International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 02 Issue: 08 August 2024
Page No: 2228- 2234
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2024.0304

A Survey on Diamond Security System Using Machine Learning and Blockchain

Swaraj Ravindra Dudhe[1], *Prajakta Aniruddha Ichole*[2], *Sudarshan Rameshwar Bansode*[3], *Kalyani Sandip Gaikwad*[4], *Kavita Mahesh Jadhav*[5]

[1,2,3,4,5]*Nutan College of engineering and Research, Talegaon Dabhade, Ta Maval Dist Pune, India.*

Emails: *swarajdudhe0501@gmail.com*[1], *prajaktaichole06@gmail.com*[2], *bansodesudarshan02@gmail.com*[3], *Kalyanigaikwad865@gmail.com*[4], *kavitajadhav@ncerpune.in*[5]

Abstract

The pervasive challenges of counterfeit in the diamond industry, cybersecurity vulnerabilities, and the looming threat of large-scale heists, akin to the infamous Antwerp money heist of 2003, underscore the critical need for innovative security solutions. This initiative addresses these pressing concerns through the fusion of machine learning algorithms and blockchain technology. Our methodology entails a comprehensive dataset comprising 32 parameters for advanced detection and 5 dominant parameters for basic detection, catering to both expert and novice users. Leveraging machine learning techniques including random forest and ensemble learning, our system meticulously analyzes these parameters to discern between natural and artificial diamonds with heightened accuracy. The accuracy of our models varies, with the highest achieving an impressive 98%, ensuring robust authentication and fraud prevention mechanisms. Furthermore, the verified data of natural diamonds is securely stored on a blockchain ledger, fortifying ownership records against tampering and ensuring transparency. This integrated approach not only safeguards the integrity of the diamond industry but also fosters trust and transparency among stakeholders. By providing a paradigm shift in security solutions, our system sets a new standard for authentication and fraud prevention in the diamond market, mitigating risks and empowering stakeholders with reliable, data-driven tools for decision-making.

Keywords: *Machine Learning (ML), Blockchain (BC), SHA256, recommendation system, Diamond industry.*

1. Introduction

The Diamond Security System is a comprehensive solution designed to authenticate and safeguard diamonds, ensuring their integrity and security within the diamond industry. It combines advanced technologies and methodologies to address various challenges such as counterfeit diamonds, cybersecurity vulnerabilities, and the risk of large-scale thefts.

a. Authentication Technology: Utilizes advanced techniques such as machine learning algorithms to differentiate between natural and artificial diamonds based on various parameters like size, shape, color, and internal characteristics.

b. Blockchain Integration: Incorporates blockchain technology to securely record ownership data and transaction history of verified natural diamonds. Blockchain ensures tamper-proof and transparent records, enhancing trust and transparency in the diamond supply chain.

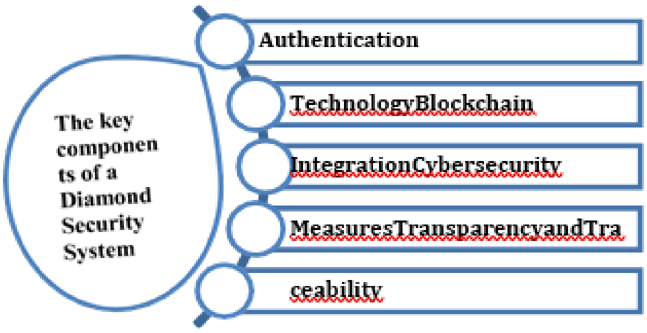
The key componen ts of a Diamond Security System
Authentication
TechnologyBlockchain
IntegrationCybersecurity
MeasuresTransparencyandTra
ceability

Figure 1 **Key Components of System**

c. Cybersecurity Measures: Implements robust cybersecurity protocols protect diamond-related data and prevent unauthorized access or cyberattacks, safeguarding sensitive information and digital assets, shown in Figure 1.

d. Fraud Detection: Employs sophisticated fraud detection mechanisms to identify and prevent counterfeit diamonds from entering the market, thereby maintaining the authenticity and value of genuine diamonds.

e. Transparency and Traceability: Promotes transparency and traceability throughout the diamond supply chain, enabling stakeholders to track the journey of diamonds from mining to retail, ensuring ethical sourcing and compliance with regulations.

