Machine learning based recommender system for e-commerce

Manal Loukili[1], Fayçal Messaoudi[2], Mohammed El Ghazi[3]
[1]Artificial Intelligence, Data Science and Emerging Systems laboratory, National School of Applied Sciences, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[2]Artificial Intelligence, Data Science and Emerging Systems laboratory, National School of Business and Management, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[3]Artificial Intelligence, Data Science and Emerging Systems laboratory, Superior School of Technology, Sidi Mohamed Ben Abdellah University, Fez, Morocco

Article Info

Article history:

Received Nov 24, 2022
Revised Jan 11, 2023
Accepted Mar 10, 2023

Keywords:

Association rules
E-commerce
Frequent-pattern-growth algorithm
Machine learning
Recommender system

ABSTRACT

Nowadays, e-commerce is becoming an essential part of business for many reasons, including the simplicity, availability, richness and diversity of products and services, flexibility of payment methods and the convenience of shopping remotely without losing time. These benefits have greatly optimized the lives of users, especially with the technological development of mobile devices and the availability of the Internet anytime and anywhere. Because of their direct impact on the revenue of e-commerce companies, recommender systems are considered a must in this field. Recommender systems detect items that match the customer's needs based on the customer's previous actions and make them appear in an interesting way. Such a customized experience helps to increase customer engagement and purchase rates as the suggested items are tailored to the customer's interests. Therefore, perfecting recommendation systems that allow for more personalized and accurate item recommendations is a major challenge in the e-marketing world. In our study, we succeeded in developing an algorithm to suggest personal recommendations to customers using association rules via the Frequent-Pattern-Growth algorithm. Our technique generated good results with a high average probability of purchasing the next product suggested by the recommendation system.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Manal Loukili
Artificial Intelligence, Data Science and Emerging Systems laboratory, National School of Applied Sciences, Sidi Mohamed Ben Abdellah University
Fez, Morocco
Email: manal.loukili@usmba.ac.ma

1. INTRODUCTION

Today, with the development of digital technologies and the ease of access to the internet, we are more and more exposed to a wealth of information [1]. This development brings a lot of diversity to users, but this multitude of information sources and information overload can become problematic [2]. In order to remedy this problem, various tools have been developed to filter information before transmitting it to users. Recommendation systems are tools that allow such filtering [3]. The main goal of these systems is to facilitate decision making for users by offering them information according to their preferences [4], [5].

Generally, there are two main categories of recommendation systems. One is done with content-based filtering and the other is done with collaborative filtering [6]. The content-based approach is implemented by taking into account the characteristics of the recommended products and creating groups of products using a similarity measure on their content [7]. Content implies a relationship between a word and a product [8]. One of the main drawbacks of this approach is that the features are usually acquired using external information that is not always available. On another note, collaborative filtering typically uses a neighborhood of similar users and recommended products based on the history of other users within the same neighborhood [9]. When systems issue their recommendations, they can do so globally or locally [10]. "Globally" means that all users will receive the same recommendations. While "locally" means that the recommended items will not be the same for all users.

Two approaches exist in the literature when developing a recommendation system using collaborative filtering [11]. The first approach is memory-based. That is, the system will issue its recommendations based on a neighborhood contained in memory. Whereas the second approach is rather model based [12]. Thus, it is necessary to first create models that resemble the users' behaviors and then use these models to issue recommendations. In practice, it has been shown that the memory-based approach offers better performance in terms of accuracy while the model-based approach is more efficient at large scale with large data sets.

There are mainly two types of information that are used by recommender systems, implicit and explicit information [13]. Explicit information is translated as information provided by the user. It is possible to provide this information through the use of ratings, wish lists, and comments. Unlike explicit information, implicit information uses only information for which the user has not been asked. For example, when a user buys a product, it creates a purchase history for that user. It is possible for implicit information recommendation systems to use this information to create different user profiles.

This paper is structured. Section 2 is devoted to some related work. Section 3 presents the background of recommender systems. Section 4 describes our methodology. Section 5 present the results obtained. Section 6 concludes the paper.

