Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 04 (2023)

Reinventing Fashion Retail with Data-Driven Solutions

Kavija Sapukotana[1], Shadhir Ameen[2], Ashen Pradeep[3],Kulana Thathsara[4], E. Weerasinghe[5],Dilshan De Silva[6]

Dept. of Computer Science and Software Engineering, Sri Lanka Institute of Information Technology, Malabe, Sri Lanka.

Abstract:-In this modern technological era, it has been normalized to have an e-commerce web application for each business around the globe. When it comes to garments business it more than a necessity it has become a fashion symbol to human nowadays, so for such kind of business having an e-commerce web application might be helpful but if it can be more customized into further aspects like smart item suggestions and having a sales prediction mechanism, the business which holds such kind of technology are the ones who dominate the market in present. Smart suggestions help to increase sales by pursuing the customers and ensures effective customer satisfaction. This smart item suggestion is achieved through collaborative filtering using user-based suggestions and item-based suggestions mainly focusing the budget as the main constraint. Also having sales predicting technology helps to manage the inventory effectively and to reduce wastage of the resources of the business. The research mainly focuses on a perfect blend of elevating sales through suggesting items to customers and getting sales predictions.

Keywords*: User preferences, machine learning, fashion retail, collaborative filtering, sales forecasting, extra tree regressor, scikit-learn, cosine-similarity.*

1. Introduction

Humans have been evolving since the stone age. At first dress was a practical necessity for humans but along with evolution it has been more than a necessity, it has been a way of expression and identity factor. The dress what was used to keep humans from external elements is now a showcase of identity, status, and beliefs. From creating a single piece of clothing for themselves humans are focused on mass production of clothing and retailing. Fashion has been through many turning points from early ages to the present. But fashion combining with digitalization is the one of the biggest achievements of humans. This allows humans to discover and access trends and styles globally. But in the dynamic landscape even though the fashion retail industry has been digitalized optimizing the sales of retail fashion has been crucial.

E-Commerce platforms have been evolving since the beginning. With machine learning and deep learning the E-commerce web sites became more intelligent and personalized, offering tailored recommendations, predicting consumer behavior, and enhancing the overall shopping experience. Previously, this topic has been researched extensively. However, the number of academic articles on this topic is limited, and existing studies lack comprehensive reviews of fashion recommendation systems and their filtering techniques. Now many filtering techniques available for recommendation systems [1].

Online platforms for fashion retails provide convenience to users but the question is whether are they being managed and optimized in the way it should be. And are the existing online platforms being able to provide personalized shopping experience. In this research the expected outcome is to harness the potential of data analytics and machine learning and aim to elevate the shopping journey by delivering optimized inventory management using sales forecasting and tailored product suggestions.

This is about the insufficiency and lack of personalization observed in online retails fashion. Instead of the traditional methods the purposed solution optimizes stock levels by accurately predicting the sales.Numerous e-

commerce platforms offer a wide range and variety of products. This makes the customer struggle when it comes to making selections aligned with the budget of customer.

This research is to bridge these gaps by developing a web application that employs machine learning technics which will have optimized inventory management and personalized dress recommendations based on user inputs.

The significance of this research is to reshape the landscape of the retail fashion business with the help of ML models. The above-mentioned web application solution can significantly shape inventory management by avoiding over or under supply using sales forecasting. The tailor dress suggestion methods which will be generation clothing bundles according to the budget and other demanding factors like gender, user preference, trending cloths, promotional offers, and seasonal discounts can ensure the user satisfaction and guide users out from their struggle of selection.

Using this user centric and data driven approach the study shows the impact on business optimized solutions in fashion retail.

The main objective of this research is to come up with novel ideas to improve sales performance when it comes to fashion retail. Implementation of the above methods will be further enhanced with the usage of machine learning. To optimize sales, the following have been identified as key components which includes cart management, inventory management, product management and clothing suggestions. For inventory management, sales forecasting will be utilized which will further improve the inventory management system which will ensure that the in-demand products are readily available for the customer to purchase and for clothing suggestion, machine learning will be utilized which will consider certain factors provided by the customer themselves and providing personalized recommendations. Within the research effect of using machine learning for generation of dynamic recommendations and usage of sales forecasting to management of the inventory, whether utilization of machine learning will improve optimization of sales or will it prove to be inefficient.