In 2003, the Antwerp Diamond Center, nestled in Belgium's diamond district, became the stage for one of history's most daring heists. Over a weekend, thieves orchestrated a meticulous plan, circumventing security measures to make off with an estimated $100 million in diamonds, gold, and precious gems. Despite exhaustive investigations, the perpetrators eluded capture, leaving the industry reeling and highlighting the urgent need for enhanced security. This paper addresses the persistent challenges posed by such audacious heists and broader concerns within the diamond industry. At its core lies the problem of counterfeit pharmaceuticals infiltrating the diamond market, exacerbating existing security vulnerabilities. To combat these threats, we've devised a comprehensive security system integrating cutting-edge technologies and innovative methodologies. Our approach centers on machine learning algorithms, particularly random forest and ensemble learning, augmented by blockchain technology. By meticulously analyzing a dataset comprising 32 parameters for advanced detection and 5 dominant parameters for basic detection, our system distinguishes between natural and artificial diamonds with remarkable accuracy. The culmination of our methodology yields promising results, with our models achieving an impressive accuracy rate of up to 98%. Furthermore, our implementation of blockchain technology ensures the secure storage of verified data for natural diamonds, fortifying ownership records and enhancing transparency throughout the supply chain. our project represents a significant leap forward in diamond security, mitigating the risks posed by counterfeit pharmaceuticals, cybersecurity vulnerabilities, and large-scale heists. By instilling trust and confidence among stakeholders, we pave the way for a more secure and transparent diamond market, safeguarding the industry's integrity for generations to come.

2. Literature Survey

Urvish Thakker et al. [1] in their scholarly exploration, the author elucidates the application of blockchain technology within the intricate context of the diamond industry. Within this industry, marked by conundrums related to provenance, supply chain transparency, third party validation, and transaction fidelity, blockchain emerges as a vanguard solution. In this multi-faceted analysis, the authors undertake comprehensive research on blockchain's integration into the diamond industry, presenting a meticulous dissection of its inherent merits and demerits. Furthermore, grounded in the extensive literature review, the authors proffer stratagems to counter the recurrent issues within diamond industry operations. In addition to this, the study expounds upon the myriad unexplored research dimensions and multifarious challenges interwoven with the assimilation of blockchain within the diamond industry. The research culminates with a case study, predominantly illuminating the real-time challenges encountered within the industry, underpinned by robust frameworks such as Everledger and Tracr. In a scholarly odyssey, authors navigate the intricate terrain of the diamond industry, unraveling the profound implications of blockchain technology. Confronted with intricate predicaments concerning provenance, supply chain transparency, third-party authentication and transaction integrity, blockchain emerges as a transformative solution, effectively bridging the chasm that separates the diamond industry from burgeoning financial markets. A cornerstone of this analysis lies in the profound trust instilled within stakeholders, who can now rest assured of the authenticity of their crystalline acquisitions and the swift resolution of property disputes. Bhabendu Kumar Mohanta. et al. [2] in a world increasingly reliant on intricate supply chains and interconnected systems, blockchain technology and smart contracts offer a compelling solution for traceability, tamper-proof record-keeping and automated execution. Exploring the capabilities of this formidable combination to reshape trust and enhance efficiency across various sectors, the following analysis investigates the potential impact.

