International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 3, June 2024, pp. 3220~3229
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i3.pp3220-3229

Alleviating cold start and sparsity problems in the micro, small, and medium enterprises marketplace using clustering and imputation techniques

Sri Lestari[1], Yulmaini[1], Aswin[2], Singgih Yulizar Ma'ruf[1], Sulyono[1], Ruki Rizal Nul Fikri[1]
[1]Faculty of Computer Science, Institute of Informatics and Business Darmajaya, Lampung, Indonesia
[2]Faculty of Economics and Business, Institute of Informatics and Business Darmajaya, Lampung, Indonesia

Article Info

Article history:

Received Nov 28, 2023
Revised Jan 29, 2024
Accepted Jan 30, 2024

Keywords:

Cold start
Imputation
K-means
Sparsity
System recommendation

ABSTRACT

Recommendation systems are often implemented in e-commerce and micro, small, and medium enterprises (MSMEs) marketplaces to improve consumer services by providing product recommendations according to their interests. However, it still faces problems, namely sparsity and cold start, thus affecting the quality of recommendations. This research proposes clustering and imputation techniques to overcome this problem. The clustering technique used is k-means, while the missing value imputation method uses average values. The imputation results are then implemented in the k-nearest neighbor (KNN) and naïve Bayes algorithms and evaluated based on performance accuracy. Experimental results show an increase in accuracy of 16.48% in the KNN algorithm from 83.52% to 100%. Meanwhile, the naïve Bayes algorithm increased accuracy by 35.30% from 64.70% to 100%.

Corresponding Author:

Sri Lestari
Faculty of Computer Science, Institute of Informatics and Business Darmajaya
ZA. Pagar Alam St., No. 93 Gedong Meneng, Bandar Lampung, Indonesia
Email: srilestari@darmajaya.ac.id

1. INTRODUCTION

Internet and mobile technology make it easy for users to access information anytime and anywhere. This also influences people's lifestyles, which are more intense in carrying out online activities, from accessing social media, e-commerce, and various other applications [1]. The same thing also happens to companies that change the way they do business; many companies utilize internet technology to provide services and market their products online. Online marketing will open up wider opportunities in terms of market reach, not only local, national, even international without being constrained by time and space. In addition, digital and social marketing offers significant opportunities in the form of lower costs for marketing, increasing brand awareness and increasing sales [2].

Digital media and social media have become important and integral components of a business marketing plan for a company [3]. With this media, companies will connect with customers, thereby increasing awareness of their brand, influencing consumer attitudes, getting feedback from consumers, and helping increase sales. Likewise with the marketplace which will connect many entrepreneurs and consumers, including micro, small, and medium enterprises (MSMEs). MSMEs are important for a country, because MSMEs are very capable of playing a role in encouraging regional economic growth and improving community welfare [4], [5]. In this case, many countries are digitizing MSMEs, including in Indonesia. There are many MSMEs marketplaces have emerged.

As the marketplace develops, the number of companies or MSMEs that join, the number of products and the number of transactions increases. This has an impact on the length of time it takes for customers to choose a product that suits their interests, because there are so many products on offer. Meanwhile, companies have difficulty recommending products that suit customer interests. Based on these problems, one solution that can be implemented is to implement a recommendation system.

Recommendation systems are efficient tools in filtering widely distributed information, changes in computer usage habits, personalization trends, and the emergence of internet access. The recommendation system has received a lot of attention recently, because it is able to direct various things to users effectively. The goals of implementing a recommendation system include increasing the number of sales, recommending to friends, increasing user satisfaction, increasing company revenue, and others [6]. A superior recommendation system is able to provide precise recommendations, but faces problems in the form of scalability, sparsity, and cold start [7]. Scalability is a condition where the system fails when the data increases in the number of items/users and is unable to provide recommendations with a reasonable response time. This is because the data dimensions are high or the data is very large in size, so it takes a lot of time to calculate the correlation between each pair of users or pairs of items [6]. Sparsity occurs because the rating matrix is very sparse, because very few users give ratings to the items they buy or view [6]. Cold start occurs when there are new users or new products, so the user's interests are not yet known [8].

Much research has been done to solve this problem, Xiong *et al.* [9] proposing k-means clustering and missing value prediction using the similarity algorithm. Rodpysh *et al.* [10] proposed multi-level singular value decomposition to solve cold start and sparsity problems. Hawashin *et al.* [11] used machine learning and user interest techniques to overcome the cold start problem. Meanwhile, in this study we propose a clustering approach and imputation techniques to solve sparsity and cold start problems. The cluster algorithm used is k-means, while the imputation technique uses the average value, and continues with testing using the k-nearest neighbor (KNN) and naïve Bayes algorithms.

2. RESEARCH METHOD

2.1. Collaborative filtering

Collaborative filtering is a recommendation technique that uses rating information from other users to produce recommendation suggestions [12]. This approach assumes that user preferences for products tend to remain constant over time, especially with users who have similar preferences. The use of collaborative filtering has proven to be very effective. In this algorithm, preference data can be grouped based on similarity of preferences, thereby enabling better decision making. The advantages of collaborative filtering include its ability to deal with content that is difficult to analyze.

