

A Weighted Hybrid Filtering Approach for Cold-Start and Diversity Evaluation in Culinary Recommender Systems

Nur Lukman^{1*}, Fiki Aprian², Khaerul Manaf³, Wildan Budiawan Zulfikar⁴,
Rifqi Syamsul Fuadi⁵

^{1,4,5}Department of Informatics, Faculty of Science and Technology, UIN Sunan Gunung Djati
Bandung, Indonesia

²PT. Sasana Solusi Digital, Indonesia

³Department of Information System Engineering, Sangga Buana University, Indonesia

ABSTRACT

Open Access

Type: Research Paper

Article:

Received: February 12, 2026

Revised: March 18, 2026

Accepted: August 13, 2026

Issued: October 30, 2026

© Lukman et al. (2026).



This is an open-access article
under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

*Correspondence Address:

n.lukman@uinsgd.ac.id

Recommender systems play a crucial role in helping users discover relevant and diverse products tailored to their preferences. However, challenges such as the cold-start problem and low recommendation diversity still persist, particularly in single-approach models such as Content-Based Filtering (CBF) and Collaborative Filtering (CF). This study implements and evaluates a Weighted Hybrid Filtering approach that linearly combines prediction scores from CBF, which recommends items based on content similarity to a user's preference profile, and CF, which recommends items based on rating patterns from similar users, using a weighting parameter α . The model was evaluated on a real-world transaction dataset from Tantri Point of Sale (POS) consisting of more than 300,000 user-item interactions, with a user-item interaction sparsity of 99.93%, using Precision@10, Recall@10, F1@10, Intra-List Diversity (ILD@10), RMSE, and MAE. Cold-start conditions were assessed through simulation by masking user interaction histories. Results show that the hybrid configuration with $\alpha = 0.9$ achieves the highest Recall@10 (0.8153) among all evaluated models, along with an F1@10 of 0.6634, an ILD@10 of 0.8030, an RMSE of 0.2984, and an MAE of 0.1924, although the standalone CF model attains higher F1@10 and ILD@10, and the standalone CBF model attains higher precision, individually. Compared to existing weighted hybrid recommender approaches, which have primarily been evaluated on non-culinary benchmark datasets, this study applies and evaluates the method on a large-scale, real-world culinary transaction dataset, providing an explicit assessment of the accuracy-diversity trade-off across multiple weighting configurations. These findings provide empirical insight into the trade-offs of weighted hybrid filtering for cold-start and diversity assessment in culinary recommender systems.

Keywords: Cold start, Collaborative filtering, Content-based filtering, Diversity, Hybrid filtering, Recommender system

1. INTRODUCTION

The growth of digital culinary platforms, including online food delivery and restaurant discovery applications, has generated large volumes of transactional and behavioral data [1], [2], [3]. Although this trend provides for enhanced accessibility and customer convenience, information overflow occurs at the same time, thus making customers unable to identify appropriate restaurants and culinary products that satisfy their needs and interests. In this situation, recommender systems prove to be highly beneficial since they help filter and personalize data according to customers' preferences.

Collaborative Filtering (CF), particularly matrix factorization methods such as Singular Value Decomposition (SVD), is widely used because it captures latent user preferences from historical interactions [4], [5], [6], [7], [8]. However, CF depends heavily on the availability of user-item interaction history. It performs poorly for new users or new items with no prior interactions (cold-start), and its accuracy degrades as the interaction matrix becomes sparse [9], [10], [11], [12].

Content-Based Filtering (CBF) addresses the cold-start limitation of CF by generating recommendations from item content, such as descriptions and category labels, without requiring interaction history [13], [14]. However, because CBF relies solely on content similarity, it tends to recommend items that are highly similar to a user's past preferences, resulting in lower recommendation diversity and limited exposure to new items [13], [15], [16].

