc}{Numbers\ of\ Word\ in\ the\ Doc}) \quad (2)$$

- TF-IDF calculation: The final step is to multiply the TF value of each word by its IDF value to obtain the TF-IDF weight of the word.

The TF-IDF model is constructed using two modules, TfidfVectorizer and TfidfTransformer, from the Python Scikit-learn library.

Sentiment lexicons can be used as an approach for feature extraction in sentiment analysis. Here's how sentiment lexicons can be utilized:

- Lexicon Selection: Choose a sentiment lexicon that is appropriate for your sentiment analysis task. Popular sentiment lexicons include AFINN, SentiWordNet, VADER, and NRC Emotion Lexicon. These lexicons contain words or phrases along with sentiment scores or labels indicating their positive or negative polarity.
- Lexicon Integration: Integrate the sentiment lexicon into your sentiment analysis pipeline. Load the lexicon into your program or use available libraries that provide access to sentiment lexicons.
- Text Pre-processing: Pre-process the text data using common NLP techniques such as tokenization, stop word removal, and lowercasing. This step ensures that the text is in a suitable format for matching against the sentiment lexicon.
- Matching Words: For each word in the pre-processed text, check if it exists in the sentiment lexicon. If a match is found, extract the sentiment score or label associated with that word from the lexicon.
- Aggregating Sentiment: Once sentiment scores or labels are obtained for each word in the text, aggregate them to determine the overall sentiment of the text. Common methods include summing up the scores, calculating the average, or considering the majority sentiment label.
- Handling Intensifiers and Negations: Sentiment lexicons may not capture the effects of intensifiers (e.g., "very good") or negations (e.g., "not good"). To account for these, the sentiment scores or labels can be modified based on the presence of intensifiers or negation words. For example, multiplying the score by a factor for intensifiers or flipping the sentiment label for negations.
- Contextual Considerations: Sentiment lexicons typically provide sentiment scores or labels for individual words. However, the sentiment of a phrase or sentence can be influenced by the context in which words appear. Consider the context and word order to fine-tune the sentiment extraction process. Techniques like n-grams or syntactic analysis can be applied to capture contextual sentiment information.

2.2 Classification

Classification involves using a machine learning model to predict the sentiment of the text based on the extracted features. In the context of sentiment analysis, machine learning algorithms such as SVM (Support Vector Machine) and logistic regression can be used to classify text data into different sentiment categories such as positive, negative, or neutral [10]. These algorithms learn to identify patterns and relationships in the data, and then use this knowledge to make predictions about the sentiment of new text data. SVM is a popular machine learning algorithm used for classification tasks, and it works by finding a hyperplane in a high-dimensional space that best separates different classes of data [11]. Logistic regression is another classification algorithm that estimates the probability of a binary outcome based on one or more input variables. Python Scikit-learn library consists of supervised machine learning methods. The SGDClassifier for the SVM model and LogisticRegression for the LR model are used to classify the sentiment. The model is fitted using TF-IDF features of all words from reviews. In the TF-IDF training

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
***Article Received**: 02 March 2023 **Revised**: 26 April 2023 **Accepted**: 12 May 2023*

model learns the frequency of words and vocabulary [11]. From Amazon and Flipkart a collection of 10000 product reviews from 10 different products are extracted. The collection undergoes text pre-processing and TF-IDF feature extraction. Cross-validation of five-fold is used while training. On every iteration after shuffling, a fifth of the data is allocated for testing. Both models are evaluated after training on the basis of accuracy, precision, recall, and F1 score. Lexicon-based approaches use pre-defined dictionaries or lexicons to identify and score the sentiment of words or phrases in a given text. VADER (Valence Aware Dictionary and sentiment Reasoner) and SentiWordNet are two popular lexicons used in sentiment analysis. VADER is a rule-based sentiment analysis tool that uses a dictionary of words and their valence scores to determine the sentiment of the text. SentiWordNet is a lexical resource that assigns sentiment scores to words based on their synset (a set of synonyms that share a common meaning) and part-of-speech tags [11]. VADER and SentiWordNet are fitted with pre-processed movie reviews, perform sentence extraction, unescaping HTML escape sequences, and expand contractions and then tokenize the tokens [11]. VADER has a scale of -4(most negative) to +4(most positive), the review sentiment is the summation of each sentiment score of the words in the review.
SentiWordNet additionally consists of a POS tag for tokens and rates the feature among positive, negative, and neutral scores. Document score is the summation of scores of individual words.

III. Results

Binary classification is a type of supervised learning task where the goal is to predict whether a given data point belongs to one of two classes. In this type of classification, there are four possible outcomes:

- True Positive (TP): The model correctly predicts a positive sample.
- False Positive (FP): The model predicts a positive sample when the true label is negative.
- True Negative (TN): The model correctly predicts a negative sample.
- False Negative (FN): The model predicts a negative sample when the true label is positive.