When integrated with blockchain, authors enable real-time, cost-effective and highly secure task execution. The work starts by elucidating the components and working principles of smart contracts. It proceeds to identify and analyze various use cases of smart contracts within blockchain applications, highlighting their advantages. Finally, the research delves into the challenges associated with implementing smart contracts in future real-life scenarios. Within this landscape, the research delves into the realm of smart contracts, which are self-executing computer programs governed by a predefined set of rules. Over the past decade, blockchain technology has found versatile applications, particularly when integrated with smart contracts. This integration offers flexibility, allowing the efficient design and implementation of real-world solutions at reduced cost and time, without the need for traditional third-party systems. Huaqun Guo et al. [3] blockchain technology offers a multitude of highly sought after attributes, including decentralization, autonomy, integrity, immutability, verification, fault tolerance, anonymity, auditability and transparency. Security within the blockchain ecosystem takes center stage, with a meticulous analysis of risk categories, real- world attacks and vulnerabilities and an overview of recent security advancements. Furthermore, the work outlines the challenges and evolving research trends aimed at creating more scalable and secure blockchain systems for large scale deployments. Within the blockchain paradigm, data resides within a distributed ledger. This technology ensures the integrity and availability of transactions, allowing participants to read, write and verify entries while preventing alterations or deletions. The security underpinning blockchain relies on cryptographic primitives and protocols, such as digital signatures and hash functions, which safeguard the integrity, authenticity and non-repudiation of recorded transactions. Avishkar Hongekar et al. [4] counterfeiting and duplication present significant risks in the global product landscape. To combat this issue, a system using blockchain technology has been developed to enable end-users to verify product authenticity via QR codes. Blockchain's decentralized, immutable ledger ensures the integrity of product data. When a user scans the QR code, a match confirms the product's authenticity, granting access to detailed information. This innovation aims to protect company reputations and consumer well-being by countering counterfeit products. Hamed Taherdoost et al. [5] emerging research explores the potential of combining blockchain technology and machine learning to bolster security. While blockchain, the backbone of cryptocurrencies, boasts robust data security, its effectiveness hinges on specific use cases and data types. Machine learning offers a promising solution by dynamically adapting to address these vulnerabilities and enhance overall security. The study reviews articles from 2012 to 2022, emphasizing the significance of machine learning impact on blockchain security. Key challenges include privacy and integration with other technologies. The research provides perception for academics and practitioners in the field. Garima Sharma et al. [6] this study uses supervised models to predict diamond prices. The Random Forest Regression model with a 97% accuracy rate is the most effective. Diamonds, known for their rarity and optical properties, are prized for their durability and fashion appeal. Their "adamantine" luster provides the distinctive sparkle. Authors also used as abrasives due to their exceptional hardness. The study showcases supervised learning's effectiveness in predicting diamond prices and suggests exploring unsupervised models for further improvements. Alper Kanak et al. [7] introduce the Diamond Accountability Model (DAM), which leverages blockchain's transformative capabilities. DAM integrates authorities into blockchain transactions, ensuring non repudiation, security, and governance. It offers a decentralized solution for secure collaboration in multi agent applications, addressing various challenges. Authorities within the blockchain enhance accountability and trust. A use case demonstrates collaborative scenarios. The authors have summarized that DAM is a promising approach for blockchain enabled multiagent ecosystems, reducing concerns about illegal blockchain usage for financial crime. Prof. Amruta et al. [8] diamonds, more than just dazzling ornaments hold a hidden value that dances with carat weight, color and market trends. This research unveils the secrets of this dance, wielding machine learning algorithms like linear regression, decision trees and K-nearest neighbors to predict diamond prices with a precision that rivals a diamond cutter's touch. Author have summarized that machine learning algorithms are promising for predicting diamond prices, enabling informed decisions by buyers and sellers. Ahmed M. Elmogy et al. [9] delve into the critical task of detecting fake reviews in the realm of E-commerce and online reputation management. Online reviews wield immense influence on consumers' decisions. Positive reviews in particular hold substantial value. However, the rise of deceptive or fake reviews has posed a significant challenge. To address this issue, the research introduces a machine learning approach that not only analyzes review attributes but also incorporates user behaviour features. Using a real Yelp dataset of restaurant reviews, the study conducts experiments with various classifiers, including KNN, Naive Bayes, SVM, Logistic Regression and Random Forest. The findings reveal that KNN, specifically with K=7, outperforms other classifiers, achieving the highest f- score of 82.40%. Importantly, the research demonstrates that by including user behaviour features the f-score increases by 3.80%. This underscores the need to continually develop techniques to identify and combat fake reviews, preserving the trust and credibility of online platforms and services. N Moratanch et al. [10] this research presents a blockchain-based solution to combat counterfeit products in the market, a problem that poses severe financial, health and safety risks. By using QR codes and barcodes linked to a blockchain, the system enables the verification