P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol.29 No.2 April - June 2024
Page: [114-121]
Original research article

Plus-Size Clothing Recommendation System Based on Sales Transaction Data Using FP-Growth Algorithm

Peeraphol Kamolnawin, Nuchjarin Intalar*, Saowaluk Watanapa, Thapana Boonchoo

Department of Computer Science, Faculty of Science and Technology, Thammasat University, Pathum Thani 12120, Thailand

Received 15 January 2024; Received in revised form 28 February 2024
Accepted 1 April 2024; Available online 25 June 2024

ABSTRACT

The recommendation systems for the fashion industry have been widely studied. However, a study on the recommendation systems for the plus-size fashion market is limited. This study investigates the plus-size clothing sales recommendation system on an e-commerce platform in Thailand using the FP-Growth algorithm. A total of 26,993 transaction records were collected during January 2021 to December 2022. The significance of the discovered rules is also examined to gain implicit knowledge about customer preferences in the plus-size clothing domain. The findings show the lists of association rules of the items that customers frequently purchase together. The model has a precision value of 7.88%, a recall of 4.23%, and an F1-score of 5.50%. The relatively low performance of the FP-Growth algorithm indicated the challenges of developing a recommendation system in this domain due to the limitation of the dataset and the diverse purchasing behavior of the plus-size customers. This study fills a research gap and provides a foundation for future research in recommendation systems for plus-size clothing.

Keywords: E-Commerce; FP-Growth; Plus-size clothing; Recommendation system

1. Introduction

The trend of the body positivity movement by accepting diverse body types and inclusivity is changing the landscape in the fashion industry. The demand for stylish and diverse apparel in larger sizes is steadily increasing. In 2023, the market value of plus-size apparel was at USD 288 billion. It is forecasted to reach USD 501.35 billion by 2033 [1].

Although the demand for plus-size clothing is growing, there are many challenges to develop an effective recommendation system in this domain. First, the available data is limited compared to standard-sized clothing. There is a small

*Corresponding author: njrin@tu.ac.th
doi: 10.14456/scitechasia.2024.30

number of fashion offers in the plus-size segments [2]. Specific datasets of plus-size purchases are often smaller and fragmented which makes it difficult for accuracy analysis and rule generation. Second, customers who buy plus-size clothing have different needs and preferences of the cloth size, materials, and style because they do not have to choose the size that fits them. Therefore, results from traditional recommendation algorithms may vary when implemented in this domain.

A recommendation system tailored to the specific preferences and needs of plus-size individuals becomes crucial for promoting a positive shopping experience and enhancing customer satisfaction.

This study addresses these challenges by applying the association rule mining (ARM) with the Frequent Pattern Growth (FP-Growth) technique to uncover hidden patterns and frequent itemset within plus-size clothing purchases.

2. Literature Review

The association rule mining technique (ARM) is used to detect customer behavior based on transaction data and generate frequent itemset that customers bought together [3]. This allows businesses to understand patterns in customer behavior and focus their marketing efforts on the right products, such as through cross-selling and upselling campaigns [4, 5].

The FP-Growth algorithm is developed from the Apriori algorithm [6, 7]. It is used to determine the set of data that most frequently appears in a data set [6, 8]. Unlike the Apriori algorithm, FP-Growth does not need to generate candidates to get a frequent itemset because it utilizes the concept of tree structure to search for frequent items, so-called FP-tree [9]. This algorithm can extract frequent itemset from the FP-tree. Therefore, it has faster runtime and higher memory consumption stability compared with the Apriori algorithm [7, 9, 10].

FP-Growth has some advantages in this context. First, it requires only one scan of the initial dataset, making it efficient for large datasets. This is particularly beneficial for plus-sized clothing data, as it might be smaller and fragmented compared to standard-sized data [11].

The key measurements of ARM are support Eq. (2.1), confidence Eq. (2.2), and lift values of rules Eq. (2.3) [11].

$$Supp(X \rightarrow Y) = \frac{Supp(X \cup Y)}{N}, \tag{2.1}$$

$$Conf(X \rightarrow Y) = \frac{Supp(X \cup Y)}{Supp(X)}, \tag{2.2}$$

$$Lift(X \rightarrow Y) = \frac{Supp(X \cup Y)}{Supp(X) \times Supp(Y)} \tag{2.3}$$

Support of a rule is the percentage of all transactions (N) which contains the itemset X and Y. Confidence values indicate how often the consequent part of the rule occurs when the antecedent part is also true. Lift value measures how many times more likely the itemset X and Y occur together, with a lift value of 1 indicating no relationship, and values greater than 1 suggesting dependence [11].