These four outcomes can together be represented in the form of a confusion matrix as shown in figure 6 to get a better understanding of the performance of the classifier.

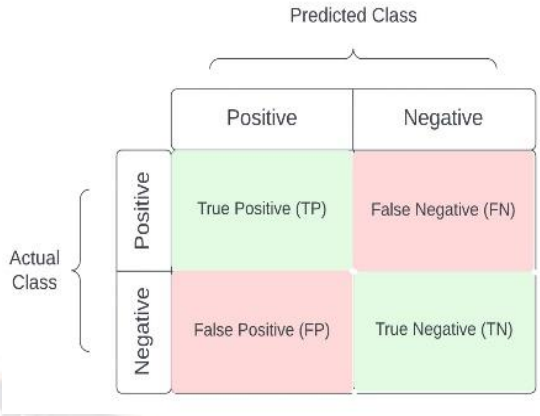

Fig. 6. Confusion Matrix

There are several metrics used to evaluate the performance of a binary classification model, including:

- Accuracy: As shown below in (3), Accuracy is the proportion of correctly classified samples among all the samples.

$$accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (3)$$

- Precision: Precision is the proportion of correctly predicted positive samples among all predicted positive samples as given below in (4).

$$Precision = \frac{TP}{(TP + FP)} \quad (4)$$

- Recall: Recall is the proportion of correctly predicted positive samples among all true positive samples shown below in equation 5.

$$recall = \frac{TP}{(TP + FN)} \quad (5)$$

- F1-score: F1-score is a harmonic mean of precision and recall. The formula for F1-score is given in equation 6.

$$F1 = 2 * \frac{(precision * recall)}{(precision + recall)} \quad (6)$$

- Area Under the Receiver Operating Characteristic Curve (AUROC): a measure of how well the model distinguishes between positive and negative samples, regardless of the chosen classification threshold.

Sentiment Analysis of the reviews performed by using the two supervised learning algorithms logistic regression (LR) and Support Vector Machine (SVM) and two majorly used lexicons in NLP, Vader and SentiWordNet. Experimental results from four different techniques to classify sentiment of Amazon and Flipkart product reviews dataset. Experimental results of all evaluation parameters for all four classification algorithms are as follows:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
Article Received: 02 March 2023 Revised: 26 April 2023 Accepted: 12 May 2023

Table 2: Performance Analysis of the models

Model	Accuracy (%)	Precision	Recall	F1 Score
Support Vector Machine	89.10	0.81	0.86	0.83
Logistic Regression	84.25	0.89	0.81	0.84
VADER lexicon	88.50	0.84	0.85	0.84
SentiWordNet lexicon	80.54	0.86	0.87	0.86