2. RELATED WORK

In this section, some recent papers on recommender systems in e-commerce are cited. Kuo and Cheng [14] proposed a personalized content-based recommender system integrating the architecture of the existing traditional content-based recommender system for e-commerce with the addition of a feedback adjuster. The results showed that the proposed system which is based on a more objective approach can determine customer preferences based on their repeated purchase behavior, thus avoiding subjective judgments. Huang [15] conducted a comparative analysis of three popular recommendation algorithms on a data set of the Alibaba website. The experimental results revealed that the convolutional neural network recommendation algorithm based on deep learning outperformed the other algorithms. In another study [16], Chehal *et al.* proposed an approach to improve recommendations by taking advantage of the product evaluation feature. This method involves extracting product features from user-provided feedback and not recommending products with irrelevant features in order to improve the recommendation list. In [17], a comparative study was conducted of various collaborative filtering algorithms to evaluate their performance, namely: K-Nearest Neighbor, Slope One, co-clustering and Non-negative Matrix Factorization. The results showed that the K-Nearest Neighbor algorithm for item-based collaborative filtering outperformed the other models. In [18], Xuecong *et al.* implemented a solution to build an e-commerce recommendation system on the cloud computing platform to manage massive commodity information and user information more efficiently. The proposed solution enables them to increase the capability of massive data mining and business intelligence analysis and achieve high-performance computing at reduced cost. The authors suggested in [19] an approach of collecting and pre-processing real-time data from multiple e-commerce platforms, which generate user data, and gathering all personalized user data to prepare for the next data mining, and then using the data mining technology in big data to automatically recommend personalized products to satisfy the personalized needs and tastes of customers. In [20], in order to identify customer preferences, the authors proposed an improved conjoint analysis method. To evaluate their method, they compared it with other prediction algorithms namely: generalized linear model, decision tree, random forest, gradient boosted trees, and support vector machine. The experimental results showed that this joint analysis network performs well and works good for choosing a product among several attributes to get the best price despite its limitations.

3. BACKGROUND

There are various types of approaches to realize a recommender system. In this section, the approaches discussed apply to both explicit and implicit information recommender systems. In general, a recommender system is a means of filtering data of user's "U", a set of products "P" and a utility function "h" that indicates the level of interest of a user in a given product. Consequently, recommender systems can be considered as a function to be maximized, which can be written in the following form (1).

$$\forall u \in U, i'_u = arg_{i \in I} \max h(u, i) \tag{1}$$

Where "I" refers to a group of products for "U", and the utility function "h" is particular to the employed type of approach. As well, the utility function "h" depends on the type of information used, either implicit or explicit. The group of products "I" plays a crucial role in the formulation of a recommender system, as it represents the set of items that can be recommended to the users. The size and composition of this group can have a significant impact on the accuracy and relevance of the recommendations generated by the system. Additionally, the choice of the utility function "h" is specific to the employed approach and is influenced by the type of information used, whether implicit or explicit. For instance, the utility function for a collaborative filtering approach is often based on the similarity between users or items, while the utility function for a content-based approach is based on the attributes or features of the items. The utility function should be chosen carefully to ensure that the recommendations generated by the system align with the user's preferences and needs.

3.1. Collaborative filtering

Collaborative filtering is the process of using the collaboration of users, to make a recommendation [21]. The way collaborative filtering works is as described,

- First, it gathers information like reviews or purchase behavior from users to determine their interests.
- Next, it checks the collected information against other users and locates the strongest similarities.
- Lastly, the system suggests products to the similar target users.

Collaborative filtering can be done either by keeping all information about similar users in memory, or by creating a preference model from this information [22]. With a system where information is kept in memory, the two most popular similarity measures are the cosine measure (2) and the Pearson correlation (3). The Pearson correlation is useful when the recommendation system uses explicit information while the cosine measure is used when there is implicit information such as a purchase history [23]. When collaborative filtering is used to create a model, other techniques such as segmentation [24] are used to discover latent information that could explain the purchase of a product or the outcome of a quote.

$$\cos(\vec{x}, \vec{y}) = \frac{\vec{x} . \vec{y}}{||\vec{x}|| * ||\vec{y}||} \tag{2}$$

$$PC(x, y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n * stdev(x) * stdev(y)} \tag{3}$$

3.2. Trusted network

Trust networks use similar principles to collaborative filtering but take the similarity between users a step further [25]. Research on trust networks is motivated by the hypothesis that there are other factors, besides the traditional similarity measures used in collaborative filtering, that can lead a user to purchase a product [26]. This hypothesis is based on the idea that in everyday life, people sometimes ask for advice from reputable people, and discuss a product with their friends. In contrast to collaborative filtering where similarity is based on the profiles of individuals, the main interest of trust networks is the trust relationship that links users together.

