International Journal of Technology

http://ijtech.eng.ui.ac.id

Research Article

A Novel Hybrid Framework for Cold-Start Resolution in Traditional Craft Recommender Systems

I Gusti Agung Gede Arya Kadyanan[1], Ni Made Ary Esta Dewi Wirastuti[2*], Gede Sukadarmika[2], Ngurah Agus Sanjaya ER[1], Is-Haka Mkwawa[3], Muhamad Asvial[4]

[1]Department of Informatics, Faculty of Mathematics and Natural Sciences, Udayana University, Badung, 80361, Indonesia

[2]Department of Electrical Engineering, Faculty of Engineering, Udayana University, Badung, 80361, Indonesia

[3]School of Engineering, Computing and Mathematics, University of Plymouth, Plymouth, PL4 8AA, United Kingdom

[4]Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia, Depok, 16424, Indonesia

[*]Corresponding author: dewi.wirastuti@unud.ac.id; Tel.: +62-361246163; Fax.: +62-361246163

Abstract: Recommender systems are essential for guiding users to relevant items and locations; however, the cold-start problem caused by missing or unrated items remains a persistent challenge. This study proposes a novel hybrid framework that integrates the item-based clustering hybrid method (ICHM) with the Slope One algorithm to specifically address cold-start scenarios in traditional craft recommender systems. A unique dataset of 48 craft locations and 60 traditional Balinese craft products, collected through direct field observation, representing an original contribution that bridges cultural heritage and advanced recommendation technologies, was used for validation. The framework predicts missing ratings using Slope One and generates recommendation scores via a weighted-sum function, providing dual recommendations for both products and production locations. The experimental results indicate high prediction accuracy, with overall mean absolute error values well below acceptable thresholds, confirming the system's reliability and robustness. Beyond technical contributions, it highlights the socio-economic and cultural potential of RSs in preserving and promoting local heritage.

Keywords: Cold start; Item-based clustering hybrid method; Recommender system; Slope one; Traditional handicraft

1. Introduction

Bali has many distinctive features unique to its people, including traditional crafts, cultural arts, and tourism, making it one of the most developed travel destinations in the world today. It is internationally recognized for its rich cultural identity and deep integration of tradition, creativity, and tourism. Bali was selected as the case study because of its internationally recognized cultural heritage and thriving creative tourism sector. Traditional crafts in Bali not only represent cultural identity but also play a significant role in supporting the local economy and tourism ecosystem. Despite this importance, systematic digital tools for recommending craft products and production sites remain underexplored. Building motivation to recommend production locations directly to tourists is very important because this not only facilitates transactions but also provides a deep cultural experience, while also serving as a means of authenticity verification to ensure that the crafts are made by legitimate local artisans.

Therefore, this study introduces a hybrid recommender framework that integrates the item-based clustering hybrid method (ICHM) with the slope one algorithm to address the cold-start problem. The framework is designed to support both cultural preservation and tourism digital innovation. Although the evaluation is conducted on a dataset of Balinese crafts, the frame-

https://doi.org/10.14716/ijtech.v17i2.8289
Received December 2025; Revised January 2026; Accepted March 2026

work is generalizable and adaptable for other cultural heritage domains and recommendation contexts. In this way, the study contributes both a methodological innovation and an applied demonstration in a real-world heritage tourism environment.

A recommendation system is a system that helps users choose or purchase an item. According to several studies, the content-based filtering method is one of the system recommendation techniques that can be employed. However, it has several drawbacks, including the inability to recommend items with content types different from those the user has already chosen. A collaborative filtering technique was developed to address the drawbacks of content-based filtering (Fareed et al., 2023; Widayanti et al., 2023; Patel et al., 2023; Ko et al., 2022). According to some studies, collaborative filtering is effective at making system recommendations; however, it still has flaws with items that have never received a user rating. This is known as the "cold-start problem," which causes the item to sink in the system and not be recommended (Tran, 2025; Sarwar et al., 2001). Theoretically, the problem formulation in this study is based on a U × I matrix (where U is the user and I is the item/location), with the aim of predicting empty rankings amid high data sparsity in local crafts. This study formally defines a double cold-start scenario, namely, item cold-start (for new craft products and locations that do not yet have ratings) and user cold-start (new users without interaction history).

This led to the development of a hybrid collaborative filtering technique that blends collaborative and content-based filtering (Afoudi et al., 2021; Aljunid and Huchaiah, 2021). The hybrid collaborative filtering approach is a popular and effective recommendation system that generates superior suggestion outputs. The hybrid collaborative filtering method can be combined in three ways: sequentially, linearly, and using the ICHM. The ICHM is a hybrid merging technique (Akbar et al., 2025; Praditya et al., 2022). To prevent the system from sinking, the item-based cluster method can solve the cold-start problem, which involves making recommendations for products that are new or have not yet been rated. The item rating indicates the quality of a traditional Balinese craft. According to certain studies, the accuracy of the suggestion results will be lowered if an item has an empty rating. The Slope One algorithm can be used to predict empty ratings on an item (Yannam et al., 2023). The item to be predicted and the prior user's rating are the two inputs required by the Slope One algorithm. When combined, these two techniques complement each other. ICHM manages new-item clustering, while Slope One estimates missing ratings, offering a balanced solution to cold start challenges.

This study combines two earlier studies on the use of the Slope One Algorithm to anticipate empty ratings (Yannam et al., 2023) and the application of the ICHM approach to solve the cold start issue in the recommendation system (Praditya et al., 2022). A mobile-based system is being developed to generate recommendations for traditional Balinese handicraft-making locations. This system uses the slope-one algorithm and ICHM. In addition to offering products to consumers, this developed recommendation system can provide information that can assist other users in selecting a location for traditional Balinese crafts. This dual recommendation capability, which covers both crafts and production sites, distinguishes the proposed system from conventional single-domain recommender models.

Although the latest cold-start solutions, such as deep hybrid models, meta-learning, and graph-based methods, show high performance, they tend to be very data-hungry, making them inadequate when applied to domains with high sparsity and limited datasets. Existing hybrid recommendation models mostly focus only on e-commerce products and cannot accommodate dual recommendation needs (physical products and geographic locations). This explicitly shows that a research gap remains regarding the cold-start solutions specifically designed for traditional craft recommendation systems. The proposed framework addresses these limitations differently by using the ICHM and Slope One approaches, which are more computationally efficient and resistant to local data gaps. The cold start issue in the recommendation system for traditional craft manufacturing locations can be resolved by combining the Slope One algorithm with the ICHM.

Although traditional Balinese crafts hold significant cultural and economic value, digital recommendation systems that address the cold-start problem and provide dual recommendations for both products and production sites remain underexplored. Most previous studies have focused on mainstream e-commerce or general tourism platforms and have not leveraged field-collected datasets to preserve and promote local heritage. Therefore, a hybrid framework that integrates the ICHM with the Slope One algorithm is needed specifically for heritage-based craft recommendation systems.

