iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** III **Month of publication:** March 2023

DOI: https://doi.org/10.22214/ijraset.2023.49904

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

Survey on Forecasting of E-Commerce Product Rating

Sanan Akther[1], Vedananda D.E[2], Skanda Prasad[3], Vikas Rao GV[4], Vishwas TS[5]
[1, 3, 4, 5]*Research Scholar, Dept. of CS&E, JNNCE, Shimoga, India*
[2]*Assistant Prof., Dept. of CS&E, JNNCE, Shimoga, India*

Abstract: ***E-commerce websites are presently the most popular online shopping platforms for a broad spectrum of products. Users are now provided with numerous different sites to access and search for the needed items on the market. Several tactics are now being applied to evaluate and understand client behaviour in order to enhance corporate growth. One such method is the online product rating, which is a critical statistic for understanding a product's acceptance by users.***
Keywords: E-commerce, online product rating, corporate growth.

I. INTRODUCTION

Rating prediction refers to predicting a star or numbered rating from the analysis of text reviews given by online shoppers. The rating is used by shoppers to assess the quality and perfection of an online product. It assists an internet buyer in deciding whether to purchase a product. It also enables the manufacturer to further modify the product throughout replication. Occasionally consumers will buy a thing online and leave a written review, but they are averse to leaving a number rating, most frequently a star rating. Yet, manufacturers must be aware of product ratings in order to do business analysis. Producers can utilize this grade for business research and to boost income. So the rating and review are significant to both the shoppers and producers.

II. RELATED WORK

Aditya Ambre at el. [1] describes in detail the process of comparing items from several e-commerce websites. To collect data from multiple websites, web crawling and web scraping techniques are used. A crawler may be constructed to get product URLs, which are then sent into the scraper to extract the necessary data from the HTML content. The gathered data is stored in a document-oriented fashion using MongoDB, a NoSQL database. JSON is a standard for storing vast volumes of unstructured data gathered during the scraping process. MongoDB is used to categorize and store product data from various e-commerce websites. The comparison logic compares items based on many factors such as name, price, features, and so on. Users may do a product search, and the query is sent to the local MongoDB database. The comparison is then conducted on many databases at the same time, and the goods are parsed according to their properties and categories, with the resulting comparison shown. Indexing is used in MongoDB to make the search process more effective and quicker.

Prashant Sanap at el.[2] to create a web application that allows customers to get the cheapest pricing for a product by uploading an image of it, rather than typing out the actual name of the product. To obtain data from various locations, several modules are needed. After all of the data from completely separate modules is received, the data with the same product name is combined and exhibited in each of the product photographs in this area with the product label and matching product names with price and product buy link. To seek for the most effective pricing from all various e-commerce sites of a product unit of measurement compared and shown best value with product link to the user at the bottom. The e-commerce website issues an API key and a secret key to allow associated users to use the website's relevant API to obtain data on various products.

S. Rajendar at el.[3] suggested system allows online consumers to locate the lowest possible price for their goods from many e-commerce websites using a single web interface. The front-end platform provides a graphical user interface in the form of a website via which clients interact with the system, whilst the backend uses web crawling and scrapping techniques to gather product data from different e-commerce websites. Another feature available on the website is a price alert, which users can set to be alerted by the website anytime an appropriate price becomes available. The data is published on the website after it has been scraped from various e-commerce websites. A single online interface is used to compare product prices from several e-commerce websites, and the results are shown.

©IJRASET: All Rights are Reserved

Md. Iqbal Hossain at el.[4] suggested an aspect-based rating prediction algorithm is proposed, which predicts ratings based on the review text. Suggested an algorithm that can remove innate views from customer reviews and apply it to the algorithm from numerous perspectives. Machine Learning Algorithms This experiment made use of three machine learning algorithms. The XGBoost algorithm achieved an accuracy of 84.7%, while the Logistic Regression algorithm achieved an accuracy of 92.1%. The Random Forest classifier outperformed the other two algorithms in this dataset. The Random Forest algorithm surpassed the others by achieving an accuracy of 94.2%.