Hybrid recommender systems combine CF and CBF to leverage the strengths of both approaches, and have been widely studied and reviewed in the literature [17], [18], [19], [20], including hybrid architectures that incorporate multiple data sources such as probabilistic graphical models [21]. Among hybridization strategies, weighted linear combination remains one of the most

commonly used due to its simplicity and interpretability. However, prior studies applying this technique have largely been developed and evaluated on non-culinary benchmark datasets such as MovieLens. For instance, existing work has implemented weighted linear combinations of content-based and item-based collaborative filtering scores on movie rating data, finding that the hybrid combination did not substantially improve accuracy over the individual methods, although it enabled predictions for items with no prior ratings [22]. Similarly, other studies have proposed weighted combinations of user-based and item-based collaborative filtering, or of multiple matrix factorization variants such as SVD and SVD++, primarily evaluated using RMSE and MAE on MovieLens datasets [23], [24]. These studies commonly emphasize prediction accuracy but provide limited evaluation of recommendation diversity, and none evaluate weighted hybrid filtering on large-scale, real-world culinary transaction data. This gap motivates the need to implement and empirically evaluate a weighted hybrid filtering approach specifically on real-world culinary transaction data, with explicit attention to both accuracy and diversity [25].

This study implements and evaluates a weighted hybrid filtering approach for culinary recommendation, combining a TF-IDF-based CBF model and an SVD-based CF model [7], [26], [27]. Prediction scores from both models are normalized and linearly combined using a fixed weighting parameter α , without an adaptive or diversity-optimization mechanism. This approach is evaluated using a real-world point-of-sale culinary transaction dataset containing more than 300,000 user-item interactions. Model performance is measured using Precision@10, Recall@10, F1@10, RMSE, MAE, and Intra-List Diversity (ILD@10), and cold-start behavior is assessed through simulation by masking user interaction histories.

The objective of this study is to implement and empirically evaluate a weighted hybrid filtering approach that combines CBF and CF for culinary recommendation, and to analyze its performance and trade-offs between accuracy and diversity, as well as its behavior under simulated cold-start conditions, using a large-scale real-world culinary transaction dataset.

The contributions of this study are as follows. First, this study implements a weighted hybrid filtering approach that linearly combines content-based (TF-IDF) and collaborative (SVD) recommendation scores for the culinary domain. Second, this approach is evaluated on a large-scale real-world culinary transaction dataset containing more than 300,000 user item interactions, with a user item interaction sparsity of 99.93%, reflecting the highly sparse nature of real-world culinary transaction data. Third, this study examines the behavior of the hybrid model under simulated cold-start conditions through controlled interaction-masking examples, providing preliminary insight into its performance for users and items with limited interaction history. Fourth, this study analyzes the trade-off between recommendation accuracy and diversity across different weighting configurations (α), including a comparison against the standalone CF and CBF models.

2. METHODS

This research employs the Cross-Industry Standard Process for Data Mining framework as the overarching methodological framework, encompassing problem diagnosis, dataset examination, data preprocessing, weighted hybrid filtering modeling, and evaluation [3], [28], [29], [30]. This study addresses the cold-start and low-diversity challenges outlined in the Introduction by implementing and evaluating a weighted hybrid filtering approach that linearly combines content-based and collaborative signals through a fixed weighting parameter, α .

2.1 Dataset Description

The dataset used in the study was obtained from Tantri Point of Sale (POS), an application that supports cafe or restaurant operations on a daily basis. The system records all transaction events and customer interactions, serving as the primary data source for constructing the recommendation engine. The dataset comprises five entities: Cafe Data, Product Data, Category Data, Product–Category Relationships, and Order Data. Table 1 summarizes the structure of the raw dataset prior to feature selection.

Table 1. Dataset Structure

Data Resource	Attribute	Format	Column Count	Row Count
Cafe Data	id, xenditAccountId, name, slug, address, addressCityProvince, optionalAddress, location, schedule, logo, description, notes, paymentMethod, sequence, tax, serviceFee, isReservation, createdAt, updatedAt, actualDeletedAt, deletedAt, isOpenBillByCustomer, isPlatformFeeByCustomer, phoneNumber, isIntegratedInventory, tzName, tzTime, email, orderType, status, typeBillingStackLog, credit, assignmentBy, isIntegratedTableManagement, scheduleStatusActiveAt, bannedReason, logoStruk, countUsedPreparationLimitOrder	CSV	38	161
Product Data	id, name, calorie, price, stock, quantity, description, type, isInStock, isHaveVariant, isActivated, createdAt, updatedAt, idKafe, isIntegratedStock, isMultiPrice, availableEndClock, availableStartClock	CSV	18	34.841
Data Category	id, name, idx, createdAt, idKafe	CSV	5	4.901
Product and Category Relationships	idx, idProduk, idCustomKategori	CSV	3	32.779
Order Data	id, idCustomer	CSV	2	100257