Therefore, this study aims to develop a hybrid recommendation system for traditional Balinese crafts that integrates the ICHM with the Slope One algorithm to address the cold-start problem. The system aims to provide accurate and reliable recommendations for both craft products and their production locations (Hendrarini et al., 2022), supporting cultural heritage promotion and offering practical guidance for tourists and consumers.

Although the combination of clustering with Slope One has been explored in previous studies and is not conceptually new in hybrid CF, this study's novelty lies in the adaptation and application of a hybrid recommendation framework that integrates ICHM with the Slope One algorithm to specifically address the cold-start problem in the context of traditional craft recommendation systems. Unlike previous studies that mainly focused on mainstream e-commerce or tourism platforms, this study applies and validates the framework using a unique dataset of traditional craft locations and products collected directly from field observations. This approach not only demonstrates technical innovation but also supports the sustainable promotion of cultural heritage through intelligent digital systems. This combination highlights not only the methodological advancement but also its originality in preserving and promoting local heritage while ensuring the framework's generalizability to broader recommender system applications.

The remainder of this article is organized as follows: Section 2 details the materials and methods related to the proposed framework. Section 3 presents the results of the experimental evaluations and discusses the system performance. Finally, Section 4 summarizes the study's conclusions and future development directions.

2. Materials and Methods

2.1 Recommendation System

The recommendation system is an application paradigm that makes suggestions about various items to help users make a decision. A recommendation system is an application model and the findings of observations that can suggest a location or an object to users based on their preferences and conditions (Di et al., 2026; Kannout et al., 2024; Yannam et al., 2023). It functions as an intelligent decision-support mechanism that personalizes content delivery by leveraging user behavior and item characteristics.

2.2 Hybrid Collaborative Filtering

The hybrid collaborative filtering method typically integrates many recommendation techniques to generate output and obtain improved recommendations (Mandalapu et al., 2023; Burke, 2007). By combining the collaborative and content-based approaches, the hybrid collaborative filtering method addresses the drawbacks of both approaches and enables them to support one another in overcoming their respective limitations. According to (Nicart et al., 2024; Q. Li and Kim, 2003), there are multiple ways to combine the hybrid collaborative filtering method. These include sequential, linear, and clustering-based combinations, each of which balances the trade-offs between accuracy, scalability, and adaptability.

2.3 Maintaining the Integrity of Specifications

The ICHM is a merging technique that computes item similarity by combining user ratings with item content information. One benefit of the ICHM is its ability to propose goods that have never been rated. The cold-start problem is the problem of recommending items that have not been rated by others. ICHM is preferred over other hybrid architectures (such as conventional linear weighting) because it uses a cluster-based approach that effectively narrows the search space and reduces local sparsity (data gaps) issues before performing the final ranking calculation.

According to Chung et al., 2014, the stages of the hybrid item-based cluster technique include the following: creating a group-rating matrix, applying the clustering method on the item content, and determining the likelihood that each item belongs to each cluster. This process allows the clustering of new or unrated items based on shared characteristics, ensuring that such items can still appear in user recommendations even without prior ratings. In this study, we used the k-means clustering algorithm. In the final stage following clustering, the probability of each item or link to the cluster is determined using Equation (1):

$$\text{Pro}(j,k) = 1 - \frac{CS(j,k)}{\max CS(i,k)} \tag{1}$$

In Equation (1), Pro(j,k) is the probability value (likelihood) of item j entering cluster k, CS(j,k) represents the Cosine Similarity value between item j and cluster center k, and maxCS(i,k) is the maximum similarity value within that cluster.

Several k clusters and item attributes are necessary inputs for generating the group-rating. The steps to create the group ratings are as follows:

1. Select a range of k values to serve as the starting point for cluster midpoints.

2. Until nothing changes, repeat steps a) and b) as follows:

   a) Sort each item into the most similar cluster according to its content.

   b) Recalculate the center value of each cluster.

   c) The probability value in relation to the cluster center value for each item is determined.

   d) The output is the number of k clusters and the probability value of each item relative to the cluster center value. Next, the following procedures were used to determine the similarity value:

      i. The group-rating matrix and item-rating matrix undergo similarity calculations, and the outcomes are subsequently aggregated for prediction computations

         a) A similarity based on Pearson's correlation is a correlation-based formula that determines item-rating similarity. It is represented in Equation (2) as follows:

$$\text{sim}(i,j) = \frac{\sum_{u=1}^{m}(R_{u,i} - \bar{R}_i)(R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u=1}^{m}(R_{u,i} - \bar{R}_i)^2} \times \sqrt{\sum_{u=1}^{m}(R_{u,j} - \bar{R}_j)^2}} \tag{2}$$

         In Equation (2), sim(i,j) is the similarity (correlation) between items i and j, $R_{u,i}$ is the rating of user u for item i, $\bar{R}_j$ is the average rating of item i, and m represents the total number of users in the system.

         b) The similarity group rating is determined using the adjusted cosine similarity equation, as shown in Equation (3):

$$\text{Pro}(j,k) = 1 - \frac{CS(j,k)}{\max CS(i,k)} \tag{3}$$

The variables in Eq. (3) are identical to those in Eq. (1), where this equation is applied to find the probability of similarity in the context of groups/clusters.

ii. We now use a linear combination equation to combine item-rating similarity data with group-rating similarity, as shown in Equation (4):

$$\text{sim}(k,l) = \text{sim}(k,l)_{\text{item}} \times (1-c) + \text{sim}(k,l)_{\text{group}} \times c \tag{4}$$

Here, sim(k,l) is the final similarity value of the hybrid, sim(k,l)$_{item}$ is the pure item-based similarity, sim(k,l)$_{group}$ is the group-based similarity, and c is the linear combination's weight parameter constant.

3. Rating prediction is performed under two cases depending on the data condition: non-cold start (for rated items) and cold start (for unrated or new items).

a) Items rated by multiple users may have a non-cold start issue. This condition uses the weighted average of deviation method, which is derived from the average of rated items, as shown in Equation (5):

$$P_{u,k} = \bar{R}_k + \frac{\sum_{i=1}^{n}(R_{u,i} - \bar{R}_i) \times \text{sim}(k,i)}{\sum_{i=1}^{n} |\text{sim}(k,i)|} \tag{5}$$

In Equation (5), P$_{u,k}$ is the final ranking prediction of user u for item k, $\bar{R}_k$ is the average rating of item k, and sim(k,i) is the similarity weight.

b) The cold-start problem is the condition of new items entering the system without any user ratings. The weighted sum technique is employed, and the equation is shown in Equation (6):

$$P_{u,k} = \frac{\sum_{i=1}^{n} R_{u,i} \times sim(k,i)}{\sum_{i=1}^{n} |\text{sim}(k,i)|} \tag{6}$$

In Equation (6), $P_{u,k}$ is the predicted rating for a new item using the weighted sum of existing interactions, without including the average deviation because the item does not yet have a rating history.