The model quality is evaluated using precision Eq. (2.4), recall Eq. (2.5), and F1-score Eq. (2.6) by selecting the top "n" rules from the training set and evaluating the test set [12, 13].

$$Precision = \frac{TP}{TP + FP}, \tag{2.4}$$

$$Recall = \frac{TP}{TP + FN}, \tag{2.5}$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{2.6}$$

Precision assesses the proportion of recommended items that are relevant or interesting to the user, focusing on the quality of the recommendations. Recall, on the other hand, measures the ability of the model to retrieve all the relevant items from the database and recommend them to the user, ensuring the completeness of the recommendations. The F-measure provides a single metric that combines both precision and recall, offering a balance between the quality and completeness of the recommendations. It is calculated as the harmonic mean of precision and recall, thus, giving equal importance to both. In summary, precision, recall, and F1-score are critical measures for evaluating the effectiveness of a recommendation model in providing relevant and useful recommendations to the user.

FP-Growth has been implemented in several applications in different industries [6-10, 14-16]. For clothing recommendation, ARM has been applied based on the transaction data [17-19]. However, the implementation of ARM in a plus-size clothing recommendation system is scarce. Previous studies have focused mainly on the standard-sized clothing [19]. Guo, et al. [20] applied the Apriori algorithm to an online dress shop from the Taobao e-Commerce platform.

There is room for developing effective clothing recommender systems, which can impact customers' shopping experiences and increase sales and revenues for sellers [21]. There is a gap in the literature regarding the recommendation system for plus-size clothing. This research aims to fill this gap by investigating the performance of ARM techniques applied specifically to the plus-size clothing dataset.

3. Methodology

The research was conducted following the predetermined research stages as shown in Fig. 1.

3.1 Data collection

The dataset was derived from the online e-commerce channel of the "CurveG.bkk" shop, a women's plus-size clothing brand in Bangkok. The dataset comprises 54 columns and 26,993 rows, covering the period from January 1, 2021, to December 31, 2022.

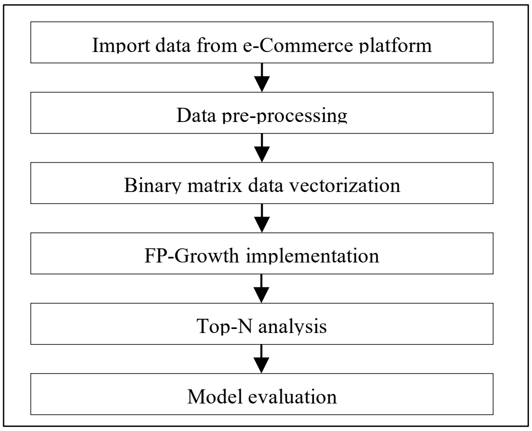

Fig. 1. Research stages.

3.2 Data Preparation

The dataset was cleaned and transformed for further analysis. We selected relevant data fields for plus-size clothing, including order status, username, order date, product name, and quantity. The irrelevant or sensitive fields such as customer personal information were excluded to ensure data privacy. Only the transactions with successfully delivered status were used in the model. With the limited amount of data, the size and color of clothes are excluded to reduce the diversification of items. Table 1 shows an example of cleaned data.

Table 2 shows the results of a binary matrix data vectorization that transformed dataset such that each row represented a transaction, and each column represented a product.

Table 1. Example of sales transaction dataset.

Transaction Date	Invoice No	Product Name	Order Status	User ID	Qty
2022-05-01	0001	Camisole	Completed	ID001	1
2022-05-01	0002	Camisole	Completed	ID002	1
2022-05-01	0003	Crop T-Shirt	Completed	ID003	1
2022-05-01	0004	Wide Leg Pant	Completed	ID003	1
2022-05-01	0005	Camisole	Completed	ID003	1

A value of 1 indicated that the product was purchased, otherwise, it was 0.

Table 2. Example of binary matrix data vectorization results.