The main advantage of adding a network of trust to a system using collaborative filtering is that it allows the cold start problem to be relaxed [27]. With the network of trust, even if the user is new to the system, he can still receive recommendations if he is connected to other individuals. Obviously, this requires that the individual knows other users who are part of the system. Also, when a product is not found in the user's immediate network, the search can move forward with a second level of trust and so on. However, even though this provides very high coverage, as the search moves further from the source, the accuracy decreases.

There are two broad categories of trust: context-specific interpersonal trust and impersonal trust. The former indicates trust where a user trusts those around him/her only for a specific context and may trust a completely different group for a different context. For example, a user might trust his or her entourage to go see a movie but might also trust a real estate agent to make a house purchase. The second indicates a systematic trust that describes a user's trust as a whole within a group. For example, two individuals trust each other because they are in the same group.

3.3. Latent model

Unlike approaches that use a trust network or collaborative filtering, recommender systems that use latent models will not hold user information in memory during recommendation [28]. Instead, they will use this information to create models using data mining techniques. In order to create a model, the system first trains its model with random parameters and then refines the parameters with one or more cost functions. Unlike the in-memory approach, the latent model approach will not necessarily attempt to find direct links between users and products. Instead, they will try to discover latent links that will motivate the user to purchase a product [29]. Following this idea, it is then possible to capture facts that could motivate some users to prefer certain products without it being obvious to a traditional system. However, it is difficult with this approach to explain the reasoning behind a recommendation, as the guidelines do not have properly identified labels. That is, a latent factor may be a mixture of different regular factors.

Latent model development in the area of recommender systems often employs the matrix factorization technique [30]. This technique allows users and products to be associated in a new space where the relationships between users and products can be defined by using the dot product. The matrix in the initial space is usually formed of users for the rows and products for the columns. Thus, to compute an estimated score for a product, one must take the user vector and the product vector in the latent space and then compute the dot product between these two vectors. However, since data sets for recommender systems are often scattered, it is important to be careful not to add too much noise to the new space.

3.4. Association rules

Association rules are typically used to discover repeating patterns in data sets that can be very large [31]. Like trust networks, association rules are often combined with a collaborative filtering approach to make the latter even more accurate. In general, association rules follow a probabilistic cause and effect model. That is, the system analyzes a multitude of events and then groups the instances that precede a fact.

Association rule-based recommender systems typically use a transaction history to extract the rules that will allow systems to make associations [32]. However, these methods often lead to a very large number of rules, which makes the knowledge obtained from these rules inefficient. In addition, association rule methods require explicit parameters such as the number of rules and the confidence level associated with these rules, in previous work, a recommendation system was created based on association rules, which are produced using a model that performs quality control on them from the moment of creation. Basically, there are two types of associations: associations between products and associations between users. For example, an association between products may be that users who like product A and product B also like product C. As for the association between users, this could be that items that are liked by user A and user B are also liked by user C. This practice allows us to find inferences between rules and analyzing the overlap between users allows us to make recommendations to users who do not have much similarity directly between them.

Although using a model to create association rules increases the quality of the association rules, the problem remains that the number of rules produced is still high. Therefore, the noise caused by a very large number of rules reduces the confidence in the recommendations significantly. An association rule can be defined as a truth table that results from the combination of two or more features [33]. Association rules are a set of "if-then" statements designed to provide probability of relationships among data features [34], in large data sets of various types of databases. Association rule mining makes a variety of use cases and is widely used to help discover sales correlations in transactional data sets. Association rules are based on the following three parameters [35]:

- Support: is a measure of the percentage of items, whether a single item or a combination of items, that appear in the data set. The support value is a measure of how frequently the item set appears in an association.
- Confidence: is a measure of the probability's accuracy of an item's occurrence relative to another item. The confidence value is between zero and one, where the closer the confidence value is to one, the more probable the two items are to occur together.
- Lift: is the ratio of the support observed compared to the support that would be expected if X and Y were independent. The lift shows the value of the rule based on random events X and Y.

Association rules are useful for analyzing and predicting customer behavior [36], [37]. They are particularly important in customer analysis, shopping cart analysis, product and customer clustering, catalog design and store management [38]–[40]. Algorithms that use association rules include AIS (Artificial Immune Systems), SETM (Self-Exciting Temporal Point Process Model), and Apriori [41].

- Apriori algorithm

The Apriori algorithm uses frequent item-sets to build association rules [42]. The algorithm is developed to work on transactional data sets [43]. Based on the association rules, it can assess the strength or weakness of the connection between two objects. Through a breadth-first search and a hash tree, the algorithm can successfully compute the itemset associations. The process is iterative in order to find frequent sets from a large data set.