Research Questions

RQ1: How will personalized recommendations help in the improvement of sales numbers and customer engagement with the system?
RQ2: How will implementation of machine learning and data analytics help in predicting sales forecasting and sales optimization.
RQ3: How will implementation of machine learning techniques affect inventory management?
RQ4: How will changing customer trends affect sales optimization techniques that are been utilized within the system.

Within the study, theyhave formulated several hypotheses to investigate the impact of various factors on sales numbers, customer engagement, sales forecasting, sales optimization, and inventory management.

Hypothesis 1:
Null hypothesis: Personalized recommendations will not increase the sales numbers and customer engagement.
Alternative hypothesis: Personalized recommendations will increase the sales numbers and customer engagement.

Hypothesis 2:
Null hypothesis: Implementation of machine learning and data analytics will not help in sales forecasting and sales optimization.
Alternative hypothesis: Implementation of machine learning and data analytics will help in sales forecasting and sales optimization.

Hypothesis 3:
Null hypothesis: Machine learning will not affect inventory management.
Alternative hypothesis: Machine learning will affect inventory management.

Hypothesis 4:
Null hypothesis: The changing customer trends will not affect the sales optimization techniques.
Alternative hypothesis: The changing customer trends will affect the sales optimization techniques significantly.

The following section includes reviewed highlights of sales optimization solutions that exist within the fashion retail industry where most of the system were utilizing predictive algorithm and artificial neural networks. Section 3 provides machine learning driven suggestions and data filtering methods for pattern analysis and filtering. For the inventory management section, methods of data collection, data refinement, data visualization and feature scaling were done. In this study the outputted results were like for each selected product up to 10 suggestions were generated and also for each item in the inventory a unique sales forecast is generated which is explained in section4.

2. Literature Review

In this section the existing solutions for sales optimized solutions for fashion retail will be considered and the steps and the paths taken by previous research's will be discussed to get an idea about how the fashion retail business optimization has evolved over the years.

The research [2]conducted by S. Sharma, L. Koehl, P. Bruniaux, X. Zeng, and Z. Wang proposed a solution to the fashion retail industry which will be using technology like 3D garment CAD which will be using parameters like design expertise, customer preferences and technical parameters and combining interactions between designer, manufacturer, and consumer to generate personalized design recommendations. The system focuses on men's shirt designs to generate personalized design recommendations.

The research[3] conducted by C. Giri, S. Jain, X. Zeng, and P. Bruniaux indicated the contribution of Artificial Intelligence over fashion retail. This research is done to find out the research that are conducted over past decades on various stages of supply chain. The research I conducted using 149 articles from "Scopus" and "Web of Science." And categorized them according to AI methods and supply chain stages. This research also focuses on Business-to-Business (B2B) and Business-to-Consumer(B2C). By this research the major objective was to benefit researchers and academic industry practitioners in fashion retail.

The research [4]conducted by Riaz, Hadiqa & Baig, Umair & Meidute-Kavaliauskiene, Ieva & Ahmed, Hassaan indicates the impact of omnichannel retailing on customer experience of Pakistan fashion retails industry. The research was conducted using 265 customers from various fashion brands to understand the connection between customer behavior and customer experience. Results analyzed using structural equation model and found out that customer behavior plays a significant crucial role in enhancing customer experience. Researchers found out that drivers of improved customer experience were omnichannel integration, order fulfillment, usability, and seamlessness. The research offers valuable insights to omnichannel strategies for fashion brands in Pakistan and enhances customer engagement in different channels.

The research [5]conducted by Bellini, Pierfrancesco & Ipsaro Palesi, Luciano & Nesi, Paolo & Pantaleo, Gianni indicates the significance of each recommendation systems available and discusses the challenges when using these kind of recommendation systems and each system is assessed individually to find out its unique advantages, disadvantages, and the questions to be taken to mind. Also, the data sets, platform and performance metrics are checked against each recommendation system. By gap analysis the research provides guidance to future efficient developments.

The research [6]handle by N.B Ganhewa, S.M.L.B Abeyrathne, G.D.S Chathurika, Dilani Lunugalage, Dilshan De Silva proposed on maximizing sales in woman's clothing segment. The research focused on developing ML to develop a web application. The study aims to provide forecasting sales, consumer segmentation and characteristics of the most sought-after fashion items.