2.4 K-Means Clustering Algorithm

The data were grouped using the k-means clustering method according to the center point (Aparna and Nair, 2016). K-means clusters are used to organize data by optimizing the data in the cluster. J.B. MacQueen first presented k-means clusters in 1976. It is widely used for its simplicity, scalability, and strong performance in partitioning high-dimensional data. The K-Means algorithm was chosen for clustering over density-based methods (such as DBSCAN) because this specific craft dataset is relatively small and has structured cultural category boundaries around the centroid, making the spatial K-Means approach much more stable and less sensitive to local data density variations. The steps involved in the k-means clustering algorithm are as follows (Sari and Armansyah, 2024; Tarigan, 2023; Larasati et al., 2019; Bock, 2007):

1. Choose the number of clusters, k, you wish to create.

2. As many as k random values should be generated for the center of the initial cluster (centroid).

3. The Euclidean distance equation is used to determine the distance between each input data point and each centroid until the closest distance between each data point and the

centroid is determined. The Euclidean distance equation is shown in Equation (7):

$$D_e = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \tag{7}$$

4. Group all of the data according to the closest centroid (shortest distance).
5. Using Equation (8), the centroid value is updated with the cluster's average:

$$C_j = \frac{\sum_{x \in C_j} x}{m_j} \tag{8}$$

6. Continue steps 2 through 5 until the members in each cluster remain the same.

2.5 Elbow Method

The Elbow technique generates information for deciding on the optimal number of clusters by examining the proportion of comparison findings between the number of clusters that would form an elbow at a point or the largest difference (Larasati et al., 2021; Shi et al., 2021). By choosing the cluster value and then adding it to be used as a data model in identifying the optimal cluster, this method provides insights into identifying the optimal cluster. The proportion of computations generated serves as a comparison of the additional clusters. With a graph as the source of information, various percentage outcomes from each cluster value can be displayed. The cluster value is optimal when the first and second cluster values form an angle in the graph that represents the largest drop.

The sum of squared error (SSE) of each cluster value is computed to obtain the comparison because the more the clusters, the lower the SSE value. Equation (9) shows the SSE equation for the k-means method (Bhuiyan and Khan, 2024; Nainggolan et al., 2019).

$$SSE = \sum_{j=1}^{k} \sum_{x \in C_j} (\text{dist}(x, m_j))^2 \tag{9}$$

The hyperparameter tuning strategy for k-means in this framework also involves the double cluster center initialization technique (k-means++) to avoid the risk of convergence to a poor local optimum during iteration.

2.6 The Slope One Algorithm

The Slope One algorithm is one method that can be used to forecast an occurrence. Typically, the Slope One method is employed to forecast an item's rating. The user's rating and the item to predict are the two inputs required by the Slope One algorithm. The Slope One approach can decrease overfitting, indicating that the model is subject to random error and produces subpar prediction outcomes (Saeed and Mansoori, 2018). The Slope One algorithm is employed to solve issues with recommendation systems that arise from a lack of user ratings. Slope One was chosen over other collaborative filtering methods, such as matrix factorization or KNN, because this algorithm is much lighter in terms of computation, does not require time-consuming model retraining when new data comes in, and is very effective in predicting data on very sparse interaction matrices.

The Slope One method can increase the accuracy of the recommendation system in hybrid collaborative filtering (Yao et al., 2022; Zhang, 2009). The Slope One schema considers information from other users by rating the same item and other items rated by the same user. For instance, let us imagine two users (A and B) and two items. User A rates both items, whereas user B rates only one item. The Slope One technique is then used to assign a prediction value to the unrated item. The prediction value is crucial information when making a prediction. Figure 1 depicts the Slope One scenario.

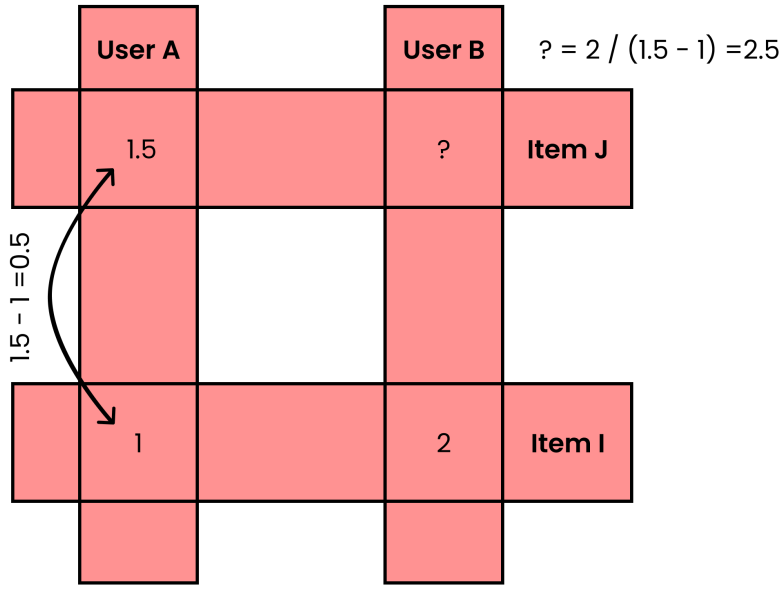

Figure 1 Base slope one schema (Lemire and Maclachlan, 2005)

The average value of the rating difference between items is determined using the Slope One algorithm technique, as represented in Equation (10) (Song and Wu, 2020):

$$\dot{x} = f(x, u)\, dev_{j,i} = \sum_{u \in S_{j,i}(x)} \frac{u_j - u_i}{\text{card}\,(S_{j,i}(X))} \tag{10}$$

The rating prediction for items without a rating can be computed using Equation (11) once the average rating difference between items has been determined:

$$p^{S1}(u)_j = \frac{\sum_{i \in S(u)-\{j\}} (dev_{j,i} + u_i)\, c_{j,i}}{\sum_{i \in S(u)-\{j\}} c_{j,i}} \tag{11}$$

In this framework, Slope One serves as the initial step to fill rating gaps before applying the HCP process.

2.7 Simple additive weighting method

The simple additive weighting (SAW) method is a popular linear combination weighting technique with an easy-to-follow procedure when handling multi-attribute decision-making (MADM) problems. The fundamental idea behind the SAW technique is the weighted sum of the value of each option across all characteristics, which first normalizes the decision matrix (X) to a scale that can be compared with all of the currently available alternative values (Loa et al., 2020).

The SAW algorithm is applied after applying the ICHM method. This is because ICHM is tasked with generating pure collaborative ranking predictions, while SAW is used to re-rank these technically accurate prediction rankings based on real-world user constraints, such as product price and actual location distance. SAW is the most appropriate ranking method for evaluating multi-criteria attributes with different dimensions (distance vs. rating vs. price) due to its straightforward and transparent approach. The steps in the SAW technique are as follows (Taherdoost, 2023):

1. Establish the relative importance of the alternatives (A) and criteria (C) that will be handled using the SAW technique.

