Shopify – Shop Augmentation and Recommendation System

Subhojeet Bhattacharjee
Department of Networking and Communications
SRM Institute of Science and Technology
Chennai, India
teejohbus07@gmail.com

Nishant Singh
Department of Networking and Communications
SRM Institute of Science and Technology
Chennai, India
ns.nishantsingh2705@gmail.com

N. Prasath
Associate Professor - Department of Networking and Communications
SRM Institute of Science and Technology
Chennai, India
prasath283@gmail.com

2023 International Conference on Computer Communication and Informatics (ICCCI) | 979-8-3503-4821-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCCI56745.2023.10128209

Abstract— **Modern times have brought about changes in lifestyle. People dislike visiting crowded markets because they take too long and are uncomfortable. Thus, an advantage of online shopping is, it saves tons of time. Online purchasing is a procedure where clients go directly to a vendor on the Internet to buy products, services, etc. Customers can browse online merchants from the convenience of their home and make purchases while seated at a computer. This research aims to introduce local people to the online shopping application so that local merchants can profit from it and expand their reach and sales through this fantastic platform. This research aims to give clients of physical stores the benefits of internet shopping. It facilitates purchasing goods from any store using a smart device and the internet. In this way, the customer will benefit from home delivery and online shopping services from his preferred retailer. The solution is different in an umpteen number of ways. It is a direct exchange between the buyer and the vendor. This system can be applied to any local business or to international brands with retail store chains in the local areas.**

Keywords- Encryption, Preference Matrix, Feature-based product analysis, Chatbot, AOT Compilation, JPA, Scalability, Security

I. INTRODUCTION

E-commerce has risen to prominence as one of the new purchasing trends on the internet. Online shopping has several benefits, including time savings and the flexibility to buy goods anywhere, anytime [1]. Any business dealing with the transfer of ownership or employment rights to goods or services through a computer network is typically considered to be part of online buying and selling over the internet [2-4]. Three subcategories of e-commerce exist: business-to-business (B2B) (Cisco), business-to-consumer (B2C) (Amazon), and consumer-to-consumer (C2C) (eBay). Developing e-commerce websites involves both front-end and back-end development. Front-end development can make use of HTML, CSS, JavaScript, J-query, and ajax. The main industry norm for creating Web pages is the hypertext markup language (HTML). It mostly serves to describe the Web pages' element-based structure [5-6]. Problems are frequently caused by procedural hiccups, an honest perception of some difficulty, a failure to keep promises, and the desire for perfection. When dealing with a street vendor, it might be difficult to tell the difference between an entrepreneur and his or her company, especially when issues develop. The problem could be associated with a certain person, thing, event, or procedure. Some problems, or problems inside the business unit, are brought on by factors outside the control of the company, and other problems are created by factors inside the control of the corporation. The first is easier to manage than the second [7]. The recommendation system, which proposes goods tailored to the consumer based on the preferences of the client, is one method of making shopping easier for customers [8]. Managing massive amounts of data for consumers is one of the RSs' objectives. RS increases EC sales by bringing in new customers, showcasing new products, developing client loyalty, raising customer happiness, and raising the likelihood that a happy customer will come back. Studies show that RSs are an essential component of the electronic industry, with personalized RSs improving sales by up to 35% through suggested items. As a result, these technologies are playing a bigger role as a way to do person-to-person marketing. [9]. One of the most well-known and commonly used recommender systems is the collaborative filter (CF), which is considered as one of the key elements of effective E-commerce systems [10-11].

The major challenges consist of:

1. Boot-start: The RS works poorly when there is minimal data or metadata available. It is constantly accessible to both new and regular users [12-13].

2. Customers' preference to simply rate a few things leads to data sparsity [14-15].

3. Scalability: Scalability issues have exploded as e-commerce websites have rapidly expanded [16].

4. Diversity: Instead of focusing on differences, suggestions emphasis overlaps. The user is thereby exposed to fewer items overall, and highly important specialized items could go missed [17].