Transaction ID	Camisole	Crop T-Shirt	Wide Leg Pant
1	1	0	0
2	1	1	0
3	0	0	1

3.3 FP-growth analysis

The FP-Growth algorithm consists of five processes, namely searching for support item values, building FP-Tree, searching for frequent itemset, searching for support and confidence values on itemset, and obtaining the results of association rules [6]. To avoid the problem of low accuracy due to the limited number of transactions, this study also applied the Top-N analysis to select the top rules ordered by the lift value. Since some customers may not purchase the exact item that the system recommends, they might purchase one in top 5 or top 10 items that the system recommends instead.

4. Results and Discussion

After implementing the FP-Growth algorithm using Python, the example of FP-Tree results is shown in Fig. 2.

We conducted the model-tuning process. We optimized the model's performance through a grid search. The dataset was split into a training set (January 2021 to August 2021) and a test set (September 2021 to April 2022). Using one-hot encoding, the training set was preprocessed and frequent itemsets were generated using the FP-Growth algorithm. During the grid search, we explored different values for the minimum support and minimum confidence thresholds. For example, the minimum support ranges from 0.002 to 0.02, while the minimum confidence ranges from 0.5 to 1 and the top-N ranges from 5 to 15. For each combination of parameter settings, we mined association rules on the training data.

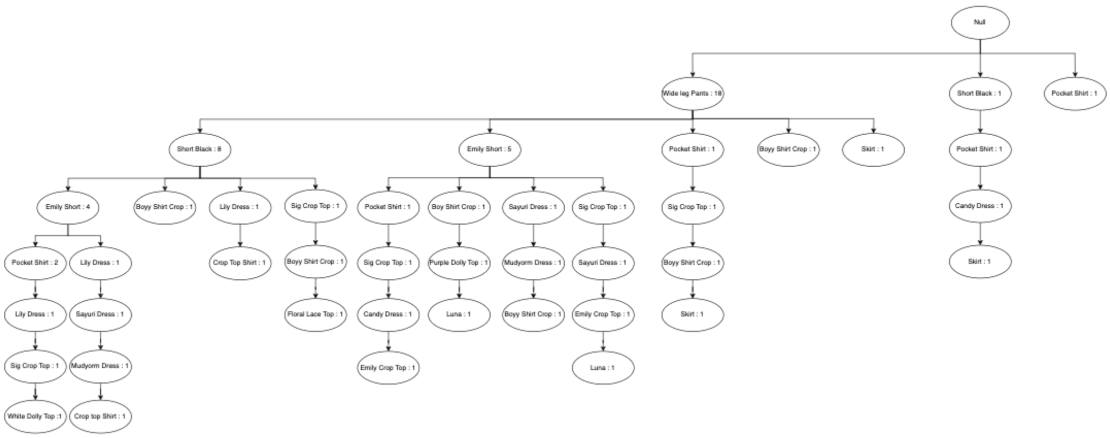

Fig. 2. FP-Tree results.

These rules were then used to generate recommendations for the users in the test set. Using these parameter settings to generate a recommended product for each user and then compare the actual product in the test set. Performance evaluation metrics, including precision, recall, and F1-score, assess the accuracy and effectiveness of the recommendations.

Table 3 shows the results of the model tuning. It included the values of the minimum support and minimum confidence parameters, as well as the corresponding recall, precision, and F1-score metrics.

Table 3. Model performance.

Top N	Min Support	Min Confidence	Precision	Recall	F1-score
1	0.002	0.5	12.49%	7.66%	9.50%
2	0.002	0.5	12.57%	8.12%	9.87%
3	0.002	0.5	12.60%	8.49%	10.14%
4	0.002	0.5	12.51%	8.55%	10.16%
5	0.002	0.5	12.48%	8.82%	10.34%
10	0.002	0.5	12.53%	8.83%	10.36%
15	0.002	0.5	12.51%	8.82%	10.34%
1	0.002	0.6	12.22%	7.48%	9.28%
2	0.002	0.6	12.50%	7.82%	9.62%
3	0.002	0.6	12.57%	8.00%	9.78%
4	0.002	0.6	12.39%	8.00%	9.72%
5	0.002	0.6	12.43%	8.14%	9.84%
10	0.002	0.6	12.42%	8.20%	9.88%
15	0.002	0.6	12.40%	8.20%	9.87%
1	0.002	0.7	9.87%	6.24%	7.65%
2	0.002	0.7	9.83%	6.62%	7.91%
3	0.002	0.7	9.90%	6.76%	8.04%
4	0.002	0.7	9.87%	6.81%	8.06%
5	0.002	0.7	9.87%	6.87%	8.10%
10	0.002	0.7	9.85%	6.87%	8.09%
15	0.002	0.7	9.84%	6.87%	8.09%

We found that increasing the minimum confidence threshold from 0.5 to 1 resulted in a decrease in the performance metrics.