Data Resource	Attribute	Format	Column Count	Row Count
Order Product Data	Id, idProduk, idOrder	CSV	3	313369

Using the raw dataset explained in Table 1, not all features were considered for the recommendation model. For example, interaction data were created by combining Order Data, Order Product Data, Product Data and Cafe Data such that explicit relations were made between users (idCustomer) and the products they purchased, enriched with contextual product information [5], [27], [31]. The resulting interaction data comprises six attributes: idProduk, idCustomer, produkName, idKafe, nameKafe, and the interaction score, where each row represents a single instance of interaction between a user and a product. The CBF model use name, description, category, and idKafe features from Product Data and id, name from Cafe Data (redefined as idCafe and nameCafe) were used for content-based item features.

2.2 Data Preprocessing

In this section, a series of systematic steps is outlined to prepare the data for the recommendation system models, for both the Content-Based Filtering (CBF) and Collaborative Filtering (CF) approaches.

- Data Preparation for Content-Based Filtering

Preprocessing in Content-Based Filtering is conducted to prepare textual data so that it can be utilized in the similarity calculation process between culinary products. The data used includes attributes from the product data (select_id, name, description, category, idKafe) where the "id" attribute from the cafe data is changed to idKafe and the "name" attribute is changed to nameKafe to clarify identity as well as several other supporting attributes related to item characteristics.

The first stage performed is merging multiple product attributes into a single text column (combined features). This merging aims to enable the system to obtain more comprehensive information regarding the characteristics of each culinary product. All textual content is in Bahasa Indonesia. Following the data merging process, text

cleaning is carried out to reduce noise in the data; this process consists of converting all text to lowercase, removing punctuation and special characters, and removing common Indonesian stopwords using the stopwords list from the NLTK library [32], [33].

	id	combined_features
0	71a07cd-5614-428f-a579-b5f1ed80db40	syrup manggis syrup buah manggis produk minuman kemasan
1	6604c0b-3c69-488f-a859-6e00cf9a6d5	nugget nugget snack
2	ee1a45d5-d77a-4747-b136-9e566947bc06	paket sahur nasi ayam lioni lemon squash special ramadhan
3	c4a3c0bc-7885-4b03-82af-1883a29ed96f	paket sahur nasi goreng kambing blue ocean special ramadhan
4	897a7e83-83f3-41a4-9df8-018e1e7d3310	paket sahur ayam sarundeng blue ocean special ramadhan

Figure 1. CBF data preparation results

The next step, after the text is cleaned, is to represent it numerically using scikit-learn's TfidfVectorizer with its default parameters (unigram tokenization, no maximum feature limit, and default document frequency thresholds). This results in a $33,262 \times 10,073$ TF-IDF matrix, where the rows correspond to the 33,262 products and the columns correspond to the 10,073 unique terms.

- Data Preparation for Collaborative Filtering

Preprocessing for CF Preprocessing for CF aims to process the interaction between a user and an item given the POS transaction data recorded between users and items. Since the available data is the implicit feedback and does not explicitly provide a rating, user and item interaction is observed when two interactions occur and recorded not when given a rating. Multi interactions between a user and a item are summarized, and users and items with a small number of interactions are removed. Thus, the sparsity in the user-item interaction matrix is decreased, which the output can be used as input for the SVD-based Collaborative Filtering model.

- Data Preparation for Weighted Hybrid Filtering

For the hybrid model, we combined the CBF and CF outputs by aligning user-item pairs based on idCustomer and idProduct. This resulted in a single data frame (merged_scores) for hybrid predictions, which included both types of scores. Both sets of scores were normalized using min-max to a scale of [0,1] for the hybrid model, as otherwise, one model would be weighted more heavily than the other due to its initially different value range. The data used was real customer interaction data to

generate the ground truth, selecting only positive user-item interaction pairs. Inverse term frequency (inverse document frequency) was generated from product content to calculate diversity similarity and visualized in the form of a similarity heatmap.

2.3 Modelling

The Modeling phase of a recommendation system development is the core stage at which Algorithms and models are created for processing these processed data and generate relevant recommendations. The next section describes the model building step of both base techniques applied in this study: Content-Based Filtering (CBF) and Collaborative Filtering (CF). Second, Content-Based Filtering (CBF) recommendations based on other content-based similarities between items in terms of the kinds of content contained at each item. To address this, We present the CBF model which trains to recommend food items that have similarity in certain attributes with either past user preferences or for a particular query. The workflow Figure 2 illustrates the CBF model construction process in detail.