of product authenticity. If the scanned code matches, the product is confirmed as genuine. Otherwise, it's considered counterfeit. Notifications are dispatched to both the customer and the manufacturer. This application enhances consumer trust and offers a more secure and user- friendly approach to trading and purchasing. In the future, further work may concentrate on simplifying the code and utilizing APIs for more efficient data extraction ultimately boosting trust and efficacy. Ms. Reema Anne Roy et al. [11] present an Artificial Intelligence-based solution for detecting counterfeit products, particularly in the context of medical items. The existing system enables non tech savvy consumers to use a mobile application to scan products, focusing on the identification of counterfeit products by analyzing logos, including both image and textual representations. The study Summarizes by emphasizing the development of a user-friendly and portable device that employs artificial intelligence for logo image analysis and text and color recognition. This technology aims to assist individuals who may not be tech-savvy in distinguishing between genuine and counterfeit products. Jayalakshmi L. et al. [12] address the critical challenge of identifying fraudulent reviews on online platforms and mitigating the adverse effects of deceptive review practices on society is the central theme of this report. With the increasing popularity of online review systems, unscrupulous actors employ tactics such as Sybil accounts and bot farms to manipulate consumers through fabricated product or service reviews. To tackle this challenge, the research employs logistic regression as a powerful technique for identifying and classifying fake reviews. Notably, logistic regression outperforms the SVM classifier in classifying fake reviews for the testing dataset, demonstrating its capacity to generalize and predict fake reviews efficiently. In summary, the detection of fake reviews is a critical concern given its adverse societal consequences. This Study approach, leveraging logistic regression alongside data processing techniques like stemming and tokenization, in conjunction with TF-IDF vectorization, presents a promising method to enhance the accuracy of fake review identification. By doing so, it contributes to mitigating the negative repercussions of deceptive reviews on online platforms. Manasi Bansode et al. [13] authors address the issue of identifying and removing deceptive online reviews. Their method classifies reviews into various categories, focusing on hotel reviews and uses data mining techniques to enhance accuracy. As online reviews are increasingly critical for consumer decision-making, combating misleading reviews is essential. The approach improves the review analysis process and introduces features such as IP address-based detection to prevent multiple submissions from the same source, a date module for accuracy and star ratings based on positive review percentages. Machine learning algorithms aid in detecting and removing fake reviews, enhancing the platform's reliability. In summary, a comprehensive solution to address deceptive online reviews, enhancing accuracy and helping consumers make informed choices while countering fake reviews' impact. Avinash Chandra Pandey et al. [14] focus on predicting future prices of precious metals such as gold and diamonds using ensemble techniques. Given the daily price fluctuations of these commodities, the goal is to apply regression models based on past price patterns to provide accurate forecasts. Precious metals, particularly gold have widespread applications in finance and investments. Predicting their prices is essential for both buyers and investors to make informed decisions. Various machine learning algorithms, particularly supervised models are employed, including linear regression, Random Forest and ensemble techniques. The comparison of ensemble models performance and the utilization of feature selection methods like PCA and RFE are key factors in enhancing prediction accuracy. Machine learning algorithms simplify the complex task of precious metal price prediction, offering efficiency for data analysts and researchers in this domain. Lin Chen et al. [15] highlight the utilization of the Alternating Least Squares algorithm through Apache Spark on Amazon Web Services to create a product recommender system. These systems are pivotal for enhancing user experiences in e- commerce and web services. By evaluating the ALS algorithm's performance and scalability with various configurations, the study shows its effectiveness in providing accurate recommendations with efficient runtime, particularly when employed on AWS's YARN cluster manager. Overall, this approach offers a promising solution for product recommendations in e-commerce platforms. Shahab Saquib et al. [16] exploring the domain of product recommendation techniques in the e-commerce sector acknowledges the rapidly digitizing world and the surge of online shopping, creating an information overload for users. Various recommendation methods, such as Collaborative Filtering, Association Rules and Web Mining are explored. The research observes the evolving trends in recommendation technology and suggests a future research direction. It Summarizes by emphasizing the need for additional efforts to address the limitations of existing techniques in the face of the information overload on the Internet. Monika K. et al. [17] focus on the significance of product recommendation systems in e-commerce, particularly for enhancing sales. It describes how product recommendation analyzes frequently purchased products to suggest items that customers are likely to buy. The Linear Regression method is emphasized in creating an effective product recommendation system. An examination is conducted on the Intel processors utilized in laptops, unravelling insights into their diverse types. The focus extends to optimizing cost values by leveraging customer ratings and price data through the application of Linear Regression techniques. The present study showcases the implementation of Linear Regression to ascertain the optimal cost figures for laptops featuring various processor types such as (i3, i5, i7). This approach leverages customer ratings and pricing data to formulate insightful recommendations.