Table-1: Comparison of Models

Algorithm Name	Precision	Recall	f1-score	Accuracy
Random Forest	0.941	0.944	0.949	94.2
XGBoost	0.858	0.856	0.851	84.7
Logistic Regression	0.929	0.921	0.922	92.1

Caroline El Fiorenza at el. [5] the intention of this research is to monitor and remove the fake reviews posted about products. In our paper, we consider the features of a product. Then divide the respective features into a series of sub-features. Many algorithms are developed for this task like Apriori, FP growth etc. Among these Apriori and FP growth have been studied on a large scale. This is based on the dataset where the ratings are on a scale of 1 to 5.

Then a dictionary is made out of the current tokens and words with the help of the SentiWordNet tool which is used for creating a dictionary of sentiment tokens. Part-Of-Speech (POS) tagger is software that reads text and determines the part of speech for each token. Create all tokens that will be stored in the database from products that will be used as input for the FP growth process at first. Polarity is generated by identifying the features with the most comments. Finally, the agreement values will be calculated. Because no algorithm can solve all of today's challenges and difficulties, certain measures such as helpfulness, usefulness, and utility must be considered when analysing each review.

Terutaka Yoshikawa at el.[6] a product recommendation system was proposed to resolve user complaints about products by analysing complaint data as well as satisfaction data from product reviews. Many e-commerce recommendation systems are based on review comments left by previous purchasers of the product. These systems assist new users in locating products and promoting their purchases. Even when users have accurate information or data, it can be difficult to find products that meet their needs. The first component of this system is product review analysis, which involves collecting product reviews and characterizing them as low-rated reviews or high-rated reviews. The second component is the extraction of negative and positive information, which involves resolving negative feature words with low ratings and positive feature words with high ratings. Eventually, product recommendations are made one after the other by analyzing the similarities between negative and positive data. This last section outlines the product review analysis approach and how to propose complaint resolution products. We tested the efficacy of our suggested complaint information extraction technique and our proposed complaint solution product suggestion approach

Rahul at el.[7] presents an overview of the technology employed in sentiment analysis classification, as well as an evaluation of the systems employed. There are other more strategies that work well on different datasets. Specifically addresses the various strategies utilized by various writers to do sentiment classification on diverse review sets. Proposed certain approaches, such as Naive Bayes and Support Vector Machine (SVM), for identifying the respective polarity on internet reviews for smartphone devices as a dataset. The recommended ways for sentiment similarity calculation include the fine-grained analysis method, the weighted average method, and the propagation algorithm for propagation trust computation. These approaches depend on graphs and are applied to Amazon.com reviews from various categories. As a result, the approach enhances analytical precision and aids in the discovery of polarity or classification of those evaluations into numerous feelings. On their particular datasets, the approaches or algorithms utilized here produced good and accurate results. Lastly, SVM produces the greatest results when computing polarity on cryptic sentiments or reviews.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Table-2: Summary of survey

TITLE	METHODOLOGY	ADVANTAGES	LIMITATIONS
Web and Android Application for Comparison of E-Commerce Products [1]	Web Crawling Web Scraping MongoDB Comparison Logic	Product data can be compared and fetched through different websites. It provides a simple interface for users to interact.	Handling a large amount of data can be more complex. Sometimes communication between database and server has long awaited time.
PRICE COMPARISON WEBSITE USING OBJECT RECOGNITION [2]	Web Scraping Search by image E-Commerce site API	Different APIs are used for extracting data rapidly Automation is used for storing data in the database	Customers can face many problems due to traffic in the network. Due to inaccurate data fetch from the database display various incorrect information to the user.
PRICE COMPARISON WEBSITE FOR ONLINE SHOPPING [3]	Web Crawling Web Scraping Python Django Framework	It provides a simple architecture for developers. Provides an interface for users to compare different prices of products from various websites	Description of products may be inaccurate while fetching Due to websites server downtime fetching of data can take more time.
Forecast the Rating of Online Products from Customer Text Reviews based on Machine Learning Algorithms[4]	Pre-processing Feature Extraction Machine Learning Algorithms	TF-IDF vectorizer is a very efficient and perfect method to convert a plain text to a vector. XG boost algorithm is more accurate compared to other algorithms.	There may be many complications while converting the same text which consists of different meanings. It may take more time for pre-processing and extraction of data.
Fake Product Review Monitoring and Removal for Genuine Online Reviews [5]	Review monitoring. Opinion spam analysis. Sentimental analysis.	Ability to remove fake and abusive reviews using various algorithms	During review filtering some of genuine comments will be removed, this can be very difficult for users.
A Product Recommendation System Based on User Complaint Analysis Using Product Reviews [6]	Extracting Complaint Information. Extracting Products with Positive Evaluation. Recommending Complaint Solution Products.	TF-IDF is very efficient in terms of feature extraction of the product-based reviews. Using this methodology positive rated products are recommended.	Because of inefficient preprocessing the data the accuracy of the prediction may reduce.
Sentiment Analysis on Product Reviews [7]	Data Acquisition. Text Pre-processing. Feature Selection and Extraction. Sentiment Classification. Polarity Detection. Validation and Evaluation.	Using the models like Support Vector Machine (SVM), Naïve Bayes and Neural Network, provides best result based on sentiment classification.	During sentiment analysis, classification based on sentiments may cause polarity shift problem.