5. One of the main challenges was selecting the project stack from a large selection of frameworks and tools [18].

979-8-3503-4821-7/23/$31.00 ©2023 IEEE

II. RELATED WORKS

Numerous difficulties with the online buying system and recommendation system have been raised in prior work, which has been cited. Performance of e-commerce systems may be improved with the help of recommendation systems. As a result, several research regarding creating diverse recommendation systems utilizing various methodologies were conducted. Some of them were consumer behavior-focused and the studies are as follows:

Yajie Hu, Mitsunori Ogihara 2011 [19] proposed a fresh strategy by gauging a song's novelty using the "Forgetting Curve" and measuring "preferences" using a user log data, the system seeks to suggest songs that the user like, is unfamiliar to them, and fits their listening habits. To ascertain the user's interest level in the next song and their listening preferences are assessed. An analysis of this article based on the RS challenges reveals that when the cold-start issue cannot be fixed, a new client is given a song at random. There was an increase in time complexity that required carrying out multiple computations due to the scalability of the number of songs and the vast diversity of customer tastes.

Bo Wang et al. in 2018 [20]. (BUIF) produced an implicit feedback-based personalized recommendation algorithm. BUIF considers the user's analysis and item sequencing in addition to their purchasing behavior. The user's behavior log is used to extract the user's purchase, comparison, and item patterns. The comparison and purchase behavior of the user are used to determine the user's similarity, and the word-to-product embedding method is used to determine the item's similarity. Similarity measurements can be used to address the challenges of variety and scalability, but not the others.

Kai Wang et al. 2019 [21] created a bespoke e-commerce product recommendation system based on learning clustering representation. The ability of the conventional KNN technique to choose close-by object sets is constrained. They provided a way to pick the set of nearby items as well as a dynamic selection model, time function, and neighborhood factor. The e-commerce product suggestion system was developed by combining RNN and attention processes.

Liao et al [22] proposed knowledge patterns, rules, and knowledge maps were proposed, explained, and solutions for online shopping were provided. Knowledge extraction through data mining. Kuo et al suggested the idea of post-purchase intentions and post-recovery consumer pleasure. Additionally, negative emotions decrease post-recovery satisfaction whereas positive emotions boost it.

Andres Ferraro et al 2018's [23] focused on music recommendations and suggested a novel approach to enhance suggestions based on a desired metric of choice that may be achieved by combining different algorithms for each user independently in accordance with their expected performance. The suggested approach necessitates predicting an error that each system will commit for every user based on that user's previous actions. For a number of measures that forecast system performance based on a variety of characteristics that describe historical user activity in the system, the study recommended a training regression model. The recommendations provided by each system are then combined using a variety of fusion techniques. The resulting hybrid system can be tuned for the desired metric using this technique. According to the evaluation metrics, the hybrid approach can help with concerns relating to sparsity and variety.

Mojtaba Salehi 2019 [24] based on user behaviour, provided some helpful tips. A preference matrix was created to measure consumer preferences based on product qualities since people communicate thoughts intuitively. Weighted association rules are also used to find these patterns, raising the standard of suggestions. The discovery of a user behavior-based recommendation technique that determines users' preferences based on implicit product attribute ratings is the contribution that is most noteworthy. This approach also makes advantage of a sequential buying pattern to raise the calibre of recommendations. This article ignores the remaining issues of the previous RSs and focuses on fixing sparsity and using the product's characteristics to assess how similar they are to one another. Simply comparing two products to determine whether they are similar is insufficient to determine whether a recommendation is sound because not all customers have the same preferences.

Duo Lin, Su Jingtao 2020 [25-26] developed a groundbreaking recommender technique for extensive online shopping platforms. Products with higher preference degrees are then offered to the client. The desired degree for each item is chosen using contextual data from the client, such as access, click, read, and purchase data. Because browsing, rating, and searching were insufficiently important to include, this essay mainly focused on the buy and click activities. Without addressing the other issues, it only managed to solve the sparsity problem.