Collaborative filtering operates by collecting and analyzing large amounts of data that reflect a user's activity, preferences, and ratings of an item. Then, based on the similarity of items selected by other users, the method projects recommendations for different items [13]. There are two main shortcomings in Collaborative Filtering, namely sparsity and scalability [14]. Sparsity arises when users provide few ratings, while scalability arises when the amount of data that needs to be searched for similarities is large enough.

2.2. Cold start

Cold start is a situation when the system faces challenges in providing recommendations for new items or users who do not yet have sufficient history or data in the system [15]. In cold start conditions for new users, the system still needs to have a history of the new user's interest direction because there is no rating data provided for existing products. So, it is difficult to provide product recommendations that suit the interests of new users. The same thing also happens to new items, where the user gives no rating to the item. This condition will result in the item not having the opportunity to be recommended to users.

2.3. Sparsity

Sparsity problems arise when new users enter the system, because information about these new users is still very limited. Evaluation results tend to be uninformative and tend to produce recommendations that are less relevant [16]. For example, new users may not be able to receive movie recommendations because no other users in the system have similar preferences to them. This study was done by following the steps as shown in Figure 1, starting with a literature study, namely looking for various references related to this research. The next step is data collection by taking public data from Grouplens. She was followed by data preprocessing, namely, taking the attributes needed in this research. After the data is preprocessed, the data is then clustered as labels and the basis for carrying out imputation. The imputation results are then implemented in the KNN and naïve Bayes algorithms and evaluated based on performance accuracy. These stages can be described in detail as follows.

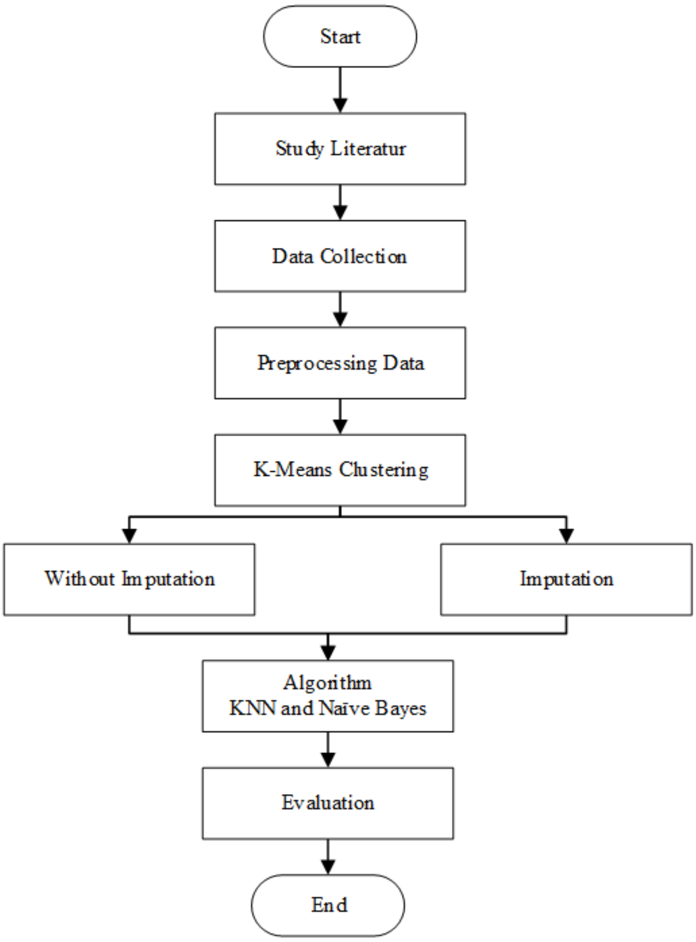

Figure 1. Research stages

2.3.1. Literature study

The first stage was to conduct a literature study to look for references from various sources such as books, the internet, journal articles and proceedings. The previous study was used as a guide in conducting this study. The method in this study involved data mining techniques, especially clustering and classification. The used algorithms are the k-means, KNN, and naïve Bayes.

2.3.2. Data collection

In this experimental study, MovieLens data was used it represented the character of data in the MSMEs marketplace to inform users, products and ratings. The MovieLens dataset was available on the website https://grouplens.org/datasets/movielens/100k/. This dataset consisted of 100,000 ratings on a scale ranging from 1-5, given by 943 users for 1,682 films. Each user provided ratings for at least 20 films, and the dataset also included demographic information such as the user's age, gender, occupation, and zip code. This dataset contained sparsity with 93.7%.

2.3.3. Data preprocessing

In this step, it was important to normalize the attributes with a specified scale in order to obtain optimal data mining results. In this context, the values in the age, gender and occupation columns were able to be transformed. In addition, abnormal data for outliers in the age column were removed to obtain higher quality data. In this case, age data is deleted if the value is below 10 and above 60. The data distribution can be seen in Figure 2.