- FP-Growth algorithm

The frequent pattern growth (FP-Growth) algorithm is an improved version of the Apriori algorithm [44]. It is an alternative method for finding frequent item-sets without using candidate generations [45], by employing a divide-and-conquer strategy. This efficient and scalable approach relies on the use of an extended prefix-tree structure to store compressed and crucial frequent pattern information, called a frequent pattern tree (FP-tree), which preserves the association information of item-sets [46].

4. METHODOLOGY

To overcome the problem of customer decision making and to increase sales, the implementation of an accurate and efficient product recommendation system is crucial [47]. In this paper, an association rule-based recommendation system, via the FP-Growth algorithm, is adopted because it provides high accuracy while being easy to implement and explain. Thus, the results of this model are understandable: all the rules specific to the activity can be read in the association table. The Figure 1 shows the steps of our methodology.

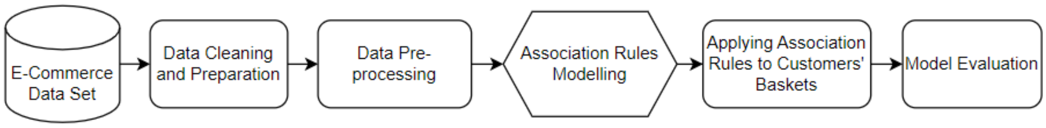

Figure 1. The proposed recommender system workflow

4.1. Data visualization

The data set used is "Online Retail" from the Machine Learning Repository. This is a transactional data set that contains 406,829 transactions for a UK-based registered e-commerce company that sells unique gifts. The dataset includes information on customer identification, invoice number, stock code, product description, quantity, invoice date, unit price, total invoice price, and country [48]. Figure 2 shows the first 10 rows of the data set used.

(406829, 9)

	InvoiceNo	StockCode	Description	Quantity	InvoiceDate	UnitPrice	CustomerID	Country	TotalPrice
0	536365	85123A	WHITE HANGING HEART T-LIGHT HOLDER	6	12/1/2010 8:26	2.55	17850.0	United Kingdom	15.30
1	536365	71053	WHITE METAL LANTERN	6	12/1/2010 8:26	3.39	17850.0	United Kingdom	20.34
2	536365	84406B	CREAM CUPID HEARTS COAT HANGER	8	12/1/2010 8:26	2.75	17850.0	United Kingdom	22.00
3	536365	84029G	KNITTED UNION FLAG HOT WATER BOTTLE	6	12/1/2010 8:26	3.39	17850.0	United Kingdom	20.34
4	536365	84029E	RED WOOLLY HOTTIE WHITE HEART.	6	12/1/2010 8:26	3.39	17850.0	United Kingdom	20.34
5	536365	22752	SET 7 BABUSHKA NESTING BOXES	2	12/1/2010 8:26	7.65	17850.0	United Kingdom	15.30
6	536365	21730	GLASS STAR FROSTED T-LIGHT HOLDER	6	12/1/2010 8:26	4.25	17850.0	United Kingdom	25.50
7	536366	22633	HAND WARMER UNION JACK	6	12/1/2010 8:28	1.85	17850.0	United Kingdom	11.10
8	536366	22632	HAND WARMER RED POLKA DOT	6	12/1/2010 8:28	1.85	17850.0	United Kingdom	11.10
9	536367	84879	ASSORTED COLOUR BIRD ORNAMENT	32	12/1/2010 8:34	1.69	13047.0	United Kingdom	54.08

Figure 2. The data set used

4.2. Data pre-processing

First, we kept only the products that the customer has actually put in his basket by deleting the products that correspond to the company's gifts to customers. Then we grouped all the products that a customer has purchased in a new data set. Each line corresponds to a transaction including the invoice number, the customer ID and all the products purchased as shown in Figure 3.

		StockCode
InvoiceNo	CustomerID	
536365	17850.0	[84406B, 85123A, 21730, 22752, 84029G, 84029E,...
536366	17850.0	[22632, 22633]
536367	13047.0	[21777, 84969, 22748, 22749, 84879, 22623, 226...
536368	13047.0	[22914, 22912, 22913, 22960]
536369	13047.0	[21756]

Figure 3. The data set of all products purchased by each customer

4.3. Association Rules modelling

For the determination of the association rules, we used the FP-Growth model which allows, from a history of transactions, to determine the set of the most frequent association rules in the data set. To do this, it needs as input parameter the set of transactions composed of the baskets of products that the customers have already purchased. To establish the table of associations, it is necessary to define 2 hyperparameters:

– minSupRatio: It represents the minimum support for an item set to be identified as frequent.
– minConf: It represents the minimum confidence for generating an association rule.