The recall, precision, and F1-score show a downward trend, indicating that higher confidence thresholds may hinder the ability to capture relevant associations in the dataset. Therefore, we set the minimum confidence threshold at 0.5 to ensure a reasonable trade-off between capturing meaningful associations and maintaining satisfactory recommendation performance.

Table 4 displays the F1-scores for various combinations of the minimum support and top-n parameters. The optimal parameter configuration for the association rule mining approach in our research is a minimum support of 0.004 and a top-n value of 5 which provided the highest F1-score of 10.37%. These settings strike a balance between capturing frequent patterns in the dataset and providing a reasonable number of recommendations to users.

Table 4. F1-score for the combination of minimum support, minimum confidence, and Top-N parameters.

Min Support	Min Confidence	Top-N 1	2	3	4	5	10	15
0.002	0.500	9.50%	9.87%	10.14%	10.16%	10.34%	10.36%	10.34%
0.004	0.500	9.64%	9.82%	10.11%	10.16%	**10.37%**	10.37%	10.34%
0.006	0.500	9.33%	9.45%	9.67%	9.71%	9.89%	10.20%	10.20%
0.008	0.500	9.55%	9.50%	9.70%	9.69%	9.93%	10.24%	10.24%
0.010	0.500	9.55%	9.50%	9.70%	9.85%	9.98%	10.24%	10.24%
0.012	0.500	9.55%	9.40%	10.24%	10.24%	10.24%	10.24%	10.24%
0.014	0.500	9.31%	9.15%	9.98%	9.98%	9.98%	9.98%	9.98%
0.016	0.500	9.21%	9.60%	9.60%	9.60%	9.60%	9.60%	9.60%
0.018	0.500	9.19%	9.19%	9.19%	9.19%	9.19%	9.19%	9.19%
0.020	0.500	8.17%	8.17%	8.17%	8.17%	8.17%	8.17%	8.17%

Table 5. Association rules results.

No	Rules	Support	Confidence	Lift
1	{Pocket Shirt} → {Emily Short, Wide Leg Pants}	0.004	1.00	13.17
2	{Pocket Shirt} → {Emily Short}	0.004	1.00	8.51
3	{Pocket Shirt, Wide Leg Pants} → {Emily Short}	0.004	1.00	8.51
4	{Short Black, Wide Leg Pants} → {Emily Short}	0.007	0.83	7.09
5	{Emily Crop Top, Wide Leg Pants} → {Emily Short}	0.006	0.80	6.81
6	{Emily Short, Gigi Set} → {Julie Set}	0.004	0.75	5.86
7	{Camisole} → {Emily Short}	0.004	0.69	5.27
8	{Gigi Set, Wide Leg Pants → {Julie Set}	0.006	0.65	5.08
9	{Emily Shirt} → {Emily Short}	0.006	0.52	3.96
10	{Pocket Shirt, Mudy Top} → {Camisole}	0.004	0.75	3.35
11	{Camisole, Mudy Top} → {Pocket Shirt}	0.004	0.64	2.94
12	{Sunday Top} → {Camisole}	0.020	0.64	2.87
13	{Paris Dress, Wide Leg Pants} → {Camisole}	0.005	0.63	2.80
14	{Sunday Top, Wide Leg Pants} → {Camisole}	0.006	0.54	2.42

Table 5 demonstrates the association rules obtained from the analysis. In total, 35 rules were generated using a minimum support of 0.004, a minimum confidence of 0.5, and a lift greater than 1. These rules highlight interesting associations and patterns in customers' purchasing behavior.