A survey paper conducted on an overview of recent progress in the field of sales forecasting by Beheshti et al[7] suggests that conventional forecasting methods face challenges in producing accurate sales data for new products and consumer-oriented goods. As a solution the research has introduced Moreover, hybrid forecasting models that perform more accurately. Moreover, the paper discusses literature focusing on the predictive value

of user-generated content and search queries. According to their conclusions the measurement of the predictive value of user-generated content is a valuable aspect in sales prediction in the fashion industry.

3. Methodology

In this section paths and methods for a web based online shopping platform is discussed. This web app's purpose is a solution for personalized fashion suggestions based on user preferences, budget of the user, seasonal offers, trending products and promotions. Also, an inventory system that runs based on sales forecasting is explained based on previous research. This system has four main comportments which are ML based fashion suggestions, cart management, inventory management and product management. Fashion suggestions and inventory management component relay on ML models. While the other components establish the foundational framework between these modules for seamless collaboration. The system is built using streamlit python framework. For the virtual environment pipenv was used.

A. ML based fashion suggestions

For suggesting customers, personalized clothes, ML will be utilized. Suggesting clothes to the customer is done by a recommendation system. The data that is needed for the recommendation system was collected by a dataset on data.world web site. It comes from NewChic.com's products catalog - E-commerce Retail site. It contains 9 categories of fashions separately for this research 3 of the 9 categories are using. Those are Men, women, and kids' data sets. These data set contain item name, item price, discounts, brands, subcategory, colors, images, and likes count. For the recommendation system, collaborative filtering and budget constraints can be implemented.

1) Collaborative Filtering[8]

Based on similarities between users and user interactions such as user preferences, rating, interactions with clothes. Collaborative filtering has two main types, User-Based collaborative filtering, and Item-Based collaborative filtering. User-Base collaborative filtering approach identifies a user who likes cloth "x" while another user who has a similar preference might like cloth "x" as shown in below "Fig 1". Item-Based filtering recommends cloths a user has interacted with, for example a user interacts with cloth "y", user might like cloths similar to "y". By using collaborative filtering user preferences can be gathered as shown in below "figure 1".

2) Budget considerations

This will recommend the clothes under the user-defined budget. The system uses budget conditions at post filtering, it would remove the cloths that exceed the user-defined budget after the recommendations are generated.

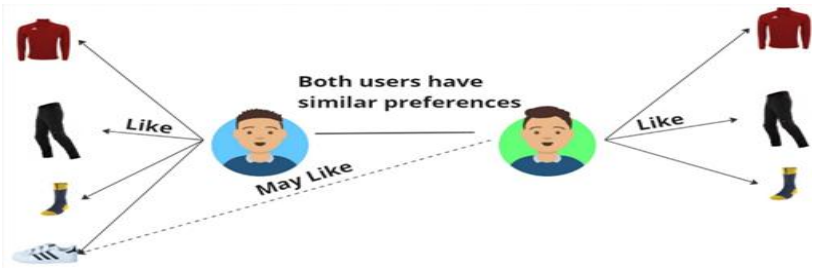


Fig. 1 User Base Collaborative Filtering

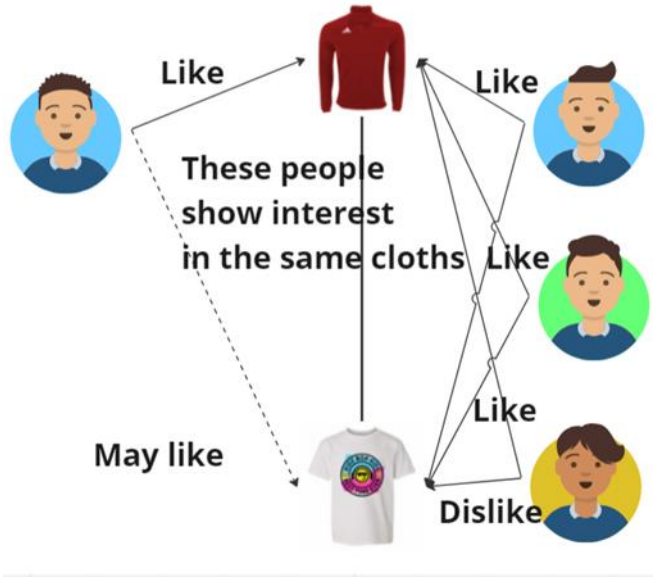


Fig. 2 User Base Collaborative Filtering

User base collaborative filtering works by getting the users rating for purchased items. First the libraries that need imported pandas, numpy, SciPy. Stats and sklearn cosine_similarity. Then get ratings of the users to the purchased items and items details. After that the two tables were merged. Next remove the items that have low ratings. Then the dataset transforms into a matrix format. Fashion products are put into columns and each row represents user ratings. If a user does not purchase or rate the item, the corresponding cell was null. Normalization was done to remove users how give much higher values than the average user. After the normalization the matrix has positive and negative values. Positive mean users rating for the product was higher than average if negative rating that given to the product was lower than average. Then the similar users can be identified by cosine similarity. To implement cosine similarity user must fill null values with 0. After that remove the user that is getting the recommendations from the matrix.