All of the earlier work was mostly concentrated on static modeling-oriented solutions, which were also less responsive and reliable. With this research, it is possible to identify alternatives to the spring framework and angular that offer more security and are more durable for the long-term development of applications. No one today compromises on security, therefore this also looks at the security side of the code and produces highly responsive web-based technology that is more swift, resilient, and secure.

III. PROPOSED WORK

To assist business owners and consumers in overcoming their difficulties, a user-friendly e-commerce website with a low cost has been developed that enables clients to select their products and make a home buy. The merchants can provide every product through a single window in a manner similar to that.

A. The Description of the Proposed System

The idea of the research paper involves a good deal of deliberation over the idea of using multiple web-based technologies present at the current scenario and cloud infrastructure. In general Shopify is an E-commerce product which tries to reduce the distance between a vendor and a buyer. An ingenious attempt to alleviate the problems of a buyer while preparing his shopping list. It offers the vendor

stores and their products at the fingertips of the customers. What really is the key idea and belief that one day this will make the notion of "Products at our fingertips" true. Shopify tries to improve the market base of vendors in their respective areas by increasing their visibility. It provides the ability to buyers to filter out shops in their areas based on multiple filtering algorithms. The two approaches to do effective resource management are: first, by leveraging elastic load balancers and the cloud to deploy servers as needed, and second, by using the spring tool suite to improve the application's security. Better shared functionality can be enabled by using the aforementioned methods in order to share the same platform or web application without interfering with other modules. Such functionality aids in system upkeep, internal cost control, or charging for third-party services. Additionally, this offers the web application sluggish loading.

This research was conducted using both primary and secondary sources. The following sections contained this collection of data.

Primary data were gathered from internet shoppers in town using a methodically created questionnaire. A questionnaire containing the respondents' preferences as well as their age, gender, and educational background was created in order to conduct statistical inquiries.

Secondary information was gathered from a variety of books, journals, research papers, and websites.

Vendors can use their credentials to sell their goods by including a detailed description and images. Items include things like phones, laptops, computers, and their accessories, as well as things like books, kitchen and home appliances, and so on. In a private or public conversation, the seller can review advertising, control bidders, rate buyers, ask for a payout, and address questions from bidders. The seller determines the sufficient price for the item rather than setting a minimum price. Additionally, the seller chooses the window of opportunity during which bidders may submit their offers; during this window, the seller may decide to feel comfortable selling the item. The buyer will be informed via email or text message that they are not able to put a bid on the chosen item if the seller chooses to block bids. The seller can create a list of pre-approved bids to ensure privacy while selling the goods to a chosen group of buyers. A trade broker who is an experienced seller supports the inexperienced sellers in their product bidding as a guidance for bidders.

Our study's goal is to make more prominent search options available and generate the most relevant results. The proposed work is represented by Algorithm 1:

B. Algorithms Used

- Algorithm 1 for shop augmentation and recommendation

```
Input: Buyer id, Vendor id, Products_id, Customer
feedback
Output: Start the recommender queue
INITIALIZE:
Buyer_id = Vendor_id = Products_id = 0
Every product in the product list has an empty queue:
for every product in the queue
      if product_match == value
       return the products's location
     end if
  end for
ADD Product
SORT the products in following orders:
Purchased products in descending order
Viewed products in descending order
Disliked products in ascending order
Sort(array, left_pointer, right_pointer):
   if left_pointer> right_pointer
      return
   middle value = (left_pointer + right_pointer)/2
   Sort(array, left_pointer, middle_value pointer)
   Sort(array, middle_value+1, right_pointer)
   merge(array, left_pointer, mid_pointer
   , right_pointer)
   recommend_List = Merge all
end
```

- Algorithm 2 for automated chatbot

```
Input: User's Natural language is taken as the input.
Output: A natural language form and the
          conversation log were used to request the
          value of the MES data.
Algorithm: User participation
1. while client ask a qustion
2.      client enters the query into queue
3.      recognized, and keyword extraction
4.      for each keyword key in queue
5.          look for key in MES data set
6.          save key for next loop
7.         return MES data value=val
8.    end for
9.    val predetermined response string concatenation
10. end while
```