2. Using criteria (C), create a decision matrix. Then, the following equation is used to normalize the decision matrix (R), as shown in Equation (12):

$$R_{i,j} = \begin{cases} \dfrac{X_{i,j}}{\max_i X_{i,j}}, & \text{if } j \text{ is a benefit attribute} \\ \dfrac{\min_i X_{i,j}}{X_{i,j}}, & \text{if } j \text{ is a cost attribute} \end{cases} \tag{12}$$

3. The preference value (V), which is the sum of the weight vector of the criteria (W) and the normalized product of the decision matrix (R) so as to choose the best option as a solution, as shown in Equation (13):

$$V_i = \sum_{j=1}^{n} W_j \times r_{i,j} \tag{13}$$

In this study, the SAW method is employed to synthesize multiple criteria, such as craft price, predicted rating, and distance, to identify the most suitable recommendation for each user.

2.8 Recommendation System Workflow

Figure 2 shows the structured workflow of this recommendation system, which is designed to provide the most relevant recommendations for craft locations to users.

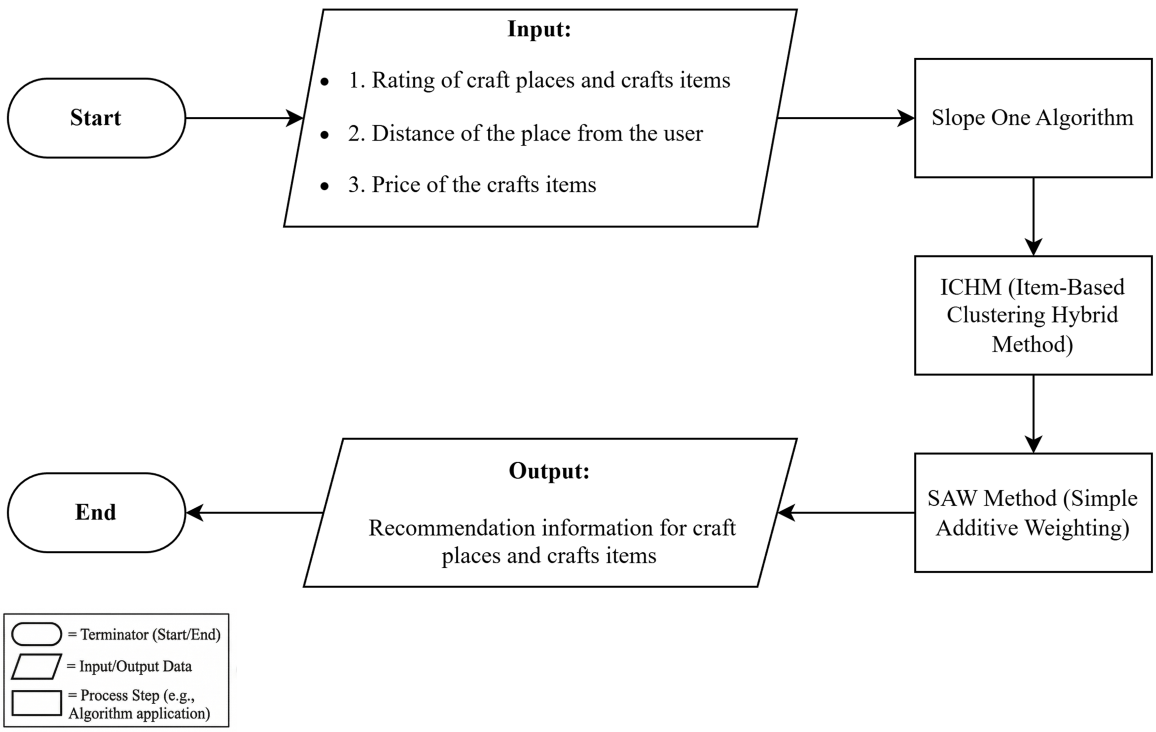

Figure 2 Workflow of traditional craft recommendation system

The proposed process begins with the collection of three main types of input data: user-provided ratings, distance between the user's location and each craft center, and product pricing information. The Slope One algorithm is used to first process the collected rating data, which serves to predict and complete the missing ratings for items that the user has not yet rated. Subsequently, the predicted rating outputs from Slope One are used as the ICHM input. As a hybrid method, the ICHM combines various approaches, such as content similarity and rating correlation, to generate a more accurate and reliable initial recommendation score. The final stage of the process involves the application of the SAW method, where the recommendation scores from the ICHM are combined with the distance and price data. The SAW method is applied to determine the final ranking by weighting these three criteria, ultimately producing a ranked list of craft location recommendations that best match the user's overall preferences.

The selection of the Slope One algorithm, followed by the ICHM, is based on the need to build a recommendation system that is efficient, accurate, and robust in addressing common challenges, such as data sparsity (a low volume of rating data) and the cold-start problem (difficulty in recommending new items). Slope One was chosen as the initial step due to its implementation simplicity and computational efficiency. This algorithm is highly effective for predicting ratings in sparse datasets, which is a common scenario in specialized recommendation systems, such as one for craft locations. Its ability to update predictions online as new ratings are added is also a significant advantage. However, Slope One alone is insufficient to handle the cold-start problem for items with no ratings at all. Here, the role of the ICHM as a hybrid method becomes crucial. The ICHM combines the strength of collaborative filtering (derived from Slope One predictions) with a content-based approach. By clustering items based on their attributes (e.g., craft type, location), the ICHM can provide recommendations even for new items, thereby overcoming the cold-start problem. This combination results in a system that considers not only item popularity (ratings) but also their intrinsic relevance, achieving higher accuracy compared with a single method.

Conceptually, the formal mathematical model of this complete hybrid system can be represented as $S_{\text{final}}(u,i) = \sum_{c=1}^{C} W_c \operatorname{Norm}(v_c)$, where the final utility S_{final} for user u on item i is a weighted aggregation of the value vector V, which includes the combined prediction P (ICHM + Slope One), spatial distance, and selling price.

All interaction data and user ratings were obtained through direct field observations. These ratings are real historical data from actual tourists who have previously visited and made direct transactions at the craft centers, not artificially simulated data. The field dataset collected in this section consists of 48 locations and 60 types of craft items. Given the distribution of art centers in the study area, this sample size has been validated and is considered representative of the region's craft population variance.

In terms of complexity analysis, this hybrid algorithm model is designed with an aggregate computational complexity of $O(|U| \cdot |I|^2)$ in the Slope One phase and an additional $O(N \cdot K \cdot I)$ for K-Means. In terms of space complexity, this model requires memory storage of $O(|U| \cdot |I|)$ to accommodate the interaction matrix. This complexity metric confirms the feasibility and scalability of the proposed algorithm within the resource constraints of mobile devices.

Finally, an experimental framework was designed that included an ablation study to objectively evaluate the reliability of this combined method. This study isolated the prediction performance testing into three separate scenarios: a model using Slope One alone, a model using ICHM alone, and a full hybrid model. This approach was taken to empirically validate the extent to which each algorithm contributes to improving the system's accuracy compared to when they work individually.