Next define the number of similar users that needed to get recommendations. To find a similar user a positive threshold must be set, which is 0.8 for the cosine similarity. Then remove the items that the user has already bought, and the similar users do not buy. Finally, the system can recommend items by weighted average of user similarity and rating[9].

Item based collaborative filtering works by selecting an item that needs to find matching recommendations. First the libraries that needed to build the system were imported, sklearn TF-IDF vectorizer and sklearn linear_kernel Then finding from which category the recommendation system get the data set from. For example, if user selects an item in men's category the men data set should be selected for find recommendations. Then drop the unnecessary columns from the data set. Then remove the rows which have null values. After that add the suggesting factors, that means the items recommending based on these factors. Which is brand, color, subcategory. When considering the factors common English word are ignored in this process like "the", "and" "is" etc. TF-IDF vectorizer learns the vocabulary and IDF (Inverse Document Frequency) from factorized dataset text-data then transform into numerical vectors. Then cosine similarity matrix is computed using linear_kernel which calculates the similarity between all pair's items based on the TF-IDF representation. Next generate a list of similarity scores for items using cosine_sim_matrix then sort the items based on the score. Then filter the data using budget constraints and discount constraints. Display the 10 items on top of the list. After displaying the recommended items, user can order those. And get similar products for that item as well.

B. Inventory management

To improve the accuracy of future sales predictions within the sales forecasting component, an ensemble learning technique known as the Extra Trees Regressor was utilized which was proven to provide more accurate predictions when compared to other machine learning models such as KNN regression model, SVR regression model[10]. The datasets used as input for the sales forecasting model are obtained from the products catalog-from E-commerce Retail site NewChic.com which contains historical records of items sold along with the necessary information. The dataset consists of three main components which have been categorized as women, men, and kids according to different clothing categories. The women's category is composed of 14810 items, men's category is composed of 10209 items and kid's category is composed of 4086 items with 22 features which includes information such as category, subcategory, name, current price, likes to count from August of 2020.

1) Data Pre-processing: Before the dataset was introduced to the model, it underwent data pre-processing where raw data was cleaned and organized in a way which suited the needs of the sales forecasting model.

a) Cleaning the data: Within this process Item records which contained inconsistencies were removed to ensure that the dataset will provide accurate outputs and to ensure the dataset is error free. Initially, item records with null values in specified columns such as name, variation_0_image, current price and subcategory were removed, and URLs of item pictures were checked and the items that consisted of URLs that contained errors were removed. Then Boolean indexing was utilized to filter out URLs that contained a specific substring.

b) Exploratory data analysis (EDA): To better understand the dataset and to summarize the dataset, exploratory data analysis was performed. Through this step the data was visualized using scatter plots and histograms and categorical data were visualized by using bar plots. For the purpose of relationship analysis, correlation matrix was used which was visualized through a heatmap. Shown in below "Fig 3."

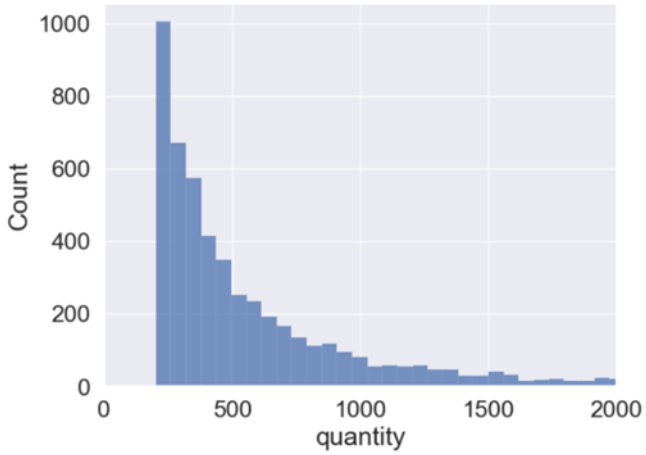


Fig. 3 Histogram for inventory quantity

c) Feature engineering: Due to the lack of the quantity sold feature within the selected dataset, likes count was utilized to generate the quantity sold. Assumptions were made including, every person who buys an item must leave a like to the item that they bought. To generate the feature units in the inventory, the likes count was multiplied by 1.5 and used as the number of quantities that are available from each item within the inventory.