2.3.4. Clustering k-means

After the data processing process was complete, this study was continued by applying the k-means algorithm to group data based on user demographic information based on the attributes age, gender and occupation. A number of techniques to be applied in determining the number of clusters are considered appropriate in the k-means algorithm include the silhouette method [17], sum of squared errors (SSE), elbow method [18], and Davies Bouldin index (DBI) [19]. In this study, the Davies-Bouldin Index is used to determine the optimal number of clusters. The use of DBI involved the optimization of the distance between clusters. It was used to minimize the distance between points within a cluster. If the distance between clusters was large, it indicated that the differences in characteristics between clusters were more clearly visible. It was because the similarity between clusters was lower. On the other hand, if the distance between clusters was small, this indicated that the objects in a cluster had a high level of similarity [20]. Evaluation of clustering results was done by determining the centroid and measuring it using the Davies Bouldin index (DBI) method. In this case, the optimal K was 4 clusters.

K-means clustering: The clustering algorithm is a technique used to group data into a number of groups (clusters), where k must be greater than or equal to 2. In this process, objects with a high level of similarity will be placed in the same cluster while increasing the distance from objects contained in other clusters [21]. Generally, the cluster method is used at an early stage to find user group profiles with a high degree of similarity. The user group profiles formed can function as preferences and have the potential to reduce the time required in the recommendation creation process [22]. Among the various cluster techniques available, the k-means method is the most frequently used because it uses a simple formula and the iteration process takes place quickly [23]. Distance calculation using Euclidean distance in k-means is shown in (1).

$$D(x_i, c_j) = \sqrt{\sum_{i=1}^{N}(x_i - c_j)^2} \tag{1}$$

where $D(i,j)$ is distance of data i to cluster center j, x_i is data to i on data attribute to k, and ej is enter point to k on attribute to k.

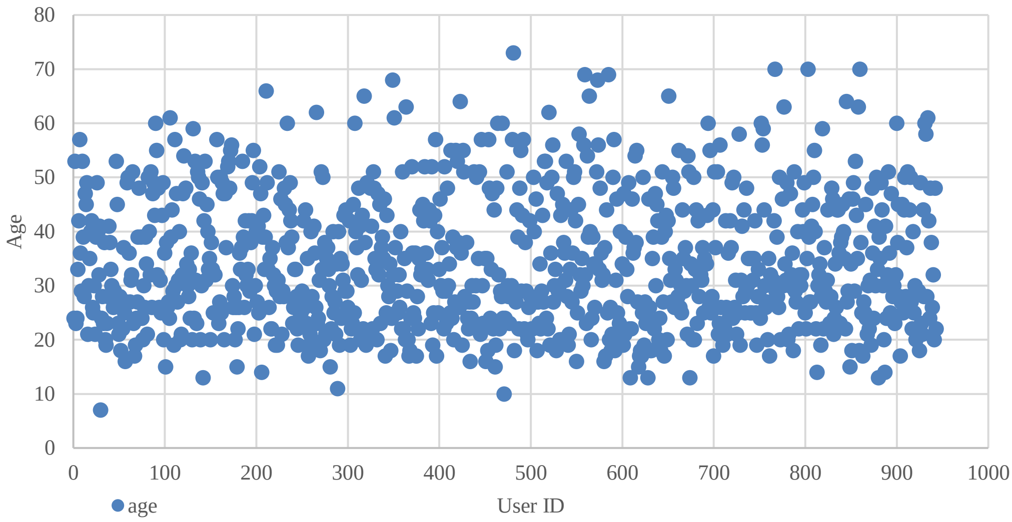

Figure 2. Data outliers in the age column

2.3.5. Imputation

The imputation process is filling in missing or empty data with specific values. In this study, the average function is used to replace empty rating values. The imputation process is carried out on each cluster; in this case, there are 4 clusters. The imputation process is carried out on each cluster with the aim of obtaining imputed values according to the character of the cluster.

2.3.6. K-nearest neighbors dan naïve Bayes

The testing stage was carried out to determine the influence of the imputation results on the quality of the recommendations produced. Testing was carried out by implementing the KNN and naive Bayes algorithms. Meanwhile, to measure the performance of the two algorithms, K-fold cross validation is used, namely by dividing the data into training data and testing data.

a. K-nearest neighbor (KNN)

The KNN algorithm is an algorithm that uses an easy similarity formula such as Euclidean distance, but has the ability to solve classification problems. The advantage of the KNN method is its ability to provide recommendations quickly and accurately with high quality. KNN generally utilizes two basic formulas that can be used to measure the similarity between training data and test data, namely Euclidean distance and cosine similarity [24]. Several researchers have tried to overcome the problem of data sparsity through a series of experiments [21]–[23], [25]. one of them is by applying the KNN algorithm. The distance calculation in the KNN algorithm is shown in (2).

$$Dist(x_1, x_2) = \sqrt{\sum_{i=1}^{n}(x_{1i} - x_{2i})^2} \tag{2}$$

where $Dist(x_1, x_2)$ is distance, x_{1i} is first data value, and x_{2i} is second data value.

b. Naïve Bayes

Naïve Bayes classifier is a Bayesian learning algorithm that has optimal speed and simplicity. This algorithm is based on Bayes' theorem, which produces a classifier by utilizing probability. The Bayesian approach is a statistical method used to predict certain parameters [26]. Naïve Bayes can be used to predict the probability of membership in a class [27], and is able to achieve a high level of accuracy with efficient time performance when applied to large databases. Calculations using the naïve Bayes algorithm are shown in (3).

$$P(H|X) = \frac{P(H|X)P(H)}{P(X)} \tag{3}$$

where X is unknown classified data, H is data hypothesis X in a class, $P(H|X)$ is hypothesis probability H on X condition, $P(H)$ is hypothesis probability $H, P(X|H)$ is probability X on H condition, and P(X) is probability X.