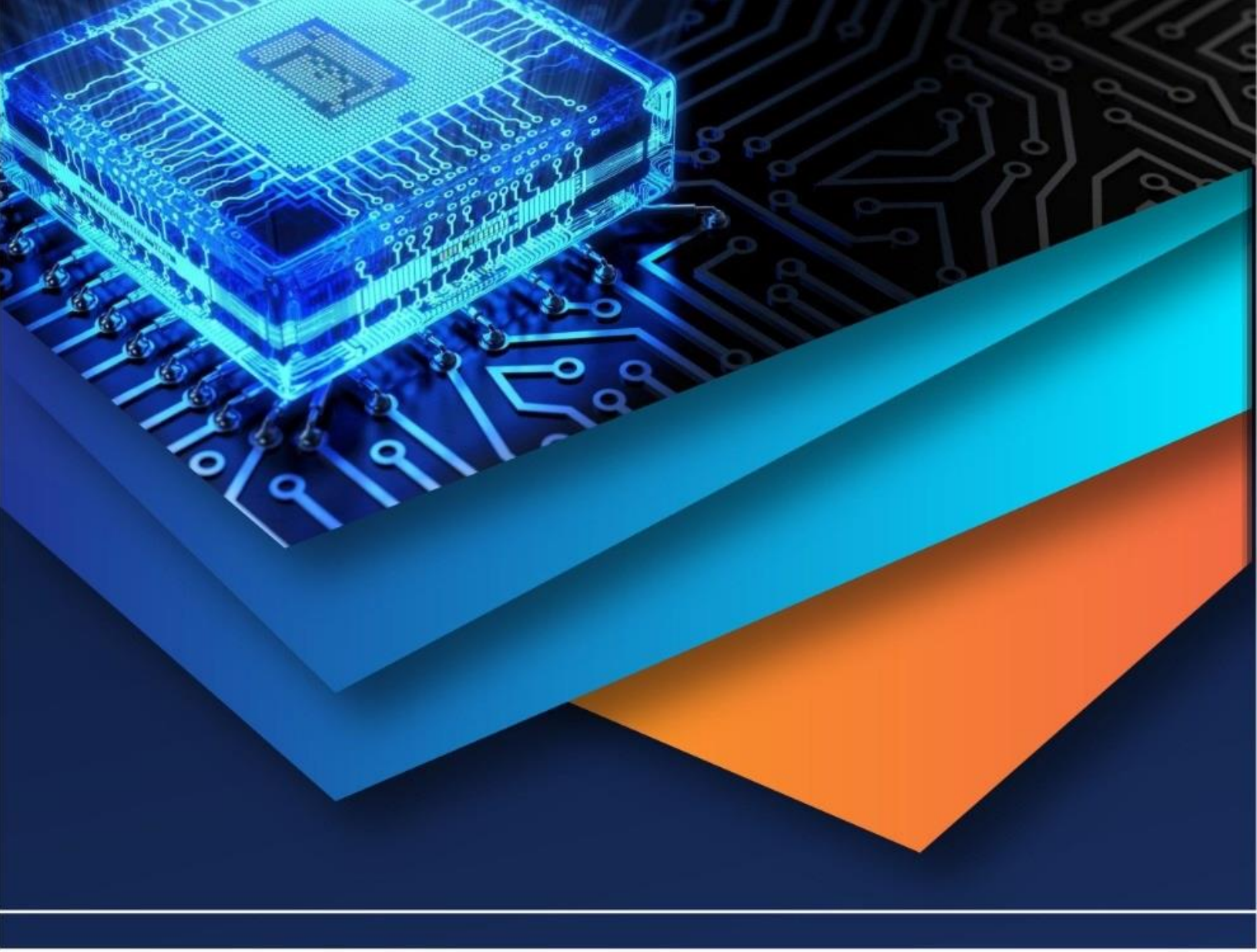
III.CONCLUSION

Predicting the rating of online products according to customer text reviews is a significant and difficult problem that machine learning systems can handle. With the rise of e-commerce platforms, online product reviews have become an important element of the customer decision-making process. Yet, it is impossible to read through all of the written reviews before purchasing. Using machine learning algorithms to create accurate rating prediction models may substantially benefit both businesses and customers. These models may assist businesses in improving their product offerings, customer service, and marketing tactics, while also enabling customers to make more educated purchase decisions. Accurate rating prediction models like Random Forest will become increasingly vital in guaranteeing client pleasure and loyalty as e-commerce grows.

REFERENCES

[1] Aditya Ambre , Praful Gaikwad , Kaustubh Pawar , Vijaykumar Patil, “Web and Android Application for Comparison of E-Commerce Products”, IJAEMS, Vol-5, April-2019
[2] Prashant Sanap, Swati Shinde, Anjali Mahajan, Rahul Vishe, Anuprita Gawande, “Price Comparison Website Using Object Recognition”, IRJMETS, Vol-4, April-2022
[3] S. Rajendar, K. Manikanta, M. Mahendar, Assistant Prof. (Mrs.) K. Madhavi, “Price Comparision Website for Online Shopping”, IJCRT, Vol-9, June-2021
[4] Md. Iqbal Hossain , Maqsudur Rahman, Tofael Ahmed and A. Z. M. Touhidul Islam, “Forecast the Rating of Online Products from Customer Text Review based on Machine Learning Algorithms”, ICICT4SD, February-2021
[5] Caroline El Fiorenza , Aditya Singh Kashyap , Karthikey Chauhan, Kishan Mokaria, Aashutosh Chandra,“Fake Product Review Monitoring and Removal for Genuine Online Reviews”, JNCET, Vol-8, April-2018