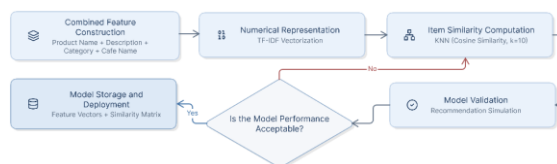


Figure 2. CBF Modeling Diagram

Figure 2 shows the procedure flow of the CBF model where the item similarity is calculated based on the TF-IDF matrix obtained in the data preparation step. CBF works in an item-based approach, this method does not produce an explicit user profile vector. For each product a user has previously purchased, the K-Nearest Neighbors (KNN) algorithm identifies the k most similar products, using cosine similarity in the TF-IDF vector space (n_neighbors = 10, metric = cosine). Cosine similarity is not an extra computation apart from KNN, it is simply the distance measure that KNN uses internally to find the nearest neighbors for each item. The trained KNN model is saved for future use, so retraining is not required. The generated sets of similar items are summed up over the purchase history of a user. This forms the CBF recommendation list of the user.

Feedback loop as shown in Figure 2. The validity of the model is checked using simulations of the recommendation output for relevance of generated items. In case the model performance is not satisfactory with respect to the pre-defined evaluation criteria, the similarity computation stage or the model parameters are tuned. When an acceptable performance is reached, the trained model and the obtained feature vectors and similarity matrix are saved and incorporated in the recommendation system..

The Collaborative Filtering (CF) approach aims to recommend items based on user-item interaction patterns [6], [33]. In this study, the Singular Value Decomposition (SVD) algorithm is used to solve matrix factorization-based recommendation problems, learning latent patterns of user-item interactions to predict preferences. The workflow for constructing the CF model is depicted in detail in Figure 3.

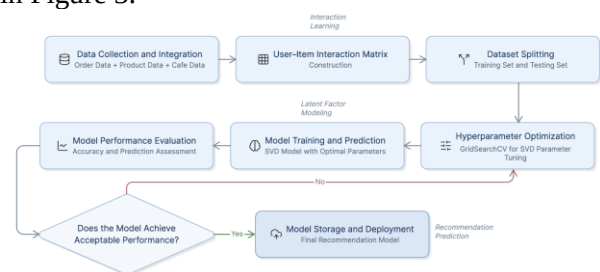


Figure 3. CF Modeling Diagram

The implicit feedback data set does not contain explicit ratings but the interactions are represented as pseudo-ratings with 1 for observed purchases and 0 for non-interactions by negative sampling, so that the model can be trained. In this context, RMSE and MAE are not measures of deviation from a subjective preference rating, but rather measures of how well the SVD model reconstructs the binary interaction labels. This serves as an internal indicator of the model's ability to distinguish likely from unlikely user-item interactions, rather than as a measure of conventional rating-prediction accuracy [34].

Based on the schema presented in Figure 3, the Collaborative Filtering methodology begins with the acquisition and aggregation of data from diverse sources, specifically encompassing order, product, and cafe datasets [13], [33]. We then use the consolidated data to create a user-item interaction matrix reflecting associations between users and food items as extracted from transaction history. Implicit

feedback in terms of this investigation refers to user engagement where a purchase action is perceived as voting for an item. After preparing the interaction matrix, dataset is split in an 80:20 ratio into training and testing data to train and evaluate the model. Then, we perform hyperparameter tuning with GridSearchCV on the SVD algorithm to guarantee a well-optimized parameter configuration. Tuning parameters: Latent factors (n_factors), learning rate (lr_all); regularization parameter reg_all; number of training iteration(n_epochs) The optimization process is carried out based on a 3-fold cross-validation environment by utilizing Root Mean Squared Error (RMSE) as the evaluation criterion. When the best parameters are found, training data is used to train the SVD model which will be then use that models in predicting user-item interaction on not yet seen data. Lastly, the predicted results are evaluated with respect to corresponding test data in order to assess how well our model has managed to generate recommendations. This process of parameter optimization is repeated with a new sample until optimal results are achieved (within the predetermined limits). When the desired model performance is achieved, we kept this CF model for further integration with culinary recommendation system.