2) *Sales forecasting model:* For the purpose of predicting future sales quantities, the Extra Trees Regressor model was used. This model was specifically chosen due to its ability to capture and handle complex relationships within the dataset. As the target variables, the quantity of items sold was used as target variable (y) and the quantity to predict the sales was used as target variable (x) which was given a constant value for 10 rather than getting the inputs from the user every single time. The Extra Trees Regressor model was an ensemble of 100 decision trees and to ensure reproducibility of the results, the random state seed was set to 0. For the inventory management component, the results generated from the sales forecasting were utilized. The results which were obtained from the sales forecasting were compared with the inventory levels and if the

forecasted value was greater than or equal to the inventory level of that particular item, then a purchase order will be generated where the ordering quantity will be same as the predicted sales value.

C. Product management

As an admin who is interacting in this web application is privilege to manage the products which are sold in this store. Admin is the privilege user to insert new product line to the shopping web application. Also, the admin can view all the available product lines that are currently being sold using the web application and has the ability to update or delete the desired product lines.

D. Cart management

In Cart management component have techniques to enhance the user experience for customers which will ultimately increase conversion rates and sales. These tactics and techniques improve the shopping cart process, from adding items to the cart to completing the purchase.

1) *Clear and Intuitive Design*: To provide a friendly user experience the shopping cart has been optimized to be intuitive and easy to navigate. With this design customers can easily view, edit, and remove items from their cart.
2) *Product Thumbnails and Information:* The cart includes clear high-quality images of the product alongside essential information such as item name, category, quantity, price, and subtotal.
3) *Real-Time Cart Updates:* The cart has been implemented with functions that update the cart total in real-time as customers add or remove items.
4) *Clear Call-to-Action Buttons:* Use prominent, clear, and persuasive buttons for actions like "Make Purchase" and "Remove Item."
5) *Purchase History Analysis:* Involves with recording and examining of customer behaviors on purchases. Suggestions and inventory predictions are based on purchase history.

4. Results & Discussion

A. Fashion suggestion

When a user selects an item that need to get matching fashions to item in their cart as shown in below Figure 4. Then user can enter a budget constraints and minimum discount rate for the recommendation that suggesting. Then 10 recommended items are displayed under the recommended item. Users can order from those

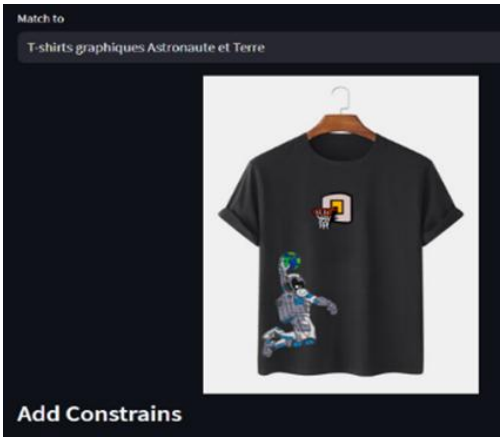
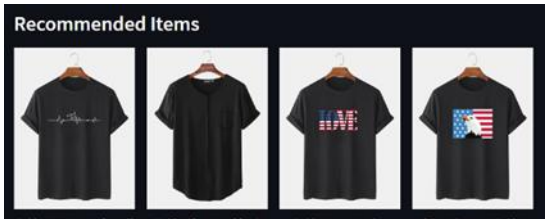


Fig. 4 Select item.

recommended items as shown in below Figure 5

Once a user rates an item, the user can view items that other users have similar user preference as the user. Items he already bought are not displayed. If there are no similar users, the system displays there are no similar users found.