The system generated 4,970 association rules as shown in Figure 4.

	basket	next_product	p
199	{22916, 22921, 22920, 22917, 22919}	{22918}	0.992537
424	{22916, 22921, 22919, 22917}	{22918}	0.986014
306	{22920, 22921, 22918, 22917}	{22916}	0.985714
310	{22920, 22921, 22916, 22917}	{22918}	0.985714
96	{22920, 22921, 22919, 22917}	{22918}	0.985401
231	{22920, 22916, 22921, 22919}	{22918}	0.985401
202	{22918, 22921, 22920, 22917, 22919}	{22916}	0.985185
201	{22916, 22921, 22918, 22920, 22919}	{22917}	0.985185
366	{22921, 22916, 22917}	{22918}	0.979866
458	{22921, 22919, 22917}	{22918}	0.979730

Figure 4. Association rules data set

Once the association rules were established using the FP-Growth model, they were applied to the customers' product baskets. This was done to recommend products to customers based on their past purchase history. Using the recommended products, the estimated prices were then calculated, as demonstrated in Figure 5.

InvoiceNo	CustomerID	Customer basket	Recommended Product	Product description	Probability	Price estimation
536365	17850.0	[85123A, 84406B, 84029G, 84029E, 22752, 21730,...	0	Null	0.000000	0.000000
536366	17850.0	[22633, 22632]	22865	HAND WARMER OWL DESIGN	0.470982	0.871317
536367	13047.0	[22310, 84969, 48187, 21754, 22623, 22749, 217...	22750	FELTCRAFT PRINCESS LOLA DOLL	0.593516	2.225686
536368	13047.0	[22912, 22960, 22913, 22914]	22961	JAM MAKING SET PRINTED	0.322280	0.467306
536369	13047.0	[21756]	21754	HOME BUILDING BLOCK WORD	0.576132	3.427984
536370	12583.0	[22631, 22629, 22726, 22661, 21913, 21724, 218...	22630	DOLLY GIRL LUNCH BOX	0.595506	1.161236
536371	13748.0	[22086]	22910	PAPER CHAIN KIT VINTAGE CHRISTMAS	0.458586	1.169394
536372	17850.0	[22633, 22632]	22865	HAND WARMER OWL DESIGN	0.470982	0.871317
536373	17850.0	[85123A, 84406B, 20679, 21871, 37370, 82482, 2...	15056BL	EDWARDIAN PARASOL BLACK	0.509158	3.029487

Figure 5. Association rules data set applied to customers' baskets

5. MODEL EVALUATION AND RESULTS

The evaluation of a recommendation system is a crucial step in determining its effectiveness in improving sales and customer satisfaction. In this study, the performance of the recommendation system was evaluated by calculating the average probability of the next product that the customer will buy ($P_{average}$). In addition, the expected incomes from the predicted and suggested products were also calculated and presented in Table 1. This evaluation allowed for an estimation of the potential revenue increase that could be achieved with the implementation of the proposed recommendation system.

Table 1. The performance of the recommender system

Number of suggested for each customer	$P_{average}$ (in %)	Generated turnover (in $)
1	69.3374553	36,891.891891

6. CONCLUSION

In today's information-intensive age, data overload has become an obstacle to making the right decisions at the right time. Particularly for e-consumers, accurately and quickly determining the information that is of interest to them in this constant chaos is a primary challenge for every e-commerce organization. For this reason, recommendation systems are called upon to solve this problem and help users make the right decisions. In this light, the objective of our study was to implement a recommendation system based on association rules via the FP-Growth algorithm. The limitation of this work is that some of the evaluation characteristics of a recommender system, such as diversity and explainability, are difficult to define. Real-world evaluation, including A/B testing and sample testing, is a powerful technique for evaluating a new recommender system. Indeed, as a next work, we will set up recommendation systems based on different machine learning models to conduct A/B testing and define the best performing model.

ACKNOWLEDGEMENTS

M. Loukili author is very grateful and appreciative for the support and guidance of her thesis supervisors, the members of the Laboratory of Artificial Intelligence, Data Science and Emerging Systems (LIASSE), not to mention all the professors of the University Sidi Mohamed Ben Abdellah of Fez, Morocco. It should be noted that this project was carried out as part of a doctoral thesis and did not receive any external research funding or grants.