After the recommendation scores from Content-Based Filtering (CBF) and Collaborative Filtering (CF) are successfully obtained separately, the next step is to integrate them into the Weighted Hybrid Filtering model. This approach aims to combine the strengths of both individual methods to overcome their respective weaknesses, such as the cold start problem in CF and low diversity in CBF.

The initial stage of the hybrid process begins by loading the recommendation scores generated by the CBF and CF models. Both scores are then normalized into a range of 0 to 1 to ensure a comparable scale [15]. This normalization process is crucial to prevent the dominance of either method due to differences in the range of the generated score values. After normalization is performed, an optimal weight search process (alpha tuning) is applied to determine the contribution proportion of each method within the hybrid model. The alpha value (α) represents the contribution weight of CBF, while the value $(1 - \alpha)$ represents the contribution of CF. In this study, testing is conducted iteratively with alpha values ranging from 0.1 to 0.9. The hybrid score is then

calculated using a linear combination equation between the CBF and CF scores:

$$score_{hybrid} = \alpha \times score_{CBF} + (1 - \alpha) \times score_{CF} \quad (1)$$

Via this formula, the system can reduce or increase contentbased amount of information to devise suggestions and consumer behaviour forms. Now, multiple static and dynamic performance metrics: RMSE for regression based classification ROI estimation as well as Precision, Recall and F1-score are used against each model configuration applied on combination of alpha values to identify the best-performing model. Evaluations are performed to verify that the model successfully addresses cold start and enhances recommendation diversity. Then the best value of alpha is selected as an optimal parameter in the Weighted Hybrid Filtering model. The hybrid model that obtained the best results is saved and finally, it will be used as our last proposed culinary recommendation system.

3. RESULTS AND DISCUSSION

3.1 Evaluation of Content-Based Filtering Recommendation System

The evaluation of the Content-Based Filtering (CBF)-based recommendation system aims to assess the relevance and accuracy of the recommendations generated based on the content characteristics of products preferred by users. In this study, the evaluation was conducted using two primary approaches: accuracy evaluation employing Precision@K, Recall@K, and F1-score@K metrics, and diversity evaluation using the Intra-List Diversity (ILD@K) metric.

- Evaluate Accuracy with Precision@K, Recall@K, dan F1@K

The first stage of the evaluation process involved comparing the recommendations generated by the system against a ground truth set of relevant products. The ground truth was constructed programmatically using a keyword-based relevance mapping approach. Specifically, all unique keywords were extracted from product names by tokenizing each name and retaining only words longer than three characters to filter out non-informative short terms (e.g., articles, conjunctions). For each extracted keyword, all products whose names contain that keyword were identified and

grouped as the relevant product set (ground truth) for that keyword query. For example, the keyword “coffee” would map to all product IDs containing “coffee” in their names. This approach produces a deterministic and reproducible ground truth that reflects content-level semantic relevance without requiring manual annotation. The system then generated recommendation lists for each keyword query based on content similarity using TF-IDF vectorization and cosine similarity measures, and the resulting recommendations were evaluated against the corresponding ground truth sets using three key performance metrics shown in Table 2.

Table 2. CBF Evaluation

Metrics	Average Value	Description
Precision@10	0.8187	Measures the proportion of relevant products among the top 10 recommendations
Recall@10	0.7597	Measures the system’s ability to retrieve all relevant products
F1@10	0.7246	Represents the harmonic mean of precision and recall

- Evaluate the Diversity of Recommendations with ILD@10

In addition to recommendation accuracy, recommendation diversity is an important factor in enhancing user experience. Diversity evaluation was conducted using the Intra-List Diversity (ILD) metric, which measures the degree of dissimilarity among items within a recommendation list. ILD was calculated based on the average pairwise dissimilarity (1–cosine similarity) between the TF-IDF vectors of the recommended products.

The resulting average ILD@10 value was 0.2819. This relatively low ILD@10 score indicates that the recommended products tend to be homogeneous or highly similar to one another. These findings suggest that the recommendation system places greater emphasis on content similarity while providing limited support for diverse product exploration. In recommendation system environments, insufficient diversity may reduce the system’s ability to introduce users to novel or distinct products.

- Cold Start Simulation for New Products and Users

To evaluate the system’s capability in handling the cold-start problem, a simulation involving the addition of new items without prior interaction history was conducted, as presented in Table 3. Three new culinary products were added to the dataset using the `add_product_simulation` (df) function, which simultaneously performed text preprocessing on the newly added product content.