2.1 Figures

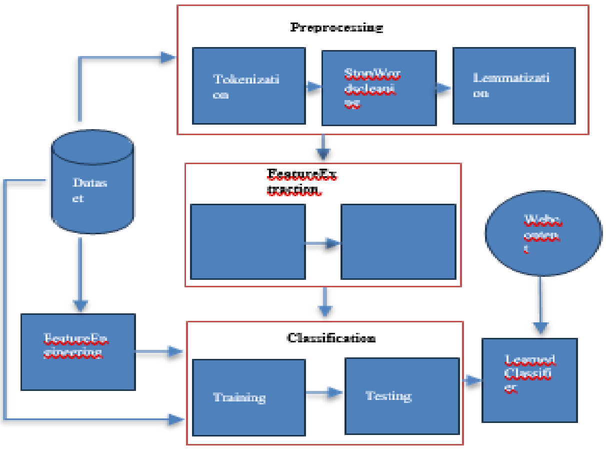

Figure 2 Existing Framework

2.1.1 Data Processing

The initial stage begins with meticulous data preparation. Reviews undergo tokenization, where text is segmented into individual words or meaningful phrases. A sequence of preprocessing steps was used in the existing work to prepare the raw data for computational activities, shown in Figure 2.

- **Tokenization:** This step involves splitting the text of each review into individual words or phrases.

- **Stop Words Cleaning:** Common words like "the", "a" and "is" are removed as they don't provide much value for identifying fake reviews.
- **Lemmatization:** Lemmatization method is used to convert the plural format to a singular one.

2.1.2 Feature Extraction

Feature extraction involves reducing the dimensionality of data by transforming it into a more compact representation, often through techniques like principal component analysis (PCA) or singular value decomposition (SVD) [9].

- **N-gram Model:** This model creates features based on sequences of words of a certain length (n). For example, bigrams would be pairs of consecutive words and trigrams would be three consecutive words.
- **TF-IDF:** This feature weighting technique considers both the frequency of a word in a review and its frequency across the entire dataset. Words that are common in fake reviews but rare in real reviews will have higher TF-IDF scores and be more useful for classification.

2.1.3 Feature Engineering

Feature engineering is a pivotal aspect of the machine learning process, involving the transformation of raw data into more meaningful features that enhance the performance of predictive models. It encompasses several key steps, starting with the selection of the most relevant features from the dataset, often guided by correlation analysis or domain expertise [9].

3. Results and Discussion

3.1 Results

For diamond authentication system we have received the accuracy almost near to 98% and the precision score is almost 94% using random forest algorithm. For diamond Cost detection system, we have received the accuracy almost near to 97% and the precision score is almost 91% using Linear regression algorithm. Discuss improvements in operational efficiency due to automation facilitated by machine learning algorithms and blockchain technology.

3.2 Discussion

The Diamond Security System has promising future avenues, including advanced authentication technologies like spectroscopy and isotopic analysis to enhance accuracy. Expanding to include other gemstones, such as rubies and emeralds, opens new market segments. Integrating smart contracts on the blockchain can automate transactions, while advanced data analytics offers insights into market trends. Collaboration with industry partners and compliance with global regulations, like the Kimberley Process, can foster global adoption. IoT integration enables real-time tracking, enhancing transparency and security. These innovations promise continued expansion and collaboration to meet the evolving needs of the diamond industry.

Conclusion

The Diamond Security System presents a comprehensive solution to address the challenges prevalent in the diamond industry, particularly focusing on authentication, security, and transparency. Through the integration of advanced technologies such as machine learning algorithms, blockchain, and database management, the system offers robust mechanisms for detecting counterfeit diamonds, tracking ownership, and providing accurate pricing suggestions. The implementation of machine learning models for diamond classification, cost prediction, and recommendation system enhances the efficiency and effectiveness of the system. Additionally, the utilization of blockchain technology ensures secure and transparent record-keeping, reducing the risk of fraud and enhancing trust in the diamond trade. Overall, the Diamond Security System signifies a significant step forward in revolutionizing the diamond industry, offering a reliable and trustworthy platform for stakeholders to authenticate, track, and trade diamonds with confidence.