2.3.7. Evaluation

The evaluation was done by checking the level of accuracy to ensure that the results obtained reach a satisfactory level of quality. In this study, the evaluation used cross fold validation. It was used to assess models or algorithms and divide the data into training data and testing data [28]. This technique was often adopted in this study because it was proven to reduce bias in sampling. K-fold cross validation continuously divides data into training data and test data. It provided each data the opportunity to become test data. K in this context referred to the number of multiples used for the division between training data and test data, this study uses K=10 [29], [30].

3. RESULTS AND DISCUSSION

This study used clustering and imputation techniques to overcome cold start and sparsity problems. Clustering was done using the k-means algorithm on user data based on demographic information such as age, gender, and occupation. The cluster results were subjected to imputation to replace the empty values. Moreover, it was also used as a label. Moreover, a comparison was made of the evaluation results of implementing the KNN and naïve Bayes classification algorithms. The tool in this study used RapidMiner. It started with reading the data and removing data outliers to become normalized the data. Then, the next step was doing the best number of clusters. The data cleaning model and determining the number of clusters was seen in Figure 3. Meanwhile, the value result of Davies Bouldin Index test was seen in Table 1.

After the clustering test was done, the number of clusters was determined based on the DBI value. It stated that the optimal K value was the smallest value and close to 0. Based on Table 1, the best number of clusters was 4. It was because the DBI value showed the smallest. Based on this, it implemented the k-means algorithm by setting the value k=4.

The next step was to test with the KNN and naïve Bayes algorithms. The first scenario was done on data without imputation. The data was first labeled based on clusters, namely cluster 0-3, and continued processing with the KNN algorithm, and evaluated using cross fold validation to determine its accuracy. For more details, the test model using the KNN algorithm was seen in Figure 4.

The second scenario for imputation is carried out on each cluster. Imputation is carried out on clusters 0 to 3. This is done to obtain imputation values that represent the characteristics of each cluster. The next step is to evaluate the imputation model as in Figures 5(a) and 5(b). Figure 5(a) shows the imputation model by replacing data with the imputation technique, namely the average function. Meanwhile, Figure 5(b) shows the evaluation model for imputation results by implementing the KNN algorithm.

Similar steps were also taken by implementing the naïve Bayes algorithm. It was done to see the effect of imputation on increasing accuracy in several algorithms. In this case, the KNN and naïve Bayes algorithms were done. Based on experimental results for performance testing, it showed an increase in accuracy. Accuracy results were seen in Table 2.