C. Architectural Diagram

The architecture design shown in Figure 1 was used to execute the proposed solution. Both the customer and the seller can log in or join up for the application as clients here. The data will be checked against the database's existing data. The data will be shown as the front-end using Angular, HTML, CSS, and TypeScript after successful validation. To build a web socket and handle Web Socket events and actions, the front-end will need a web socket server. It offers methods comparable to the Web Socket client API, and most programming languages support it. The back-end server is built with Spring Boot and Java, and it sends and receives data from and from the front-end. The database is used to hold all of the client and product records in separate tables. The server retrieves this information via data transfer. Amazon Aurora, an AWS service, is used to deploy the database to the cloud.

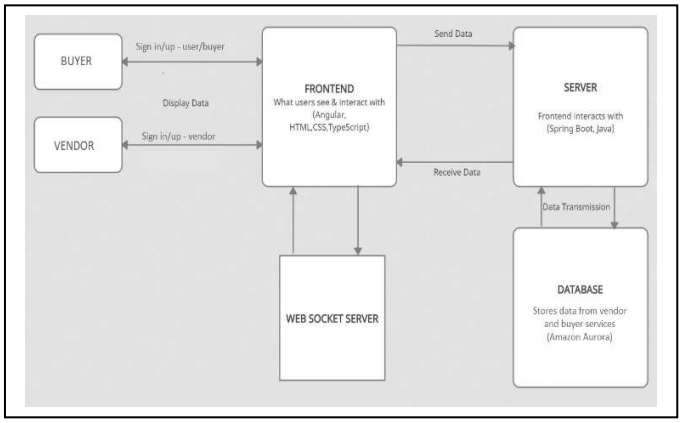

Figure 1. The complete architectural flow of shop augmentation and recommendation System

The vendor module was run using the architectural design depicted in Figure 2. The customer will log in/sign up and then be taken to the dashboard. The customer may use the dashboard to add, update, remove, and view products, among other things. Following these procedures, a preference matrix will be constructed, which will aid in the creation of the recommendation system. The operations are transmitted to the server, which retrieves the data from the database. A feedback system will be maintained to identify the client base's preferences in his/her searches and activities. The recommendation system will create a list of suggestions that will be utilized in the future.

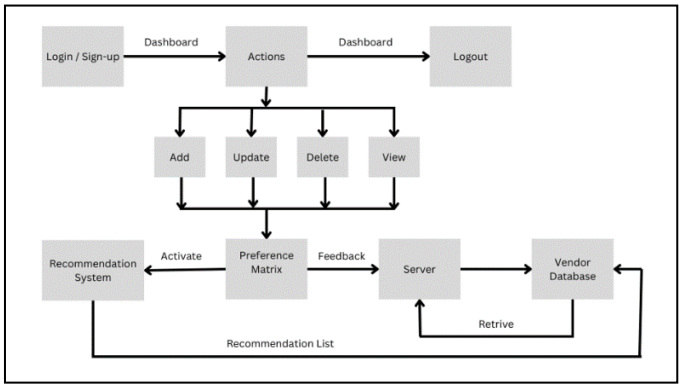

Figure 2. Describes complete vendor module along with multiple services

The buyer module was executed using the architecture concept depicted in Figure 3. The customer will log in/sign up and then be taken to the dashboard. The customer may use the dashboard to browse, search, filter, and evaluate items, among other things. Following these procedures, a preference matrix will be constructed, which will aid in the creation of the recommendation system. The operations are transmitted to the server, which retrieves the data from the database. Based on his or her previous searches, the buyer will be suggested preferred goods. The user may also search for products based on region, zip code, or product category.