For example, one notable rule reveals a strong association between "Emily Short, Gigi Set" and "Julie Set" with a support of 0.004, a confidence of 0.75, and a lift of 5.86. This indicated that customers who purchased the "Emily Short" and "Gigi Set" items, were highly likely to purchase the "Julie Set" as well. Another interesting rule involves "Wide Leg Pants" and "Emily Short" with a support of 0.004, a confidence of 0.69, and a lift of 5.27. This suggested that customers who purchase the "Emily Short Beige" item are also likely to purchase the "Emily Short" item.

The proposed model provides a precision value of 7.88%, a recall value of 4.23%, and F1-score of 5.50%, which are relatively low due to the limited number of transactions. Moreover, about 65.33% of the customers purchased only 1 item per transaction, 18.42% of them purchased 2 items per transaction, and the rest purchased at least 3 items. These association rules, along with others, shed light on customers' purchasing patterns and present valuable opportunities to enhance recommendation systems specifically for plus-size clothing.

For the implementation stage, we investigated the purchase history of each customer and listed the recommended item obtained from the association rules.

Table 6 shows an example of product recommendations. For example, in the case of Customer ID001, based on her purchase history, we recommend Wide Leg Pants, Short Black, and Emily Short for her. The actual purchases were Emily Short and Wide Leg Pants, which were listed in the recommended items. Customer ID002, we recommend Emily Short, Short Black, Wide Leg Pants. Her actual purchase was Pocket Shirt, Wide Leg Pants. The results showed that customers purchased at least one item in the recommended list obtained from the FP-Growth.

We also applied the Apriori algorithm to this dataset. It provided the same set of rules. Therefore, we compared the computation time (in seconds) as the data size increased to observe the model performance between the FP-Growth and Apriori techniques. Fig. 3 shows that the FP-Growth algorithm had a lower average computation time as the data size increased.

Table 6. Results of recommended items and actual purchased item.

Customer ID	Recommended Items	Actual Purchase Items
ID001	Wide Leg Pants, Short Black, Emily Short	Emily Short, Wide Leg Pants
ID002	Emily Short, Short Black, Wide Leg Pants	Pocket Shirt, Wide Leg Pants
ID003	Wide Leg Pants	Short Black, Pocket Shirt, Wide Leg Pants
ID004	Emily Short, Wide Leg Pants, Mudyorm Dress, Emily Short	Wide Leg Pants
ID005	Emily Short, Short Black, Wide Leg Pants	Emily Short, Candy Dress

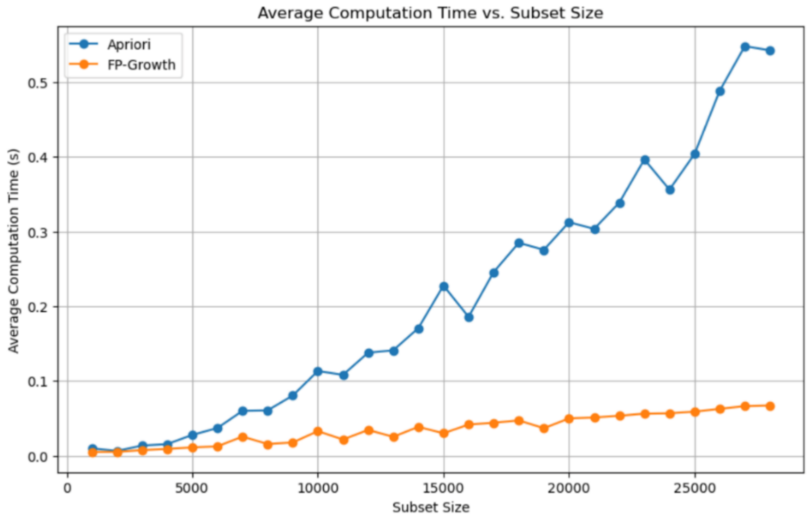

Fig. 3. Average computation time comparison between FP-Growth and Apriori techniques.

5. Conclusion

In this study, a recommendation system in the context of plus-size clothing online transactions using the FP-Growth algorithm has been proposed. The goal is to find frequent itemset for the plus-size clothing to generate the recommended list for each customer.

The model generated a reliable association rule that a user can use to recommend to customers by referring to the purchase history. However, the values of precision, recall, and F1-score of the model are relatively low because of the diverse purchasing behavior of the plus-size customers and most of the customers purchased only 1-2 items per transaction. These are the challenges of developing a recommendation system in the context of plus-size clothing.