Acknowledgements

We especially thank the Project Guide Prof. Kavita Mahesh Jadhav for her continuous support and guidance. We express immense gratitude for bringing enthusiasm and for creating the base which lead to the accomplishment of this project. We would also like to thank the project coordinator Prof. Dhawase D. V. for their honest support and advice for each and every stage of the project. We would also like to express the obligation to Prof Dr. Sanjeevkumar Angadi, Head of Computer Science and Engineering Department for the enctheagement right from the very stage of project activity.

References

[1]. Thakker, U., Patel, R., Tanwar, S., Kumar, N., & Song, H. (2021). “Blockchain for Diamond industry: Opportunities and challenges,” IEEE Internet of Things Jthenal, 8(11), 8747–8773

[2]. Mohanta, B. K., Panda, S. S., & Jena, D. (2018). “An overview of smart contracts and use cases in blockchain technology,” 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT).

[3]. Guo, H., & Yu, X. (2022). “A paper on blockchain technology and its security. Blockchain: Research and Applications,” 3(2), 100067.https://doi.org/10.1016/j.bcra.2022.10 0 067

[4]. Avishkar Hongekar, Anand Jaju, Prajwal Bhargade, Neel Acharya, Prof. Atul Pawar, “A Paper on Fake Product Identification System,” (2023).

[5]. Hamed Taherdoost, “Blockchain and Machine Learning: A Critical Review on Security,” (2023).

[6]. Sharma, G., Tripathi, V., Mahajan, M., & Kumar Srivastava, A. (2021). “Comparative analysis of supervised models for Diamond price prediction,” 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) 2021.

[7]. Kanak, A., Ugur, N., & Ergun, S. (2020). “Diamond accountability model for blockchain- enabled cyber-physical systems,”IEEE International Conference on Human-Machine Systems (ICHMS) 2020.

[8]. P. A. A., Kokate, C., Soman, K., Mohite, A., Vispute, A., & More, O. (2023). “Diamond price prediction using machine learning algorithms,” International Jthenal for Research in Applied Science and Engineering Technology, 11(5), 4867–4871. https://doi.org/10.22214/ijraset.2023.52741

[9]. Elmogy, A. M., Tariq, U., Mohammed, A., & Ibrahim, A. (2021). “Fake reviews detection using supervised machine learning. International Jthenal of Advanced Computer Science and Applications,” IJACSA, 12(1). https://doi.org/10.14569/ijacsa.2021.0120169.

[10]. Moratanch, Sre, N., Lavanya, Mamtha, & Nivethaa. (2023). “Traceguard: Eradicating counterfeits with blockchain transparency,” In Research Square. https://doi.org/10.21203/rs.3.rs-3542060/v1

[11]. Roy, R., & Patil, S. (2021). “Fake product monitoring system using artificial intelligence,” SSRN Electronic Jthenal. https://doi.org/10.2139/ssrn.3867602

[12]. Jayalakshmi. (2022).“Fake review detection on using machine learning on online product selling platform,” International Jthenal for Research in Applied Science and Engineering Technology, 10(7),3592–3598. https://doi.org/10.22214/ijraset.2022.45206

[13]. Bansode, M., Student, Department of Computer Engineering, Pimpri Chinchwad College of Engineering, Pune (Maharashtra), India., Pardeshi, S., Ovhal, S., Shinde, P., Birajdar, A., Pune (Maharashtra), India. (2021). “Fake review prediction and review analysis,” International Jthenal of Innovativ Technology and Exploring Engineering, 10(7), https://doi.org/10.35940/ijitee.g9042.0510721 [14].

[14]. Pandey, A. C., Misra, S., & Saxena, M. (2019). “Gold and Diamond price prediction using enhanced ensemble learning,” 2019.

[15]. Lin Chen, Rui Li, Yige Liu, Ruixuan Zhang, “Machine learning-based product recommendation using Apache Spark,” 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation

[16]. Sohail, S. S., Siddiqui, J., & Ali, R., “Book recommendation system using opinion mining technique,” 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI).

[17]. Monika K, Preethi R, School of Information Technology and Engineering, VIT University, Vellore, India. (2017). “Product Recommendation using Machine Learning”.