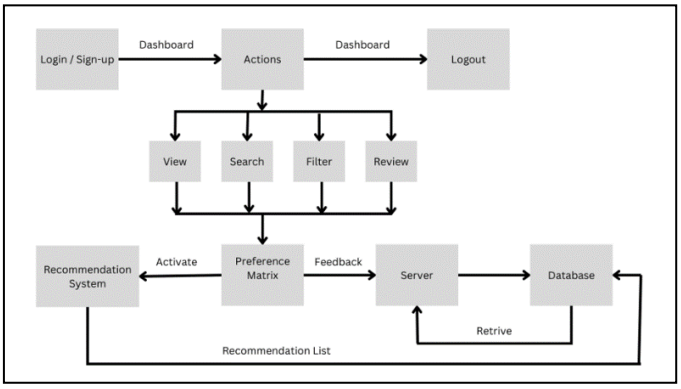

Figure 3. Describes complete vendor module along with multiple services

D. Database Design of Shopify

The system is made up of four tables which are built using MySQL relational database: one with basic information about the buyer, one with basic information about the product categories, one with basic information about the suppliers, and one with the order list. The connection between the linked fields in the tables allows us to determine the relationship. Additionally, Figure 4 illustrates the relationship between the tables:

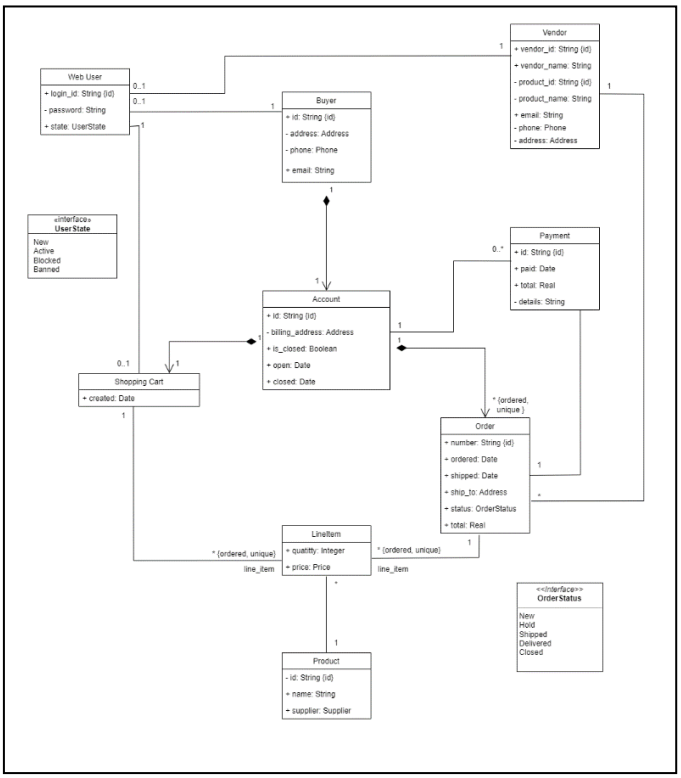

Figure 4. System Database Design

IV. RESULTS DISCUSSIONS

RS determines how similar the users are based on how many common items they have purchased. RS uses the quantity of common consumers who bought the same thing to assess how comparable it is. listing of goods from various industries

that can be obtained from nearby stores serves as a common platform for both buyers and sellers when it comes to local shopping. These industries cover a variety of areas, such as apparel, food products, raw materials, and groceries. Anyone interested in selling products is welcome to sign up on Shopify.

Clustering coefficients for several graph classes are displayed in Figure 5. Except for babies, all other categories have a higher clustering coefficient than the corresponding random graph (with the same number of edges).