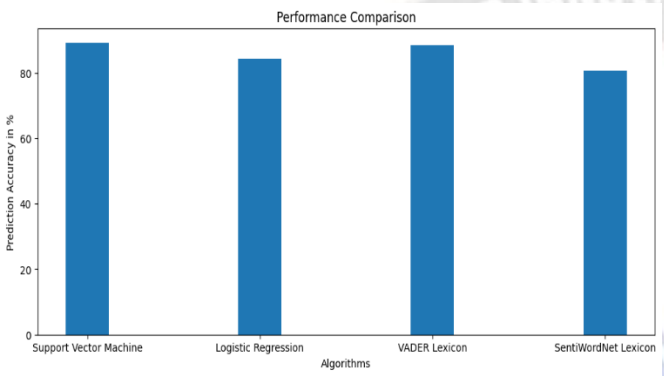

Fig. 7 Performance Comparison of the models

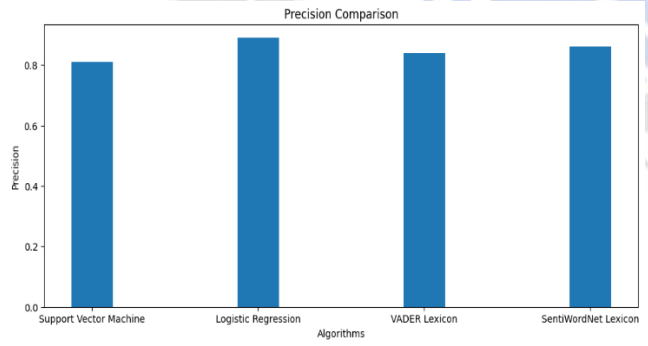

Fig. 8 Precision Comparison of the models

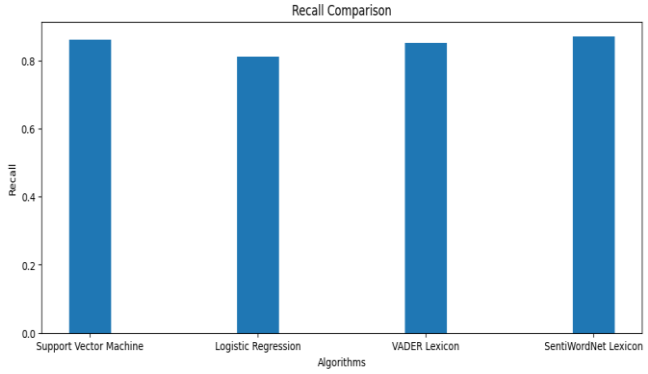

Fig. 9 Recall Comparison of the models

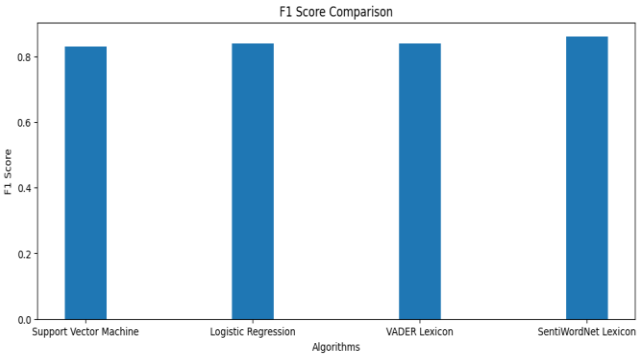

Fig. 10 F1 Score Comparison of the models

IV. Conclusion

The proposed system aims to empower users by providing them with the ability to extract product information from various e-commerce sites and conduct sentiment analysis on the latest reviews. The proposed system employs web scraping techniques to gather relevant product details. Users will have the convenience of accessing product names, prices, ratings, review counts, and images across different e-commerce platforms. Additionally, by utilizing VADER for sentiment analysis, the system can uncover the sentiment polarity and subjectivity of the most recent reviews. This versatile system holds significant potential in the realms of e-commerce and sentiment analysis, offering businesses a valuable tool to analyse customer feedback and enhance their products and services.

References

[1] Nougarahiya, Shrey and Shetty, Gaurav and Mandloi, Dheeraj, A Review of E – Commerce in India: The Past, Present, and the Future (March 15, 2021). Research Review International Journal of Multidisciplinary. Volume 06 Issue 03, March 2021, 12-22, Available at SSRN: https://ssrn.com/abstract=3809521

[2] Ikechukwu Onyenwe, Ebele Onyedinma, Chidinma Nwafor and Obinna Agbata "Developing Products Update-Alert System for e-Commerce Websites Users Using HTML Data and Web Scraping Technique ", IJNLC Vol-10, Issue-10, 2021 https://doi.org/10.48550/arXiv.2109.00656

[3] Fang, X., Zhan, J. Sentiment analysis using product review data. Journal of Big Data 2, 5 (2015). https://doi.org/10.1186/s40537-015-0015-2

[4] Daroch, B., Nagrath, G. and Gupta, A. (2021), "A study on factors limiting online shopping behaviour of consumers", Rajagiri Management Journal, Vol. 15 No. 1, pp. 39-52. https://doi.org/10.1108/RAMJ-07-2020-0038

[5] Moaiad Ahmad Khder, "Web Scraping or Web Crawling: State of Art, Techniques, Approaches and Application", Int. J. Advance Soft Compu. Appl, Vol. 13, No. 3, November 2021 Print ISSN: 2710-1274

[6] Susan M. Mudambi, "Mudambi & Schuff/Consumer Reviews on Amazon.com RESEARCH NOTEWHAT

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
DOI: https://doi.org/10.17762/ijritcc.v11i5.6693
Article Received: 02 March 2023 Revised: 26 April 2023 Accepted: 12 May 2023

MAKES A HELPFUL ONLINE REVIEW? A STUDY OF CUSTOMER REVIEWS ON AMAZON.COM", MIS Quarterly Vol. 34 No. 1, pp. 185-200/March 2010

[7] NN Sakhare, SA Joshi, "Criminal Identification System Based On Data Mining" 3rd ICRTET, ISBN, Issue 978-93, Pages 5107-220, 2015

[8] NN Sakhare, SA Joshi, "Classification of criminal data using J48-Decision Tree algorithm" IFRSA International Journal of Data Warehousing & Mining, Vol. 4, 2014

[9] NN Sakhare, SS Imambi, Technical Analysis Based Prediction of Stock Market Trading Strategies Using Deep Learning and Machine Learning Algorithms, International Journal of Intelligent Systems and Applications in Engineering, 2022, 10(3), pp. 411–42.

[10] Sakhare,N.N., Shaik,I.S.,Saha,S.: Prediction of stock market movement via technical analysis of stock data stored on blockchain using novel History Bits based machine learning algorithm. IET Soft.1–12(2023). https://doi.org/10.1049/sfw2.1209212

[11] NN Sakhare, SS Imambi, S Kagad, H Malekar, M Dalal, "Stock market prediction using sentiment analysis" International Journal of Advanced Science and Technology, Vol. 4, issue 3, 2020.