REFERENCES

[1] M. Loukili, F. Messaoudi, and M. El Ghazi, "Supervised learning algorithms for predicting customer churn with hyperparameter optimization," *International Journal of Advances in Soft Computing and its Applications*, vol. 14, no. 3, pp. 49–63, 2022, doi: 10.15849/IJASCA.221128.04.

[2] P. M. Alamdari, N. J. Navimipour, M. Hosseinzadeh, A. A. Safaei, and A. Darwesh, "A systematic study on the recommender systems in the e-commerce," *IEEE Access*, vol. 8, pp. 115694–115716, 2020, doi: 10.1109/ACCESS.2020.3002803.

[3] Z. Batmaz, A. Yurekli, A. Bilge, and C. Kaleli, "A review on deep learning for recommender systems: challenges and remedies," *Artificial Intelligence Review*, vol. 52, no. 1, pp. 1–37, 2019, doi: 10.1007/s10462-018-9654-y.

[4] M. Loukili, F. Messaoudi, and M. El Ghazi, "Sentiment analysis of product reviews for e-commerce recommendation based on machine learning," *International Journal of Advances in Soft Computing and its Applications*, vol. 15, no. 1, pp. 1–13, 2023, doi: 10.15849/IJASCA.230320.01.

[5] S. Rawat, U. Tyagi, and S. Singhal, "Recommender systems in e-commerce and their challenges," *Proceedings - 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021*, pp. 1598–1601, 2021, doi: 10.1109/ICAC3N53548.2021.9725681.

[6] S. Salloum and D. Rajamanthri, "Implementation and evaluation of movie recommender systems using collaborative filtering," *Journal of Advances in Information Technology*, vol. 12, no. 3, pp. 189–196, 2021, doi: 10.12720/jait.12.3.189-196.

[7] A. H. Espejel and F. J. Cantu-Ortiz, "Data mining techniques to build a recommender system," *Proceedings - 2021 International Symposium on Computer Science and Intelligent Controls, ISCSIC 2021*, pp. 217–221, 2021, doi: 10.1109/ISCSIC54682.2021.00047.

[8] Q. Ai, V. Azizi, X. Chen, and Y. Zhang, "Learning heterogeneous knowledge base embeddings for explainable recommendation," *Algorithms*, vol. 11, no. 9, 2018, doi: 10.3390/a11090137.

[9] Z. Cui *et al.*, "Personalized recommendation system based on collaborative filtering for IoT scenarios," *IEEE Transactions on Services Computing*, vol. 13, no. 4, pp. 685–695, 2020, doi: 10.1109/TSC.2020.2964552.

[10] "XXX."

[11] G. George and A. M. Lal, "Hy-MOM: Hybrid recommender system framework using memory-based and model-based collaborative filtering framework," *Cybernetics and Information Technologies*, vol. 22, no. 1, pp. 134–150, 2022, doi: 10.2478/cait-2022-0009.

[12] Z. Noshad, A. Bouyer, and M. Noshad, "Mutual information-based recommender system using autoencoder," *Applied Soft Computing*, vol. 109, 2021, doi: 10.1016/j.asoc.2021.107547.

[13] M. Riyahi and M. K. Sohrabi, "Providing effective recommendations in discussion groups using a new hybrid recommender system based on implicit ratings and semantic similarity," *Electronic Commerce Research and Applications*, vol. 40, 2020, doi: 10.1016/j.elerap.2020.100938.

[14] R. J. Kuo and H. R. Cheng, "A content-based recommender system with consideration of repeat purchase behavior," *Applied Soft Computing*, vol. 127, 2022, doi: 10.1016/j.asoc.2022.109361.

[15] G. Huang, "E-commerce intelligent recommendation system based on deep learning," *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers, IPEC 2022*, pp. 1154–1157, 2022, doi: 10.1109/IPEC54454.2022.9777500.

[16] D. Chehal, P. Gupta, and P. Gulati, "An approach to utilize e-commerce product reviews to remove irrelevant recommendations," *2022 IEEE Delhi Section Conference, DELCON 2022*, 2022, doi: 10.1109/DELCON54057.2022.9753277.

[17] M. Al-Ghamdi, H. Elazhary, and A. Mojahed, "Evaluation of collaborative filtering for recommender systems," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 3, pp. 559–564, 2021, doi: 10.14569/IJACSA.2021.0120367.