3. Results and Discussion

During the data collection phase, the literature study approach was employed to gather information from books, journals, research findings, theses, and other related works. Additionally, information on traditional Balinese handicraft locations was gathered through observation, yielding 48 craft locations and 60 craft data. The Slope One algorithm and the ICHM method were implemented using the rating information provided by users during application use. A total of 60 people provided the customer data, which generated 920 rating data points. These data provide a valuable real-world foundation for evaluating the hybrid recommendation model in a culturally specific and small-scale market environment, reflecting authentic consumer interactions with traditional crafts.

To ensure scientific rigor and testing validity, this dataset was divided using a train-test split evaluation scheme with a proportion of 80% as training data (used to build the centroid cluster model and Slope One deviation matrix) and 20% as test data. Furthermore, the system's ability to solve cold-start problems is not only demonstrated procedurally but also isolated experimentally through synthetic cold-start simulations. This simulation was conducted by deliberately masking 100% of the interaction history on a specific subset of test items to force the system to validate predictions purely based on ICHM attributes without relying on past collaborative data.

a) Prediction value

The next phase, which is to obtain the expected value in the empty rating matrix column, uses the average value of the rating difference as the input. User 1 predicted that craft places 1 would receive a rating of **3.03** in the initial iteration. Table 1 shows the predicted rating values for each empty matrix cell.

Table 1 shows that the Slope One algorithm successfully fills all the empty cells in the rating matrix. For example, the system predicts that User 1 will give Item 1 a rating of 3.03, even though User 1 has never rated it. This step is crucial for overcoming the problem of data sparsity (i.e., rare data). With this matrix filled, each item now has a base score that can be used for calculations in the next stage, namely the ICHM stage. This collaborative filtering foundation is obtained because the algorithm calculates the average rating deviation (Supplementary File - Table 1) of all other users who have rated Item 1 relative to other items that have also been rated by User 1, considering the amount of supporting data (Supplementary File - Table 2).

Table 1 Predicted rating value in the empty matrix cell (n = 10)

Users and Craft Places	$\sum_{i\in S(u)-j} (dev_{j,i} + u_i)C_{j,i}$	$\sum_{i\in S(u)-j} C_{j,i}$	Prediction $(P^{S1}(u,j))$
1,1	109.00	36	3.03
1,10	61.00	30	2.04
2,2	79.67	28	2.85
3,8	148.00	41	3.61
4,1	87.00	36	2.42
4,3	106.67	37	2.89
4,4	83.00	42	1.98
5,2	57.67	28	2.06
5,3	86.67	37	2.35
5,7	67.00	42	1.60
5,10	94.00	41	2.30
6,9	128.00	40	3.20
7,2	75.67	28	2.71
7,5	70.00	42	1.67
7,10	44.00	30	1.47

1. Item-based cluster hybrid method (ICHM)

   The ICHM is utilized as a suggestion value solution technique following the completion of the Slope One procedure by integrating the content and rating values of craft locations, which predicts the empty rating value. This integration forms the backbone of the hybrid model, combining user-driven patterns (from ratings) with content-based characteristics (i.e., price and distance). The steps involved in the ICHM are as follows:

   a) Group rating

      Group rating is the initial procedure followed, using item content inputs (price and distance) in the first step of ICHM. The k-means clustering method is used to process the input craft place content and the likelihood that each item will cluster to produce a group rating matrix.

First, normalization is applied to ensure uniform data scaling, followed by clustering using the k-means algorithm. The Elbow method determines the optimal number of clusters based on the SSE analysis. The following is the equation-based explanation.

i. Normalization of the content values of craft places
Value normalization is used to bring the imbalanced value range into line with the same range. With a lower limit value range of 1 and an upper limit value of 10, value normalization employs the Min-Max method. Equation (14) represents the normalization procedure:

$$\text{norm}(x) = \frac{(x - \min) \times (\max R - \min R)}{(\max - \min) + \min R} \quad (14)$$

The outcomes after the normalization of the content value of the items in the craft place are shown in Supplementary File (Table 4).

2. K-means clustering

By maximizing the k value using the Elbow technique and the following steps, the k-means clustering algorithm groups the content value of the craft place items that have been standardized in Supplementary File - Table 4 according to the closeness value between items. According to the Euclidean distance computation, the first craft place data became Cluster group 1 because there was the least distance between the first craft place data and Cluster 1. These are the outcomes of every clustered craft location dataset during the initial iteration. The results of the first and final iteration clustering are shown in Supplementary Files 5 and 6, respectively.

An evaluation was conducted using the elbow approach with the SSE equation, following the acquisition of the cluster findings in the preceding procedure. The method for computing the SSE is shown in Eq. (9).

The procedure of clustering craft locations with the addition of 1 cluster up to 10 clusters was repeated after determining the SSE value for two clusters. Supplementary File-Table 7 displays the SSE values for Clusters 2–10. The largest SSE value difference among the ten clustering studies was 22.53 when there were four clusters, as shown in Supplementary File–Table 8.

3. Probability of each item belonging to the cluster

As shown in Supplementary File-Table 9, the resulting group-rating values indicate strong differentiation across clusters, confirming that based on their normalized attributes, the clustering algorithm effectively distinguishes between craft items. This probabilistic grouping, which is based on the optimal clustering result of k = 4, contributes to the system's ability to make accurate hybrid recommendations even for unrated or new items.

a) Similarity item-rating
This phase utilizes the correlation-based similarity Equation (2) to obtain the similarity value between item-ratings based on the craft place ratings that were processed using the Slope One algorithm. The steps of the similarity process based on Pearson correlation are shown in Supplementary File-Table 10, using the rating data provided in Table 1.

The Slope One technique is used in the first iteration to obtain the similarity rating value between craft places 1 and 2. The similarity rating values for the ten craft locations are shown in the supplementary file (Table 11).

b) Similarity group-rating

The similarity rating value between craft places 1 and 2 is determined in the first iteration using the input group rating value already processed in Supplementary File - Table 11. Supplementary File-Table 12 Determined similarity in group ratings using cosine similarity.

c) Linear Combination of Similarity

The next procedure is a linear combination. Equation (4) can be used to create linear combinations by entering the the result of the similarity rating for each item in Supplementary File (Table 11). Similarity group-rating values are shown in Supplementary File (Table 13). Craft Place 1 and Craft Place 2 have a linear similarity combination value of 0.07 as shown in Supplementary File-Table 13. The importance of linear similarity combinations for every location of the craft is evident.

d) Prediction of cold-start problem

We used the item rating in Supplementary File-Table 11 and linear combinations of similarity to make a prediction in the event of a cold start using Equation (6). Using the crafting place of Item 1, a cold-start problem prediction value is obtained for User 1. Table 2 shows the detailed results of the cold-start problem prediction value.