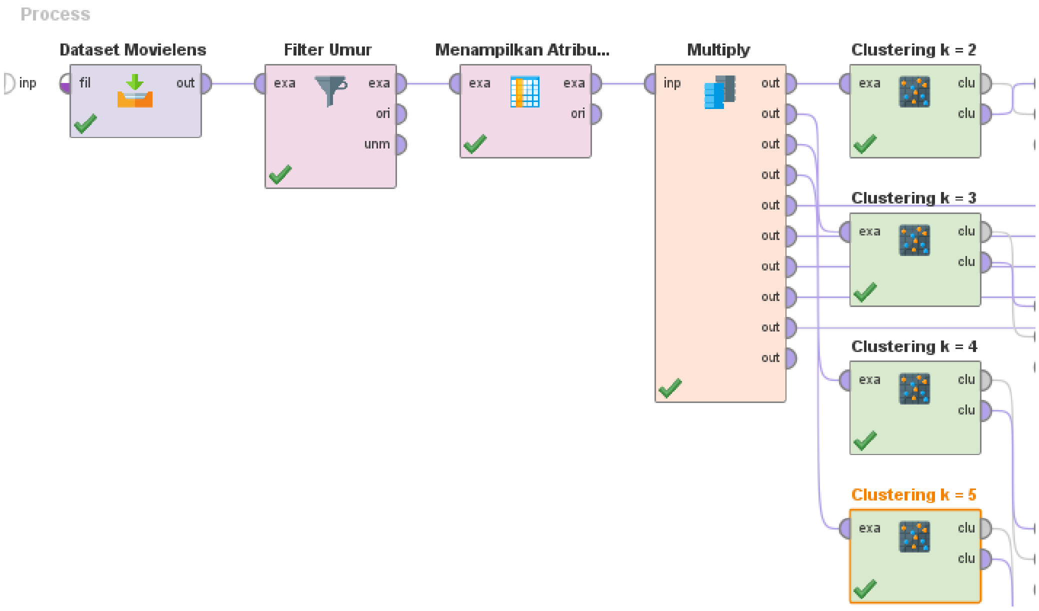

Figure 3. Data cleaning and determining the number of clusters

Table 1. The value results of DBI test

No	Cluster	DBI
1	2	0.803
2	3	0.846
3	4	0.788
4	5	0.855

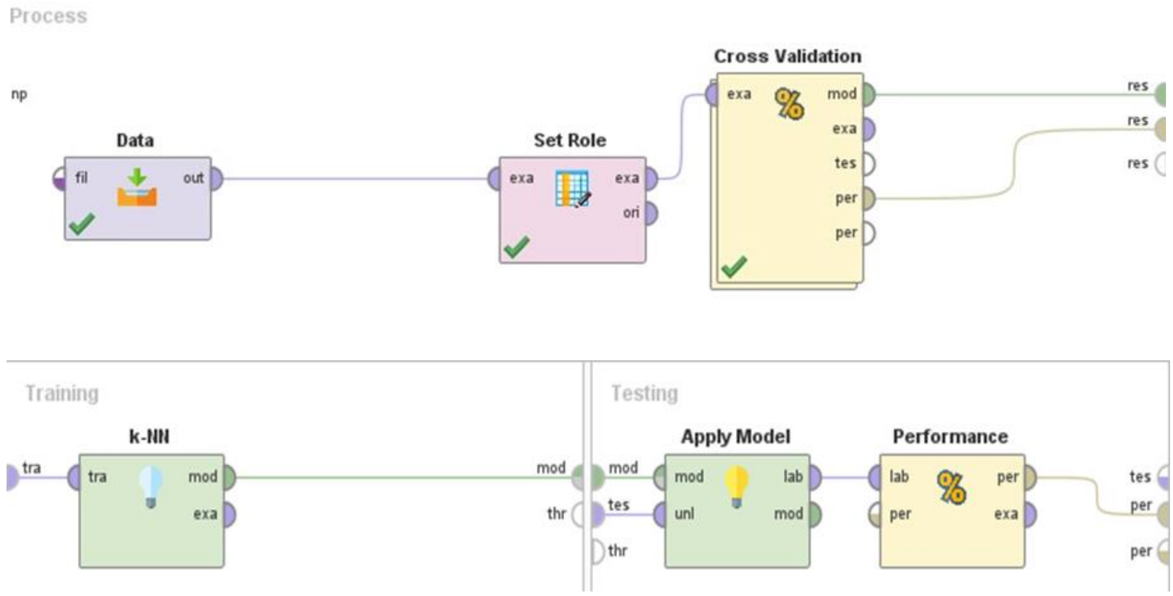

Figure 4. The testing model using the KNN algorithm

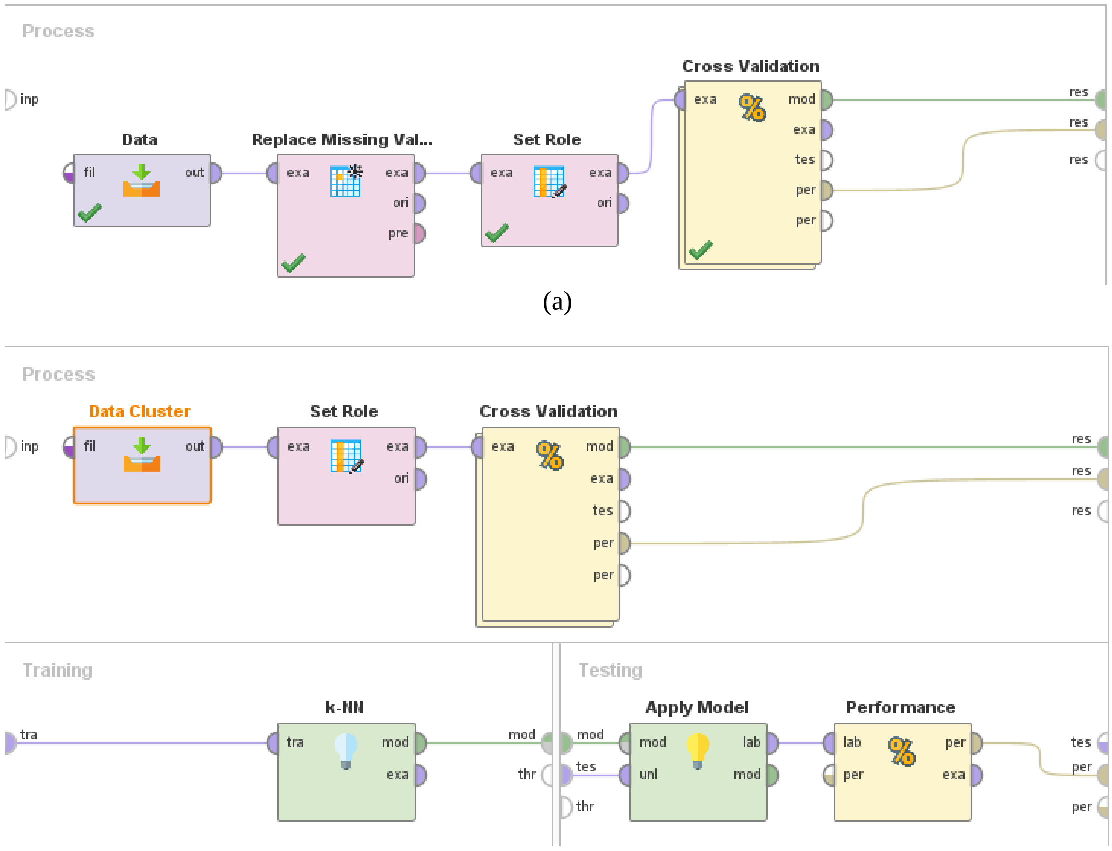

Figure 5. The evaluate the imputation model: (a) Imputation model and (b) the testing model using the KNN algorithm on data imputation

Table 2 showed that the accuracy value of the KNN algorithm without imputation was 83.53% and after imputation it showed an increase in accuracy to 100%. A similar thing also happened with the naïve Bayes algorithm. The accuracy value increased from 64.70% to 100%. It showed that imputation solved the sparsity problem. Besides, it also solved the cold start problem for new items. Previously, it had a minimal or no rating. Thus, there was a rating to be used as a reference in providing recommendations.