[18] C. Xuecong, L. Zhaoming, and C. Sisi, "Design and implementation of e-commerce recommendation system model based on cloud computing," *Proceedings of IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers, IPEC 2021*, pp. 1100–1103, 2021, doi: 10.1109/IPEC51340.2021.9421260.

[19] Z. Wang, A. Maalla, and M. Liang, "Research on e-commerce personalized recommendation system based on big data technology," *Proceedings of 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence, ICIBA 2021*, pp. 909–913, 2021, doi: 10.1109/ICIBA52610.2021.9687955.
[20] A. B. Osmond and F. Hidayat, "Electronic commerce product recommendation using enhanced conjoint analysis," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, pp. 666–673, 2021, doi: 10.14569/IJACSA.2021.0121176.
[21] L. Long, F. Huang, Y. Yin, and Y. Xu, "Multi-task learning for collaborative filtering," *International Journal of Machine Learning and Cybernetics*, vol. 13, no. 5, pp. 1355–1368, 2022, doi: 10.1007/s13042-021-01451-0.
[22] Y. Koren and R. Bell, "Advances in collaborative filtering," in *Recommender Systems Handbook*, Boston, MA: Springer US, 2015, pp. 77–118.
[23] M. Ahmadian, M. Ahmadi, and S. Ahmadian, "A reliable deep representation learning to improve trust-aware recommendation systems," *Expert Systems with Applications*, vol. 197, 2022, doi: 10.1016/j.eswa.2022.116697.
[24] B. Y. Pratama, I. Budi, and A. Yuliawati, "Product recommendation in offline retail industry by using collaborative filtering," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 9, pp. 635–643, 2020, doi: 10.14569/IJACSA.2020.0110975.
[25] W. Li, J. Cao, J. Wu, C. Huang, and R. Buyya, "A collaborative filtering recommendation method based on discrete quantum-inspired shuffled frog leaping algorithms in social networks," *Future Generation Computer Systems*, vol. 88, pp. 262–270, 2018, doi: 10.1016/j.future.2018.05.070.
[26] M. Raizada, "Survey on recommender systems incorporating trust," *Proceedings - International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2022*, pp. 1011–1015, 2022, doi: 10.1109/ICAAIC53929.2022.9792731.
[27] Q. Wu *et al.*, "Dual graph attention networks for deep latent representation of multifaceted social effects in recommender systems," *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, pp. 2091–2102, 2019, doi: 10.1145/3308558.3313442.
[28] Z. Khan, N. Iltaf, H. Afzal, and H. Abbas, "Enriching non-negative matrix factorization with contextual embeddings for recommender systems," *Neurocomputing*, vol. 380, pp. 246–258, 2020, doi: 10.1016/j.neucom.2019.09.080.
[29] G. Adomavicius, K. Bauman, A. Tuzhilin, and M. Unger, "Context-aware recommender systems: From foundations to recent developments," *Recommender Systems Handbook*, pp. 211–250, 2022, doi: 10.1007/978-1-0716-2197-4_6.
[30] N. Idrissi and A. Zellou, "A systematic literature review of sparsity issues in recommender systems," *Social Network Analysis and Mining*, vol. 10, no. 1, 2020, doi: 10.1007/s13278-020-0626-2.
[31] Y. Zheng and D. (Xuejun) Wang, "A survey of recommender systems with multi-objective optimization," *Neurocomputing*, vol. 474, pp. 141–153, 2022, doi: 10.1016/j.neucom.2021.11.041.
[32] E. Saha and P. Rathore, "Discovering hidden patterns among medicines prescribed to patients using Association Rule Mining Technique," *International Journal of Healthcare Management*, 2022, doi: 10.1080/20479700.2022.2099335.
[33] I. H. Sarker, "Machine learning: algorithms, real-world applications and research directions," *SN Computer Science*, vol. 2, no. 3, 2021, doi: 10.1007/s42979-021-00592-x.
[34] K. H. T. Dam, T. Given-Wilson, A. Legay, and R. Veroneze, "Packer classification based on association rule mining," *Applied Soft Computing*, vol. 127, 2022, doi: 10.1016/j.asoc.2022.109373.
[35] S. Das, X. Sun, S. Goel, M. Sun, A. Rahman, and A. Dutta, "Flooding related traffic crashes: Findings from association rules," *Journal of Transportation Safety and Security*, vol. 14, no. 1, pp. 111–129, 2022, doi: 10.1080/19439962.2020.1734130.