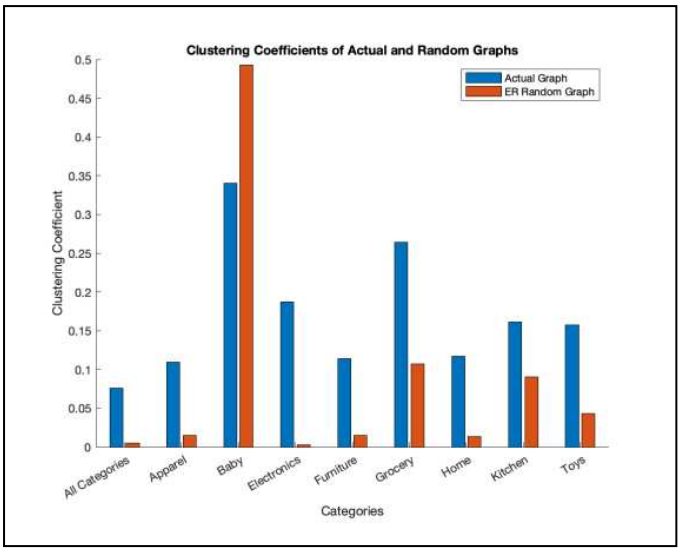

Figure 5. Categories vs Clustering Coefficient

Probability of a graph edge across many classes is depicted in Figure 6. In terms of likelihood, the baby category is over 10% more likely than the grocery category. The likelihood is the sum of all potential edges in the undirected graph (n(n-1)), where n is the number of vertices) of products that were both seen or purchased together.

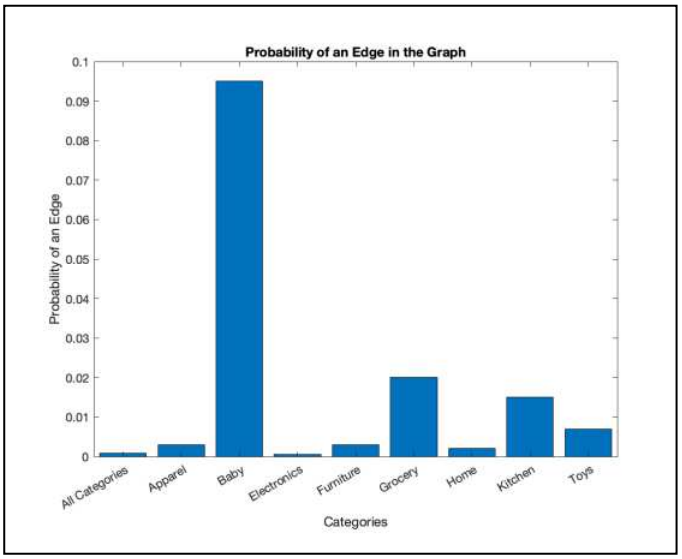

Figure 6. Categories vs Probability of an Edge

Figure 7 details the quantity of products in each category, allowing sellers to calculate their potential revenue and determine how much of each product should be kept in store.

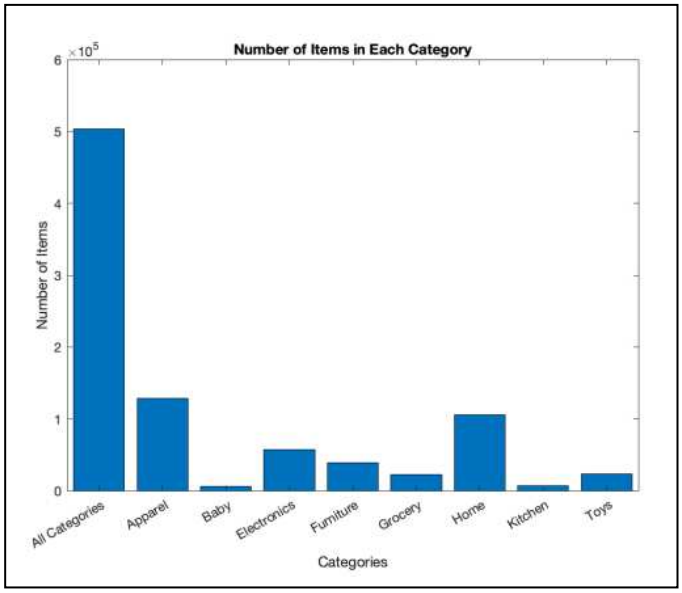

Figure 7. Categories vs Number of items

The Ranking Technique was used to assess the data that had been obtained using Preference Ranking Algorithm below.

$$W_i = (1-d) + d \sum_{i=1\, i \neq j}^{N} l_{i,j} \frac{W_i}{n_i} \tag{1}$$

Using a matrix table, the resultant rank for each issue is converted into a score. The sum of each respondent's scores is divided by the total number of respondents whose scores are added together. The rankings are presented in ascending order according to importance, with the mean scores for each factor. The ranking approach is used to examine online purchasers' preferences for products.