Table 2 Result of the cold-start problem prediction values

User	Craft Placement Items									
	1	2	3	4	5	6	7	8	9	10
1	3.66	3.46	2.12	2.00	3.76	3.40	1.58	1.87	4.40	1.34
2	1.86	2.67	3.85	2.70	1.38	1.88	2.60	4.45	1.81	2.24
3	2.97	1.48	4.28	2.23	1.25	1.09	3.93	4.21	1.98	3.55
4	3.09	3.77	2.61	2.18	2.68	2.62	2.09	3.01	3.39	1.31
5	3.81	2.86	3.05	1.63	1.99	2.20	2.43	3.66	2.40	2.13
6	2.68	3.44	3.15	2.66	1.50	1.86	2.55	4.27	1.89	1.58
7	2.38	2.93	3.18	3.58	2.10	1.96	2.43	3.13	3.13	1.64

Table 2 shows the final prediction values generated by the ICHM, which successfully provided values for all items, including the cold-start scenario. The ability of the system to predict values (e.g., 3.66 for User 1, Item 1) even though there was no previous rating data from that user is because the ICHM intelligently combines two sources of information: item-rating similarity (Supplementary File - Table 11) and group-rating similarity (Supplementary File - Table 12) through the linear combination in Supplementary File - Table 13. However, the numerical predictions' predicted values are not directly translated into final decisions that are displayed to users. The system applies an acceptance threshold of > 3.5 on a 5-point scale. This means that only items and locations that achieve a utility score above 3.5 are technically considered "eligible" and will be recommended to users to ensure the quality of the suggestions provided.

This is a key demonstration of the hybrid model's strength; when collaborative data is absent, the system does not fail and relies more heavily on content data. The results show that prediction values are successfully generated for all unrated items, validating the hybrid model's ability to address both user-based and item-based cold-start scenarios. This capability is crucial in traditional craft domains, where many products or artisans are newly introduced to the market and have limited initial user interaction data.

4. SAW Method

Based on the recommendation values and distance values from the prior ICHM, the SAW technique was employed to determine the recommendations for craft places. The SAW method's inputs were the cold-start problem prediction values in Table 2. In this analysis, the cold-start problem prediction value data of User 1 were used to determine the cold-start problem prediction value based on the active user (login). The steps involved in the SAW method are as follows:

a) Determine the weight of criteria (C) and alternatives (A)

The weight of the criteria is based on the price of the craft item (30%), the estimated value of the cold-start problem (40%), and the distance to the craft location (30%). In addition, craft stores serve as alternatives to the SAW procedure.

These weights were not set arbitrarily but were justified through a preliminary questionnaire to respondents regarding their travel preferences. In addition, sensitivity analysis testing was conducted to prove that these weighting ratios produced the most stable recommendations and were not dominated by a single criterion.

b) Decision matrix based on criteria (C)

value of each craft place to find the maximum and minimum value of alternatives (craft places) from each criterion.

c) Normalization of decision matrix (R)

Considering that the values of each alternative on each criterion vary greatly, Equation (11) must be used to normalize the values. This is done using the normalization equation for the cost attribute after determining that the distance criterion (C_1) is a cost attribute (cost) from the equation above. The normalization equation for the profit attribute is used as the criterion (C_2) is a profit attribute. Supplementary File-Table 16 displays the outcomes of the decision matrix standardization for each craft place.

d) Preference Value (V)

A preference value is generated using Equation (12) by entering the weight value of the criteria and the decision matrix's normalization value. The preference value for craft place Item 1 in the above equation is **0.37**. The values for each craft place preference are arranged from highest to lowest. The recommended craft place is the one with the highest preference rating. With the highest preference value of **0.73** among all craft place preference values, Item 9 emerges as the optimal craft place. Additional information is presented in Table 4.

The final preference results using the SAW method (Table 4) clearly place item 9 (V = 0.73) as the top recommendation for user 1. Analysis of the input data (Table 3 and Supplementary File-Table 16) explains why item 9 wins by such a large margin. This item has the highest prediction value (C2; 4.40) and the largest weight (40%). Item 9 also has the lowest price (C3; 1000), which also has a high weight (30%). Although its distance (C1) is not the closest, its dominance in the two main criteria makes it the clear winner. This result illustrates how the hybrid framework not only predicts ratings accurately but also ranks recommendations effectively. The system successfully integrates users' personal preferences (from the C2 prediction value) with real-world constraints (price C3 and distance C1), resulting in practical and affordable recommendations that are not only liked but also practical.

An in-depth analysis of this SAW decision reveals an interesting finding: the

Table 3 Decision matrix based on the criteria

Craft Placement Items	Criteria		
	C1 = distance (m)	C2 = prediction value	C3 = craft price (IDR)
1	10,600	3.66	17,000
2	7,870	3.46	18,000
3	1,353	2.12	28,000
4	657	2.00	20,000
5	5,885	3.76	17,000
6	6,913	3.40	6,000
7	3,109	1.58	65,000
8	4,180	1.87	12,000
9	7,524	4.40	1,000
10	4,082	1.34	10,000
Max	10,600	4.40	65,000
Min	657	1.34	1,000

"distance" criterion (30%) proved to have stronger discriminatory power on the final ranking than "price" (30%). Although Item 9 is quite far away (7,524 m), its lowest price (Rp 1,000) and highest predicted ranking (4.40) made it win first place. However, in this dataset, the variance in geographical distance in Bali imposes a significant penalty on the SAW method. The system often cancels recommendations for locations with high predicted rankings if they are considered too remote or logistically inaccessible to tourists, ensuring that highly pragmatic recommendations are provided.

Table 4 Craft Place Preference Score

Craft Placement Items	Preference value (V)
1	0.37
2	0.36
3	0.35
4	0.50
5	0.39
6	0.39
7	0.21
8	0.24
9	0.73
10	0.20

The MAE equation is used in accuracy testing. The Slope One algorithm and the ICHM are used to assess the accuracy level of a recommended item (Nur Rohmah and Baizal, 2025; R. Li and Liu, 2023; Sharma and Sharma, 2019). The results of the item rating and recommendation values of 60 consumers are used to calculate the MAE equation. The MAE values fall between 0 and 1. If the MAE number is greater than or equal to 1, then the recommendations are no longer accurate. As a result of the computation, the Slamet Gamelan handicraft manufacturing location has an MAE value of 0.45. The 0 rating value is not computed because 0 is not the actual value. The resulting MAE value is the value of the item with the actual rating. Figure 3 shows the MAE value graph for the craft place.

Figure 3 MAE value of craft place

Murtika Lukisan achieved the lowest MAE score of 0.31 and the most accurate suggestion of craft place information. However, **Sumber Aneka Kreasi Bamboo** has the highest MAE value of 0.66. The graphs and specifics of craft names and MAE values are shown in Figure 4.

The lowest MAE value of 0.09 for **Gender Bali** provides the most accurate craft. However, **Gender Made** has the highest MAE value of 0.69. This low MAE result (well below the threshold of 1.00) validates the hybrid model's accuracy. This shows that the predictions generated by the system (Table 2) are very close to the users' original ratings. This low error value is due to the ICHM's ability to balance collaborative information and content, thereby not producing extreme predictions that deviate significantly and keeping the average error to a minimum. An accurate MAE number should not exceed 1.00; a lower MAE value means a more accurate prediction value and offers higher quality and more effective recommendations.