Table 2. The comparison of accuracy results using KNN and naïve Bayes

No	Status	Performance Accuracy	
		KNN	Naïve Bayes
1	Without imputation	83.52%	64.70%
2	Imputation	100%	100%

5. CONCLUSION

Recommendation systems are an effective way to improve services in e-commerce and marketplaces by providing personalized service. However, it faces problems, namely cold start and sparsity, which affects the quality of the recommendations produced. Based on this, this research proposes clustering and imputation techniques, namely the k-means algorithm and the missing value average imputation method. The results of the imputation are then implemented in the KNN and naïve Bayes algorithms and evaluated using accuracy performance**.** The experimental results showed an increase in accuracy after imputation of 16.48% on the KNN algorithm from 83.52% to 100%. Meanwhile, the naïve Bayes algorithm increased accuracy by 35.30% from 64.70% to 100%. It stated that clustering and imputation techniques solved sparsity and cold start problems that often occurred in recommendation systems. It was widely used in e-commerce and marketplaces, including MSMEs marketplaces.

ACKNOWLEDGEMENTS

This study was funded by the Directorate of Research, Technology, and Community Service (or *DRTPM*) under the Excellence-Applied Research for Higher Education scheme with the number 0536/E5/PG.02.00/2023.

REFERENCES

[1] M. Rostami, V. Farrahi, S. Ahmadian, S. M. Jafar Jalali, and M. Oussalah, “A novel healthy and time-aware food recommender system using attributed community detection,” *Expert Systems with Applications*, vol. 221, Art. no. 119719, Jul. 2023, doi: 10.1016/j.eswa.2023.119719.

[2] Y. K. Dwivedi *et al.*, “Setting the future of digital and social media marketing research: Perspectives and research propositions,” *International Journal of Information Management*, vol. 59, p. 102168, Aug. 2021, doi: 10.1016/j.ijinfomgt.2020.102168.

[3] A. T. Stephen, “The role of digital and social media marketing in consumer behavior,” *Current Opinion in Psychology*, vol. 10, pp. 17–21, Aug. 2016, doi: 10.1016/j.copsyc.2015.10.016.

[4] S. Juminawati, A. Hamid, E. Amalia, Ma. Mufraini, and A. Sopyan Mulazid, “The effect of micro, small and medium enterprises on economic growth,” *Budapest International Research and Critics Institute*, vol. 4, no. 3, pp. 5697–5704, 2021, doi: 10.33258/birci.v4i3.2368.

[5] T. Tambunan, “Recent evidence of the development of micro, small and medium enterprises in Indonesia,” *Journal of Global Entrepreneurship Research*, vol. 9, no. 1, Jan. 2019, doi: 10.1186/s40497-018-0140-4.

[6] M. Singh, “Scalability and sparsity issues in recommender datasets: a survey,” *Knowledge and Information Systems*, vol. 62, no. 1, pp. 1–43, Oct. 2018, doi: 10.1007/s10115-018-1254-2.

[7] 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.

[8] M. Mohamed, M. Khafagy, H. Elbeh, and A. Abdallh, “Sparsity and cold start recommendation system challenges solved by hybrid feedback,” *International Journal of Engineering Research and Technology*, vol. 12, pp. 2734–2741, Dec. 2019.

[9] C. Xiong, Y. Liu, Z. Yang, and E. Liu, “Improving recommendation performance with clustering and missing value prediction,” Mar. 2020. doi: 10.1109/ciss48834.2020.1570617370.

[10] K. Vahidy Rodpysh, S. J. Mirabedini, and T. Banirostam, “Resolving cold start and sparse data challenge in recommender systems using multi-level singular value decomposition,” *Computers & Electrical Engineering*, vol. 94, Art. no. 107361, Sep. 2021, doi: 10.1016/j.compeleceng.2021.107361.

[11] B. Hawashin, S. Alzubi, A. Mughaid, F. Fotouhi, and A. Abusukhon, “An efficient cold start solution for recommender systems based on machine learning and user interests,” Apr. 2020. doi: 10.1109/sds49854.2020.9143953.

[12] M. F. Aljunid and M. Dh, “An efficient deep learning approach for collaborative filtering recommender system,” *Procedia Computer Science*, vol. 171, pp. 829–836, 2020, doi: 10.1016/j.procs.2020.04.090.

[13] J. S. Breese, D. Heckerman, and C. Kadie, “Empirical analysis of predictive algorithms for collaborative filtering,” *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence (UAI1998)*, Jan. 2013.

[14] A. Fareed, S. Hassan, S. B. Belhaouari, and Z. Halim, “A collaborative filtering recommendation framework utilizing social networks,” *Machine Learning with Applications*, vol. 14, Art. no. 100495, Dec. 2023, doi: 10.1016/j.mlwa.2023.100495.