[36] A. Telikani, A. H. Gandomi, and A. Shahbahrami, "A survey of evolutionary computation for association rule mining," *Information Sciences*, vol. 524, pp. 318–352, 2020, doi: 10.1016/j.ins.2020.02.073.
[37] S. Singh and A. Yassine, "Big data mining of energy time series for behavioral analytics and energy consumption forecasting," *Energies*, vol. 11, no. 2, 2018, doi: 10.3390/en11020452.
[38] Y. A. Ünvan, "Market basket analysis with association rules," *Communications in Statistics - Theory and Methods*, vol. 50, no. 7, pp. 1615–1628, 2021, doi: 10.1080/03610926.2020.1716255.
[39] A. Griva, C. Bardaki, K. Pramatari, and D. Papakiriakopoulos, "Retail business analytics: Customer visit segmentation using market basket data," *Expert Systems with Applications*, vol. 100, pp. 1–16, 2018, doi: 10.1016/j.eswa.2018.01.029.
[40] A. Doniec, S. Lecoeuche, R. Mandiau, and A. Sylvain, "Purchase intention-based agent for customer behaviours," *Information Sciences*, vol. 521, pp. 380–397, 2020, doi: 10.1016/j.ins.2020.02.054.
[41] and F. M. S. Mukherjee, P. Gupta, "Integrating application with algorithms of association rule used in descriptive data modelling, through which data mining can be implemented for future prediction," *International Journal of Applied Engineering Research*, vol. 13, no. 17, pp. 13272–13281, 2018.
[42] N. K. Verma and V. Singh, "A rational approach to improve access time of apriori algorithm by applying inner join in a arm to redefining fis in textual data," *ECS Transactions*, vol. 107, no. 1, pp. 7749–7758, 2022, doi: 10.1149/10701.7749ecst.
[43] H. Xie, "Research and case analysis of apriori algorithm based on mining frequent item-sets," *Open Journal of Social Sciences*, vol. 09, no. 04, pp. 458–468, 2021, doi: 10.4236/jss.2021.94034.
[44] M. Patil and T. Patil, "Apriori algorithm against FP growth algorithm: a comparative study of data mining algorithms," *SSRN Electronic Journal*, 2022, doi: 10.2139/ssrn.4113695.
[45] L. Shabtay, P. Fournier-Viger, R. Yaari, and I. Dattner, "A guided FP-growth algorithm for mining multitude-targeted item-sets and class association rules in imbalanced data," *Information Sciences*, vol. 553, pp. 353–375, 2021, doi: 10.1016/j.ins.2020.10.020.
[46] A. Javed and A. Khokhar, "Frequent pattern mining on message passing multiprocessor systems," *Distributed and Parallel Databases*, vol. 16, no. 3, pp. 321–334, 2004, doi: 10.1023/B:DAPD.0000031634.19130.bd.
[47] M. L. and F. Messaoudi, "Machine learning, deep neural network and natural language processing based recommendation system," *2022 International Conference on Advanced Intelligent Systems for Sustainable Development, Lecture Notes in Networks and Systems AI2SD*, 2023.
[48] Daqing Chen, "Online retail data set," *UCI Machine Learning Repository*, 2012, [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Online+Retail.

BIOGRAPHIES OF AUTHORS

Dr. Eng. Manal Loukili is an IT engineer and PhD student at the University of Sidi Mohamed Ben Abdellah in Fez, Morocco. She is a member of the laboratory: Artificial Intelligence, Data Science and Emerging Systems (IASSE). Her principal research areas are Machine Learning and E-Marketing. She can be contacted at email: manal.loukili@usmba.ac.ma.

Pr. Fayçal Messaoudi is an accredited professor at the National School of Business and Management in Fez, Morocco. He is a member of the Research Laboratory in Management, Finance and Audit of Organizations, and Artificial Intelligence, Data Science and Emerging Systems laboratory. His main teaching and research interests concern Artificial Intelligence, Data Analysis, Database Management, and E-Marketing. He can be contacted at email: faycal.messaoudi@usmba.ac.ma.

Pr. Mohammed El Ghazi is an accredited professor at the Superior School of Technology in Fez, Morocco. He is a member of the Artificial Intelligence, Data Science and Emerging Systems laboratory. His major teaching and research focus involve Artificial Intelligence and Machine Learning, Networks and Telecommunications, and Computer Science. He can be contacted at email: mohammed.elghazi@usmba.ac.ma.