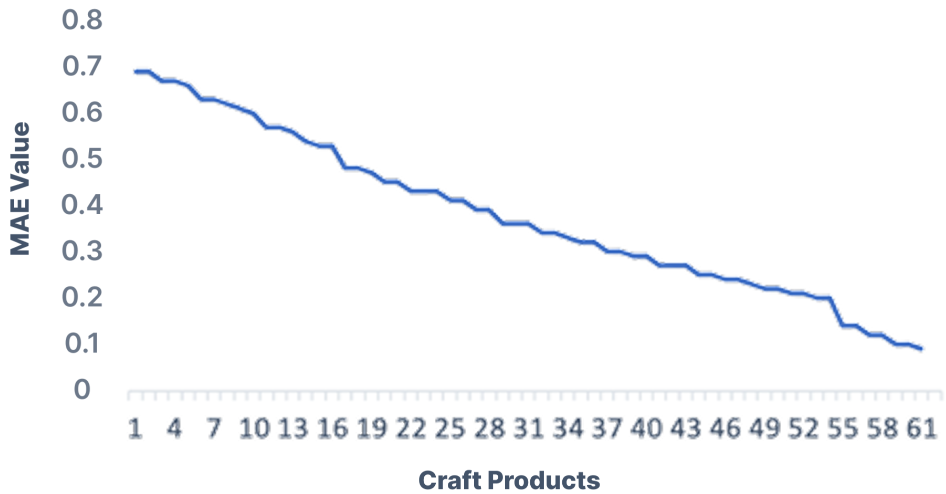

Figure 4 MAE value of craft or product item

Supplementary File – Table 18 presents the results of a comparative MAE study between the proposed model and the baseline algorithm to prove that the "Hybrid" approach is quantitatively superior.

Supplementary File–Table 18 presents a step-by-step presentation of how the hybrid model dramatically reduces the error rate to 0.38 (aggregate average). Based on a rating scale of 1 to 5, an MAE value below 1.0 academically indicates a very high level of accuracy, as the system deviation is less than one star rating (e.g., predicting 4.2 when the actual value is 5.0).

This metric's interpretation rejects misleading empirical assumptions and proves that Slope One running on dense ICHM clusters can overcome data sparsity.

Although the quantitative evaluation results show excellent performance, the proposed system has several operational limitations. First, data collection still relies on manual processes in the field. Second, the dataset size (48 locations, 60 products) represents a relatively small test scale. Evaluation on a large-scale dynamic dataset is needed.

4. Conclusions

The hybrid recommendation system, which integrates the ICHM with the Slope One and SAW algorithms, was successfully developed to provide recommendations for both craft products and production locations. The system effectively addresses the cold-start problem, enabling recommendations based on item content for new or unrated handicraft items. The system achieves a low mean absolute error value, with 0.09 for craft products and 0.31 for production locations, indicating high prediction accuracy across the dataset. This low prediction error rate is not only algorithmically crucial but also essential in building user trust. Trust in the reliability of digital recommendations is a major driver for tourists and local communities to adopt smart technology to explore and preserve the traditional craft sector. Future research should focus on scalability. If the craft ecosystem is expanded to include 10,000 products and locations, the current Hybrid Slope One + ICHM architecture, which relies on a quadratic distance matrix, may experience a degradation in computational efficiency (bottleneck). The integration of distributed matrix factorization models or deep learning techniques is highly recommended to maintain future real-time performance efficiency.

Acknowledgement

This research was supported by the BPI – Doctoral Completion Scholarship Program (Beasiswa Program Doktor Penyelesaian Studi di Dalam Negeri) 2025 funded by the Kementerian Pendidikan Tinggi, Sains, dan Teknologi, Republic of Indonesia. The authors sincerely appreciate the financial support that enabled the completion of this study BPI Number 202528112523, year 2025. We would like to thank Edy Anjasmara, Dwi Adnyana, and the team for their support in software development and datasets.

Author Contributions

The authors confirm their individual contributions as follows: I Gusti Agung Gede Arya Kadyanan was responsible for the computational programming; Ni Made Ary Esta Dewi Wirastuti contributed to the final manuscript drafting; Gede Sukadarmika and Ngurah Agus Sanjaya ER were involved in verifying the results and references; and Is-Haka Mkwawa and Muhamad Asvial were responsible for the results verification. All authors have read and approved the published version of the manuscript.

Conflict of Interest

The authors declare no conflicts of interest..

References

Afoudi, Y., Lazaar, M., & Al Achhab, M. (2021). Hybrid recommendation system combined content-based filtering and collaborative prediction using artificial neural network. *Simulation Modelling Practice and Theory*, *113*, 102375. https://doi.org/10.1016/j.simpat.2021.102375

Akbar, S., Rahmayanti, G., & Winarko, E. (2025). Clustering narrow-domain scientific text using unsupervised and similarity-based approaches. *International Journal of Technology*, *16*(5), 1467–1483. https://doi.org/10.14716/ijtech.v16i5.7110

Aljunid, M. F., & Huchaiah, M. D. (2021). An efficient hybrid recommendation model based on collaborative filtering recommender systems. *CAAI Transactions on Intelligence Technology*, *6*(4), 480–492. https://doi.org/10.1049/cit2.12048

Aparna, K., & Nair, M. K. (2016). Incorporating stability and error-based constraints for a novel partitional clustering algorithm. *International Journal of Technology*, *7*(4), 691–700. https://doi.org/10.14716/ijtech.v7i4.1579

Bhuiyan, B. U., & Khan, I. (2024). Household energy consumption clustering using k-means algorithm: A bangladesh case study. *2024 IEEE International Conference on Power, Electrical, Electronics and Industrial Applications (PEEIACON)*, 38–42. https://doi.org/10.1109/PEEIACON63629.2024.10800004

Bock, H. H. (2007). Clustering methods: A history of k-means algorithms. In *Selected contributions in data analysis and classification* (pp. 161–172). https://doi.org/10.1007/978-3-540-73560-1_15

Burke, R. (2007). Hybrid web recommender systems. In *The adaptive web* (pp. 377–408). https://doi.org/10.1007/978-3-540-72079-9_12

Chung, K. Y., Lee, D., & Kim, K. J. (2014). Categorization for grouping associative items using data mining in item-based collaborative filtering. *Multimedia Tools and Applications*, *71*(2), 889–904. https://doi.org/10.1007/s11042-011-0885-z

Di, Y., Liang, C., & Zhang, Y. (2026). Fedrl: A reinforcement learning federated recommender system for efficient communication using reinforcement selector and hypernet generator. *ACM Transactions on Recommender Systems*, *4*(1), 1–31. https://doi.org/10.1145/3682076

Fareed, A., Naveed, A., & Rashid, M. (2023). A collaborative filtering recommendation framework utilizing social networks. *Machine Learning with Applications*, *14*, 100495. https://doi.org/10.1016/j.mlwa.2023.100495