Future research should investigate a broader range of recommendation algorithms, such as collaborative filtering or support vector machine, incorporating additional variables to improve the performance of the plus-size clothing model. However, the results may differ in different countries or even cities due to the diversity of customer purchasing behaviors.

Acknowledgements

We would like to thank anonymous reviewers for their constructive feedback and valuable suggestions.

References

[1] S. Verghese. "Plus Size Clothing Market Outlook (2023 to 2033)." https://www.futuremarketinsights.com/reports/plus-size-clothing-market (accessed 7 July, 2022).

[2] A. J. Büttner and M. A. Linardi, "Female Plus Size Consumer Challenge," ed, 2019.

[3] C. Zhang and S. Zhang, Association rule mining: models and algorithms. Springer, 2002.

[4] O. Dogan, F. C. Kem, and B. Oztaysi, "Fuzzy association rule mining approach to identify e-Commerce product association considering sales amount," Complex & Intelligent Systems, 2022;8(2):1551-60.

[5] A. W. Mbugua and A. O. Omondi, "An Application of association rule learning in recommender systems for e-Commerce and its effect on marketing," 2017.

[6] K. Malik, N. Raheja, and P. Garg, "Enhanced FP-growth algorithm," International Journal of Computational Engineering and Management, 2011;12:54-56.

[7] D. Hunyadi, "Performance comparison of Apriori and FP-Growth algorithms in generating association rules," in Proceedings of the 5th European conference on European computing conference, 2011, pp. 376-81.

[8] H. Li, Y. Wang, D. Zhang, M. Zhang, and E. Y. Chang, "PFP: parallel FP-Growth for query recommendation," in Proceedings of the 2008 ACM conference on Recommender systems, 2008, pp. 107-14.

[9] A. M. Said, P. Dominic, and A. B. Abdullah, "A comparative study of FP-Growth variations," International Journal of Computer Science and Network Security, 2009;9(5):266-72.

[10] R. Ranjan and A. Sharma, "Evaluation of frequent itemset mining platforms using Apriori and FP-growth algorithm," arXiv preprint arXiv:1902.10999, 2019.

[11] R. Agrawal, T. Imielinski, and A. Swami, "Mining associations between sets of items in large databases," in Proceedings of the ACM SIGMOD International Conference on Management of Data, 2019, pp. 207-16.

[12] Z. Khanzadeh and M. Mahdavi, "Utilizing association rules for improving the performance of collaborative filtering," International Journal of E-Entrepreneurship and Innovation (IJEEI), 2012;3(2):14-28.

[13] Y. K. Choi and S. K. Kim, "A Recommendation System for Repetitively Purchasing Items in E-commerce Based on Collaborative Filtering and Association Rules," Journal of Internet Technology, 2018;19(6):1691-8.

[14] J. Heaton, "Comparing dataset characteristics that favor the Apriori, Eclat or FP-Growth frequent itemset mining algorithms," in SoutheastCon 2016, 2016: IEEE, pp. 1-7.

[15] M. R. Pradana, M. Syafrullah, H. Irawan, J. C. Chandra, and A. Solichin, "Market basket analysis using FP-Growth algorithm on retail sales data," in 2022 9th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), 2022: IEEE, pp. 86-89.

[16] M. Wojciechowski, K. Galecki, and K. Gawronek, "Concurrent processing of frequent itemset queries using FP-Growth algorithm," in Proceedings of the 1st ADBIS Workshop on Data Mining and Knowledge Discovery. Publishing House of Poznan University of Technology, 2005, pp. 35-46.

[17] C. Guan, S. Qin, W. Ling, and G. Ding, "Apparel recommendation system evolution: an empirical review," International Journal of Clothing Science and Technology, 2016;28(6):854-79.

[18] 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, 2021;21(12):4239.

[19] Y. Wakita, K. Oku, H.-H. Huang, and K. Kawagoe, "A fashion-brand recommender system using brand association rules and features," presented at the 2015 IIAI 4th International Congress on Advanced Applied Informatics, 2015.

[20] Y. Guo, M. Wang, and X. Li, "Application of an improved Apriori algorithm in a mobile e-Commerce recommendation system," Industrial Management & Data Systems, 2017;117(2):287-303.

[21] Y. Deldjoo et al., "A review of modern fashion recommender systems," ACM Computing Surveys, 2023;56(4):1-37.