[15] M. Rahman, I. A. Shama, S. Rahman, and R. Nabil, “Hybrid recommendation system to solve cold start problem,” *Journal of Theoretical and Applied Information Technology*, vol. 100, no. 11, pp. 3562–3580, 2022.

[16] G. Naik, K. N. Narasinga Rao, and A. Baje, “ICT intervention challenges in education in Rural India,” in *Lecture Notes in Networks and Systems*, Springer Singapore, 2020, pp. 183–197. doi: 10.1007/978-981-15-0630-7_18.

[17] W. Kwedlo, “A clustering method combining differential evolution with the k-means algorithm,” *Pattern Recognition Letters*, vol. 32, no. 12, pp. 1613–1621, Sep. 2011, doi: 10.1016/j.patrec.2011.05.010.

[18] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, “Integration k-means clustering method and elbow method for identification of the best customer profile cluster,” *IOP Conference Series: Materials Science and Engineering*, vol. 336, Art. no. 12017, Apr. 2018, doi: 10.1088/1757-899x/336/1/012017.

[19] M. Mughnyanti, S. Efendi, and M. Zarlis, “Analysis of determining centroid clustering x-means algorithm with Davies-Bouldin index evaluation,” *IOP Conference Series: Materials Science and Engineering*, vol. 725, no. 1, Art. no. 12128, Jan. 2020, doi: 10.1088/1757-899x/725/1/012128.

[20] S. Kapil and M. Chawla, “Performance evaluation of k-means clustering algorithm with various distance metrics,” Jul. 2016. doi: 10.1109/icpeices.2016.7853264.

[21] J. Pérez-Ortega, N. N. Almanza-Ortega, and D. Romero, “Balancing effort and benefit of k-means clustering algorithms in Big Data realms,” *PLOS ONE*, vol. 13, no. 9, Art. no. e0201874, Sep. 2018, doi: 10.1371/journal.pone.0201874.

[22] U. Kuzelewska, “Clustering algorithms in hybrid recommender system on MovieLens data,” *Studies in Logic, Grammar and Rhetoric*, vol. 37, no. 1, pp. 125–139, Jan. 2014, doi: 10.2478/slgr-2014-0021.

[23] S. Theodoridis and K. Koutroumbas, Eds., “Copyright,” in *Pattern Recognition*, Fourth Edi., Elsevier, 2009, p. iv. doi: 10.1016/b978-1-59749-272-0.50001-3.

[24] D. A. Adeniyi, Z. Wei, and Y. Yongquan, “Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method,” *Applied Computing and Informatics*, vol. 12, no. 1, pp. 90–108, Jan. 2016, doi: 10.1016/j.aci.2014.10.001.

[25] J. Cheng and L. Zhang, “Jaccard coefficient-based bi-clustering and fusion recommender system for solving data sparsity,” in *Lecture Notes in Computer Science*, Springer International Publishing, 2019, pp. 369–380. doi: 10.1007/978-3-030-16145-3_29.

[26] H. Yücel, S. S. Tüzüner, and C. Massey, “Application of peak based-Bayesian statistical method for isotope identification and categorization of depleted, natural and low enriched uranium measured by LaBr3:Ce scintillation detector,” *Nuclear Engineering and Technology*, vol. 55, no. 10, pp. 3913–3923, Oct. 2023, doi: 10.1016/j.net.2023.07.006.

[27] I. L. Qurnia, E. Prasetyo, and R. F. Zainal, “Classification of diabetes disease using Naive Bayes case study: Siti Khadijah Hospital,” *JEECS (Journal of Electrical Engineering and Computer Sciences)*, vol. 1, no. 2, pp. 147–151, Dec. 2016, doi: 10.54732/jeecs.v1i2.177.

[28] M. T R, V. K. V, D. K. V, O. Geman, M. Margala, and M. Guduri, “The stratified K-folds cross-validation and class-balancing methods with high-performance ensemble classifiers for breast cancer classification,” *Healthcare Analytics*, vol. 4, p. 100247, Dec. 2023, doi: 10.1016/j.health.2023.100247.

[29] S. M. Malakouti, M. B. Menhaj, and A. A. Suratgar, “The usage of 10-fold cross-validation and grid search to enhance ML

methods performance in solar farm power generation prediction," *Cleaner Engineering and Technology*, vol. 15, Art. no. 100664, Aug. 2023, doi: 10.1016/j.clet.2023.100664.

[30] S. M. Malakouti, "Babysitting hyperparameter optimization and 10-fold-cross-validation to enhance the performance of ML methods in predicting wind speed and energy generation," *Intelligent Systems with Applications*, vol. 19, Art. no. 200248, Sep. 2023, doi: 10.1016/j.iswa.2023.200248.