B. Inventory sales forecasting

When the administrator of the website visits the sales forecasting section, they will be able to select items from a drop box and the inventory level of that selected item, predicted sales amount, last month's sales figures and the item status will be displayed as shown in below "Figure 6". After the system predicts the sales quantity, it will compare the obtained value with the current inventory levels of the selected item and if the quantity of the item in the inventory is less than or equal to the value that was predicted, the system will automatically generate an order where the predicted amount of quantity of the selected item will be ordered. "Figure 7"shows a snapshot of the database before sales prediction was conducted while "Figure 8" shows a snapshot of the database after sales prediction was carried out and the ordered item was received.

Fig. 6 Sales prediction demo

category	subcategory	name	price	discount	quantity
men	Shirts	Chemises à manches court…	17.99	50	212

Fig. 7 Snapshot of database before sales prediction

category	subcategory	name	price	discount	quantity
men	Shirts	Chemises à manches court…	17.99	50	702

Fig. 8 Snapshot of database after sales prediction

5. Conclusion

This conducted research is helpful to get insights to the fashion retail e commerce site to enhance and optimize the business and to increase economic growth. The research states the optimized solutions for the business using machine learning techniques like collaborative filtering achieved through user-based filtering and item-based

filtering based on budget considerations through this model the business gets the ability to make effective sales to the customers and pursue the customers to buy more products which leads the business to grow economically. The research shows application of data analysis and usage of machine learning model to optimization of sales. This was achieved through sales forecasting which utilized the extra trees regressor model which was fed with the previous month's sales figures along with quantity to predict, to predict future sales. The results that were generated through the model were then utilized to manage the inventory. The forecasted value is first compared with the available inventory quantity levels and a decision is made whether that particular item will be in stock or out of stock.

In future the research can be further expanded and be modified to get the predictions of user suggest products through userbase collaborative filtering, the collected results can be sorted and get the most future demanding and give the ability to make sales predictions and manage inventory effectively.

References

[1] S. Chakraborty, Md. S. Hoque, N. R. Jeem, M. C. Biswas, D. Bardhan, and E. Lobatón, "Fashion Recommendation Systems, Models and Methods: a review," *Informatics (Basel)*, vol. 8, no. 3, p. 49, Jul. 2021, doi: 10.3390/informatics8030049.

[2] S. Sharma, L. Koehl, P. Bruniaux, X. Zeng, and Z. Wang, "Development of an intelligent Data-Driven system to recommend personalized fashion design solutions," *Sensors*, vol. 21, no. 12, p. 4239, Jun. 2021, doi: 10.3390/s21124239.

[3] C. Giri, S. Jain, X. Zeng, and P. Bruniaux, "A detailed review of artificial intelligence applied in the fashion and apparel industry," *IEEE Access*, vol. 7, pp. 95376–95396, Jan. 2019, doi: 10.1109/access.2019.2928979.

[4] H. Riaz, U. Baig, I. Meidutė- Kavaliauskienė, and H. Ahmed, "Factors Effecting Omnichannel Customer Experience: Evidence from Fashion Retail," *Information*, vol. 13, no. 1, p. 12, Dec. 2021, doi: 10.3390/info13010012.

[5] D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," *Journal of Big Data*, vol. 9, no. 1, May 2022, doi: 10.1186/s40537-022-00592-5.

[6] N. B. S. Ganhewa, M. L. B. Abeyratne, G. D. S. Chathurika, D. Lunugalage and D. D. Silva, "Sales Optimization Solution for Fashion Retail," *Third IEEE International Conference on Advancements in Computing*, pp. 443-448, 2021.

[7] S. Beheshti-Kashi, H. R. Karimi, K. Thoben, M. Lütjen, and M. Teucke, "A survey on retail sales forecasting and prediction in fashion markets," *Systems Science & Control Engineering*, vol. 3, no. 1, pp. 154–161, Dec. 2014, doi: 10.1080/21642583.2014.999389.

[8] Baldha and Shivam, "Introduction to Collaborative Filtering," Analytics Vidhya, 25 February 2022. [Online]. Available: https://www.analyticsvidhya.com/blog/2022/02/introduction-to-collaborative-filtering/.

[9] GrabNGoInfo, "Recommendation System: User-Based Collaborative Filtering," Medium, Apr. 2022. [Online]. Available: https://medium.com/grabngoinfo/recommendation-system-user-based-collaborative-filtering-a2e76e3e15c4.

[10] S. Raizada and J. R. Saini, "Comparative analysis of supervised machine learning techniques for sales forecasting," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, Jan. 2021, doi: 10.14569/ijacsa.2021.0121112.