Hendrarini, N., Asvial, M., & Sari, R. F. (2022). Wireless sensor networks optimization with localization-based clustering using game theory algorithm. *International Journal of Technology*, *13*(1), 213–224. https://doi.org/10.14716/ijtech.v13i1.4850

Kannout, E., Al-Otaibi, S., & Al-Zahrani, A. (2024). Clustering-based frequent pattern mining framework for solving cold-start problem in recommender systems. *IEEE Access*, *12*, 13678–13698. https://doi.org/10.1109/ACCESS.2024.3355057

Ko, H., Lee, S., & Park, S. (2022). A survey of recommendation systems: Recommendation models, techniques, and application fields. *Electronics*, *11*(1), 141. https://doi.org/10.3390/electronics11010141

Larasati, A., Maren, R., & Wulandari, R. (2021). Utilizing elbow method for text clustering optimization in analyzing social media marketing content of indonesian e-commerce. *Jurnal Teknik Industri*, *23*(2), 111–120. https://doi.org/10.9744/jti.23.2.111-120

Larasati, A., Supriadi, H., & Wulandari, R. (2019). Profiling academic library patrons using k-means and x-means clustering. *International Journal of Technology*, *10*(8), 1567–1575. https://doi.org/10.14716/ijtech.v10i8.3440

Lemire, D., & Maclachlan, A. (2005). Slope one predictors for online rating-based collaborative filtering. *Proceedings of the 2005 SIAM International Conference on Data Mining*, 471–475. https://doi.org/10.1137/1.9781611972757.43

Li, Q., & Kim, B. M. (2003). An approach for combining content-based and collaborative filters. *Proceedings of the Sixth International Workshop on Information Retrieval with Asian Languages*, 17–24. https://doi.org/10.3115/1118935.1118938

Li, R., & Liu, Q. (2023). A weighted slope one collaborative filtering algorithm for improved user relevance. *2023 International Conference on Artificial Intelligence and Automation Control (AIAC)*, 1–4. https://doi.org/10.1109/AIAC61660.2023.00009

Loa, A., Gunawan, I., & Setiawan, A. (2020). Comparing saw and cpi method in decisions systems support to evaluate teachers performance. *bit-Tech*, *2*(3), 121–130. https://doi.org/10.32877/bt.v2i3.141

Mandalapu, S. R., Narayanan, B., & Putheti, S. (2023). A hybrid collaborative filtering mechanism for product recommendation system. *Multimedia Tools and Applications*, *83*(5), 12775–12798. https://doi.org/10.1007/s11042-023-16056-8

Nainggolan, R., Perangin-angin, R., & Tarigan, E. B. (2019). Improved the performance of the k-means cluster using the sum of squared error (sse) optimized by using the elbow method. *Journal of Physics: Conference Series*, *1361*(1), 012015. https://doi.org/10.1088/1742-6596/1361/1/012015

Nicart, E. B. B., Arellano, B. R., & Acunin, M. L. (2024). Machine algorithm-based journey assistant: An intelligent interface for tourism website. *International Journal of Technology*, *15*(3), 709–719. https://doi.org/10.14716/ijtech.v15i3.6218

Nur Rohmah, S. N., & Baizal, Z. K. A. (2025). A hybrid culinary recommender system using slope one and ichm (case study: Greater jakarta metropolitan area). *2025 International Conference on Data Science and Its Applications (ICoDSA)*, 637–642. https://doi.org/10.1109/ICoDSA67155.2025.11157062

Patel, D., Patel, F., & Chauhan, U. (2023). Recommendation systems: Types, applications, and challenges. *International Journal of Computing and Digital Systems*, *13*(1), 851–868. https://doi.org/10.12785/ijcds/130168

Praditya, N. W. P. Y., Budhi, G. S., & Adipranata, R. (2022). Collaborative and content-based filtering hybrid method on tourism recommender system to promote less explored areas. *International Journal of Applied Engineering and Technology*, *4*(2), 59–65. https://doi.org/10.5281/zenodo.7404490

Saeed, M., & Mansoori, E. G. (2018). A new slope one based recommendation algorithm using virtual predictive items. *Journal of Intelligent Information Systems*, *50*(3), 527–547. https://doi.org/10.1007/s10844-017-0470-7

Sari, M., & Armansyah. (2024). Manhattan metric in k-means clustering for data grouping. *Journal of Information Systems and Informatics*, *6*(3), 1152–1164. https://doi.org/10.51519/journalisi.v6i3.841

Sarwar, B., Karypis, G., & Konstan, J. (2001). Item-based collaborative filtering recommendation algorithms. *Proceedings of the 10th International Conference on World Wide Web*, 285–295. https://doi.org/10.1145/371920.372071

Sharma, R., & Sharma, M. (2019). Improved cf based prediction technique for recommendation systems. *International Journal of Advanced Trends in Computer Science and Engineering*, *6*(55), 342–350. https://doi.org/10.19101/IJATEE.2019.650038

Shi, C., Zhao, B., & Wang, Y. (2021). A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *EURASIP Journal on Wireless Communications and Networking*, *2021*(1). https://doi.org/10.1186/s13638-021-01910-w

Song, Y. T., & Wu, S. (2020). Slope one recommendation algorithm based on user clustering and scoring preferences. *Procedia Computer Science*, *166*, 539–545. https://doi.org/10.1016/j.procs.2020.02.042

Taherdoost, H. (2023). Analysis of simple additive weighting method (saw) as a multiattribute decision-making technique: A step-by-step guide. *Journal of Management Science and Engineering Research*, *6*(1), 21–24. https://doi.org/10.30564/jmser.v6i1.5400

Tarigan, D. A. (2023). Optimization of the k-means clustering algorithm using davies bouldin index in iris data classification. *Kajian Ilmiah Informatika dan Komputer*, *4*(1), 545–552. https://doi.org/10.30865/klik.v4i1.964

Tran, N. T. (2025). *Disentangling user preferences towards self-interpretable recommender systems* [Doctoral dissertation, Singapore Management University]. https://ink.library.smu.edu.sg/etd_coll

Widayanti, R., Santoso, S., & Haryanto, A. (2023). Improving recommender systems using hybrid techniques of collaborative filtering and content-based filtering. *Journal of Applied Data Sciences*, *4*(3), 289–302. https://doi.org/10.47738/jads.v4i3.115

Yannam, V. R., Satapathy, S. C., & Zhang, Y. D. (2023). Enhancing the accuracy of group recommendation using slope one. *The Journal of Supercomputing*, *79*(1), 499–540. https://doi.org/10.1007/s11227-022-04664-4

Yao, D., Deng, X., & Qing, X. (2022). A course teacher recommendation method based on an improved weighted bipartite graph and slope one. *IEEE Access*, *10*, 129763–129780. https://doi.org/10.1109/ACCESS.2022.3228957

Zhang, D. (2009). An item-based collaborative filtering recommendation algorithm using slope one scheme smoothing. *2009 Second International Symposium on Electronic Commerce and Security*, 215–217. https://doi.org/10.1109/ISECS.2009.173