BIOGRAPHIES OF AUTHORS

Sri Lestari obtained her doctorate (Dr) from the Electrical Engineering Doctor Program, Universitas Gadjah Mada, Yogyakarta, Indonesia in 2019. She is a lecturer in the Department of Computer Science, Institut Informatika and Bisnis Darmajaya, Bandar Lampung, Indonesia. Her research interests include artificial intelligence, recommendation system, collaborative filtering, data mining, decision support systems, and software engineering. Her representative published articles listed as follows: "Imputation missing value to overcome sparsity problems in the recommendation system" (Jurnal Rekayasa Sistem dan Teknologi Informasi (RESTI), Vol 7 No. 1, 2023). "Implementation of the C4.5 algorithm for micro, small, and medium enterprises classification" (International Journal of Electrical and Computer Engineering (IJECE), Vol 12 No. 6, 2022). "The improvement of personal service in the MSMEs marketplace using a rank-based collaborative filtering approach" (International Conference on Electrical and Information Technology (IEIT), Politeknik Negeri Malang, 2022). "Porat^{Rank} to improve performance recommendation system" (Lecture Notes in Electrical Engineering, Springer 2021), Decision support system for service quality using SMART and fuzzy ServQual methods" (JUITA: Jurnal Informatika, 2021), "WP-Rank: rank aggregation based collaborative filtering method in recommender system" (International Journal of Engineering & Technology (UAE), 2018), "Performance comparison of rank aggregation using Borda and Copeland in recommender system" (International Workshop on Big Data and Information Security (IWBIS 2018)), "NRF: normalized rating frequency for collaborative filtering" (The 2018 International Conference on Applied Information Technology and Innovation (ICAITI 2018)). She can be contacted at email: srilestari@darmajaya.ac.id.

Yulmaini obtained her master's degree from the Postgraduate Program, at Universitas Gadjah Mada, Yogyakarta, Indonesia in 2011. She is a lecturer in the Department of Computer Science, Institut Informatika and Bisnis Darmajaya, Bandar Lampung, Indonesia. Her research interests include artificial intelligence, decision support systems, and software engineering. Her representative published articles lists as follows: "Application of Tsukamoto's fuzzy inference system in determination of concentration for students' thesis topics" (Journal of Critical Reviews , 2019), "Formulating a higher education competitiveness model" (Journal of Talent Development and Excellence, 2020), "Competitiveness universities strategy development based on the research and information technology" (Journal of Engineering and Technology Management, 2020), "Improvement of one-dimensional Fisher face algorithm to extract the features (Case study: Face Recognition)" (in Proceedings of International Conference on Science and Technology 2019), "Implementation of analytic hierarchy process for determining priority criteria in higher education competitiveness development strategy based on RAISE++ model (Proceedings of The 2nd Joint International Conference on Emerging Computing Technology and Sports (JICETS) 2019), "Application of Tsukamoto's fuzzy inference system in determination of concentration for students' thesis topics" (Proceedings of the International Conference on Information Technology and Business (ICITB), 2019). She can be contacted at email: yulmaini@darmajaya.ac.id.

Aswin is a lecturer in the Department of Management specialization in marketing, Institut Informatika and Bisnis Darmajaya, Bandar Lampung, Indonesia. Her research interests include consumer behavior, customer satisfaction, service quality, brand equity, and social media marketing. Her representative published articles listed as follows: "Role of social media marketing to enhance the supply chain and business management" (International Journal of Supply Chain Management (IJSCM), 2020), "Analisis brand equity perguruan tinggi swasta di Bandar Lampung" (National Journal of Gentiaras dan Akuntansi (GEMA), 2021), "Analysis of student and student interest in the Darmajaya IIB management study program (S1) concentration in Bandar Lampung" (National Journal of Applied Accounting and Business), 2019), "Analysis of community influence on the use of family planning devices in Bandar Lampung City" (National Journal of GEMA Economic, 2017). She can be contacted at email: aswin@darmajaya.ac.id.

Singgih Yulizar Ma'ruf obtained a bachelor's degree from the Informatics Engineering undergraduate program, IIB Darmajaya, Bandar Lampung, Indonesia in 2022. Currently he is continuing his studies in the Master of Informatics Engineering with a concentration in Data Science at IIB Darmajaya. He can be contacted at email: singgihyulizar07@gmail.com.

Sulyono Graduated in the Informatics Engineering Study Program, Faculty of Computer Science, Darmajaya Institute of Informatics and Business in 2007, graduated master's program in informatics engineering with a specialization in software engineering, Darmajaya Institute of Informatics and Business in 2015. Became a programmer since sitting on the lecture bench, the application product has been widely used by users. Currently, he is a permanent lecturer in the Informatics Engineering Study Program, Faculty of Computer Science, Darmajaya Institute of Informatics and Business, Bandar Lampung City, Lampung Province. Capable of programming-based courses from the first semester to the final semester. Has attended several pieces of training that are appropriate in the field of teaching, especially programming. He can be contacted at email: sulyono@darmajaya.ac.id.

Ruki Rizal Nul Fikri Graduated in the Informatics Engineering Study Program, Faculty of Computer Science, Darmajaya Institute of Informatics and Business in 2012, graduate master's program in Informatics Engineering with a specialization in Software Engineering, Darmajaya Institute of Informatics and Business in 2015. Became a programmer since sitting on the lecture bench, the application product has been widely used BU users. Currently, he is a permanent lecturer in informatics and business, Bandar Lampung city, Lampung Province. Capable of programming-based courses from the first semester to the final semester. Has attended several pieces of training that are appropriate in the field of teaching, especially programming. He can be contacted at email: rukirizal@darmajaya.ac.id.