[6] Terutaka Yoshikawa, Yuanyuan Wang, Yukiko Kawai “A Product Recommendation System Based on User Complaint Analysis Using Product Reviews”, GCCE, 2019
[7] Rahul, Vasundhara Raj, Monika, “Sentiment Analysis on Product Reviews”, ICCCIS, 2019.
[8] Riya Shah, Karishma Pathan, Anand Masurkar, Shweta Rewatkar, P.N. Vengurlekar, “Comparison of E-Commerce Products using Web Mining”, International Journal of Scientific and Research Publications, Volume 6, Issue 5, May 2016 640 ISSN 2250-3153.
[9] A. Rafay, M. Suleman, and A. Alim, “Robust review rating prediction algorithm based on machine and deep learning: Yelp dataset,” in Intern. Conf. on Emerging Trends in Smart Technologies, March 26-27, 2020, pp. 8138-8143.
[10] “Grammarand Product Reviews,” [Online]. Accessed on: June 23, 2020. Available: https://www.kaggle.com/datafiniti/grammar-andonline-productreviews
[11] B. Zhang, H. Zhang, X. Sun, G. Feng, and C. He, “Integrating an attention mechanism and convolution collaborative filtering for document context-aware rating prediction,” IEEE Access, vol. 7, December 17, 2018, pp. 3826-3835.
[12] E. Tutubalina, V. Malykh, S. Nikolenko, A. Alekseev, and I. Shenbin, “AspeRa: aspect-based rating prediction algorithm,” in Proceedings of the Workshop on Eidening NLP, 2019, pp. 11-13.
[13] R. Haji, K. Daanyaal, G. Deval, and G. Rushikesh, “Rating prediction based on textual review: machine learning approach, lexicon approach and the combined approach,” in International Research Journal of Engineering and Technology (IRJET), vol. 6 (3), March 2019, pp. 5437-5443.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)