A. Discussions

The proposed work is able to declare that diversity and scalability issues are under control because the suggested website that sells computers and the things that go with them uses the suggested RS (Recommended System) system. Problems like these can be effectively addressed in more comprehensive e-commerce settings including many categories, such as apparel, electronics, and personal care items, by employing additional factors categorized according to the products' categories. Figure 6 depicts the probability of an edge, while Figure 5 depicts the clustering coefficient (CC). There is a strong relationship between these two metrics (R = 0.85). These two quantities are symbolic of the cohesiveness of the product-to-product network that how often two products are considered or purchased together. The grocery, kitchen, and toy categories appear to have a good level of clustering, but when compared to the CC of all categories, there is space for improvement.

V. CONCLUSION

The In this study, a recommendation system is offered along with a discussion of the RS (Recommended Systems) concerns. Scalability, diversity, sparsity, and cold start are a few of these challenges. As seen in linked papers, this paper addresses some of these issues but not all of them. The suggested system computes a number of characteristics (consumer behavior) using statistical methods and analysis, and then presents a list of recommendations that are customized to the users' tastes.

Future work might be to analyze how the various categories are combined. Observing how the products cluster when the graph is made up of pairs of categories may be especially useful. This would aid in identifying the product categories that a buyer is most likely to purchase and boost vendor revenue.

REFERENCES

[1] Liao, Shu-hsien, Pei-hui Chu, Yin-ju Chen, and Chia-Chen Chang. "Mining customer knowledge for exploring online group buying behavior." Expert Systems with Applications 39, no. 3 (2012): 3708-3716.

[2] Niranjanamurthy, M., N. Kavyashree, S. Jagannath, and Dharmendra Chahar. "Analysis of e-commerce and m-commerce: advantages, limitations and security issues." International Journal of Advanced Research in Computer and Communication Engineering 2, no. 6 (2013): 2360-2370.

[3] Sivaji, Ashok, Alan G. Downe, Muhammad Fahmi Mazlan, Shi-Tzuaan Soo, and Azween Abdullah. "Importance of incorporating fundamental usability with social & trust elements for e-commerce website." In 2011 International Conference on Business, Engineering and Industrial Applications, pp. 221-226. IEEE, 2011.

[4] Gupta, Swati. "Online shopping cart application." (2013).

[5] Buzzi, Maria Claudia, Marina Buzzi, Barbara Leporini, and Fahim Akhter. "User trust in ecommerce services: perception via screen reader." In 2009 international conference on new trends in information and service science, pp. 1166-1171. IEEE, 2009.

[6] Kiran, PB Narendra, and G. N. P. V. Babu. "Problems and prospects of street vendors: a study with reference to Visakhapatnam city." International Journal of Management, Technology and Engineering 9, no. 6 (2019): 2249-7455.

[7] Elahi, Ehtsham, and Ashok Chandrashekar. "Learning representations of hierarchical slates in collaborative filtering." In Fourteenth ACM Conference on Recommender Systems, pp. 703-707. 2020.

[8] Abbas, Ayad R. "An Adaptive E-Commerce: Applying of Psychological Testing Method to Improve Buying Decision Process." Engineering and Technology Journal 33, no. 2 Part (B) Scientific (2015).

[9] Bortko, Kamil, Piotr Bartków, Jarosław Jankowski, Damian Kuras, and Piotr Sulikowski. "Multi-criteria evaluation of recommending interfaces towards habituation reduction and limited negative impact on user experience." Procedia Computer Science 159 (2019): 2240-2248.

[10] Lee, Dokyun, and Kartik Hosanagar. "How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment." Information Systems Research 30, no. 1 (2019): 239-259.

[11] Shareef, Sarah M., and Soukaena H. Hashim. "Proposed hybrid classifier to improve network intrusion detection system using data mining techniques." Engineering and Technology Journal 38, no. 1 (2020): 6-14.

[12] Zhang, Hongwei, Xiangwei Kong, and Yujia Zhang. "Selective Knowledge Transfer for Cross-Domain Collaborative Recommendation." IEEE Access 9 (2021): 48039-48051.

[13] Wang, Donghui, Yanchun Liang, Dong Xu, Xiaoyue Feng, and Renchu Guan. "A content-based recommender system for computer science publications." Knowledge-Based Systems 157 (2018): 1-9.

[14] Mohammed, Mazin Abed, Dheyaa Ahmed Ibrahim, and Akbal Omran Salman. "Adaptive intelligent learning approach based on visual anti-spam email model for multi-natural language." Journal of Intelligent Systems 30, no. 1 (2021): 774-792.

[15] Sahoo, Abhaya Kumar, Chittaranjan Pradhan, Rabindra Kumar Barik, and Harishchandra Dubey. "DeepReco: deep learning based health recommender system using collaborative filtering." Computation 7, no. 2 (2019): 25.

[16] Fayyaz, Zeshan, Mahsa Ebrahimian, Dina Nawara, Ahmed Ibrahim, and Rasha Kashef. "Recommendation systems: Algorithms, challenges, metrics, and business opportunities." applied sciences 10, no. 21 (2020): 7748.

[17] Mostafa, Salama A., Saraswathy Shamini Gunasekaran, Aida Mustapha, Mazin Abed Mohammed, and Wafaa Mustafa Abduallah. "Modelling an adjustable autonomous multi-agent internet of things system for elderly smart home." In International conference on applied human factors and ergonomics, pp. 301-311. Springer, Cham, 2019.

[18] Kuanr, Madhusree, B. Kesari Rath, and S. Nandan Mohanty. "Crop recommender system for the farmers using mamdani fuzzy inference model." International Journal of Engineering & Technology 7, no. 4.15 (2018): 277-280.

[19] Sulikowski, Piotr, and Tomasz Zdziebko. "Deep learning-enhanced framework for performance evaluation of a recommending interface with varied recommendation position and intensity based on eye-tracking equipment data processing." Electronics 9, no. 2 (2020): 266.

[20] Srifi, Mehdi, Ahmed Oussous, Ayoub Ait Lahcen, and Salma Mouline. "Recommender systems based on collaborative filtering using review texts—a survey." Information 11, no. 6 (2020): 317.

[21] Wei-wei, Guo, and Liu Feng. "Application Research of Hadoop's Weibo Recommendation System Prototype Based on Customer Dynamic Behavior." In 2018 3rd International Conference on Smart City and Systems Engineering (ICSCSE), pp. 562-566. IEEE, 2018.

[22] Park, Chanyoung, Donghyun Kim, Min-Chul Yang, Jung-Tae Lee, and Hwanjo Yu. "Click-aware purchase prediction with push at the top." Information Sciences 521 (2020): 350-364.

[23] Li, Xixi, Jiahao Xing, Haihui Wang, Lingfang Zheng, Suling Jia, and Qiang Wang. "A hybrid recommendation method based on feature for offline book personalization." arXiv preprint arXiv:1804.11335 (2018).

[24] Jiao, Ming-hai, Xiao-fang Chen, Zhong-hai Su, and Xu Chen. "Research on personalized recommendation optimization of E-commerce system based on customer trade behaviour data." In 2016 Chinese Control and Decision Conference (CCDC), pp. 6506-6511. IEEE, 2016.

[25] Li, Dancheng, Guangming Zhao, Zhi Wang, Wenjia Ma, and Ying Liu. "A method of purchase prediction based on user behavior log." In 2015 IEEE International Conference on Data Mining Workshop (ICDMW), pp. 1031-1039. IEEE, 2015.

[26] Xu, Jing, Jie Wang, Ye Tian, Jiangpeng Yan, Xiu Li, and Xin Gao. "SE-stacking: Improving user purchase behavior prediction by information fusion and ensemble learning." Plos one 15, no. 11 (2020): e0242629.