Table 3. CBF cold start simulation

Product ID	Product Name	Description	Category	Café Name
1234567	Mangosteen Syrup	Fresh and healthy mangosteen syrup	Packaged beverage products	Cafe Sehat
1234568	Passion Fruit Ice	Authentic passion fruit ice drink with little sugar	Cold fruit drinks	Cafe Sehat
1234569	Fresh Pineapple Juice	Fresh pineapple juice with no added sugar	Natural healthy drinks	Cafe Sehat

The CBF model was subsequently retrained to incorporate the newly added products through their content representations. The evaluation was conducted under a user-side cold-start scenario by providing preference input in descriptive form, namely “healthy and refreshing beverages”. The system then generated recommendations based on content similarity between the input query and product representations, as presented in Table 4.

Table 4. Cold start simulation results

Rating	Product Name	Similarity Score
1	Mangosteen Syrup	0.4713
2	Healthy Juice	0.4140
3	Small Orange	0.3732
4	Fresh Tea	0.3657
5	Fresh Pineapple Juice	0.3151

These results demonstrate the CBF approach can generate content-based recommendations for newly added products even without interaction history. However, the evaluation was performed on only three new products and one preference query. The results

should therefore be taken as a limited illustrative simulation, rather than conclusive evidence that the method solves the cold-start problem in general.

3.2 Evaluation of Collaborative Filtering-Based Recommendation System

The Recommendation System is looked at in three ways. First we look at how accurate the predictions are, Then we look at the quality of the recommendations. Lastly we look at the variety of items in the list of things that the CBF Recommendation System suggests. In this study we use the Singular Value Decomposition approach to look at how users interact with items and make suggestions. The Collaborative Filtering-Based Recommendation System uses this method to make recommendations that're useful, to users. The Singular Value Decomposition approach was chosen to improve the Collaborative Filtering-Based Recommendation System's performance.

- Evaluation of Accuracy and Quality of Recommendations

Prediction accuracy is measured using two common error metrics, namely Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Both of these metrics measure the extent to which the system's predictions are close to actual interaction values, with RMSE being more sensitive to large errors, while MAE showing absolute average errors. Meanwhile, the quality of the recommendation list was measured using Precision@10, Recall@10, and F1@10. Precision@10 measure the proportion of relevant items out of 10 recommended items, Recall@10 measure the scope of relevant items that were successfully found, and F1@10 draw a balance between the two.

In addition to accuracy and relevance aspects, the evaluation also considers the diversity of the recommended items using the Intra-List Diversity (ILD@10) metric. This metric measures the average of the disparity between items in a recommendation list based on the cosine distance from the TF-IDF representation of the product content. A high ILD value indicates that the recommendations given are not homogeneous or too uniform, the results of the evaluation are presented in Table 5.

Table 5. CF Evaluation Results

Metric	Value
RMSE	0.2984
MAE	0.1924
Precision@10	0.7974
Recall@10	0.8021
F1@10	0.7824
ILD@10	0.9215

Based on the table above, the low RMSE and MAE values indicate that the system is able to predict user interactions with a small error rate. A high Precision@10 and Recall@10 value indicate that the system can generate relevant recommendations and include items that match the user's preferences. In addition, the ILD@10 value of 0.9215 indicates that the list of recommendations generated is quite diverse, thus preventing users from being saturated due to recommendations that are too uniform.

- Cold Start Simulation

The cold start simulation was carried out to test the system's ability to provide recommendations to new users who do not have a history of interaction, the simulation results are shown in figure 4. In this trial, new users are put into the system without explicit data, and then given recommendations by the SVD model.

```

Simulasi Rekomendasi untuk User Cold Start: user_coldstart_001

Top Rekomendasi:
1. Dora Donburi (ID: b4745df4-6a6c-4753-a34c-84074c568fb1) - Score: 1.0000
2. Kopi Lu (ID: 47c20276-1237-4f62-ab3b-6d7d6808059d) - Score: 1.0000
3. Dinsum Ayam (ID: a81e480a-448c-4075-9917-765f18ed2623) - Score: 1.0000
4. Nasi Goreng Hongkong (ID: 78a2cb56-be03-42aa-9e33-0ffdf2d533cf2) - Score: 1.0000
5. Kwetiau Goreng Hongkong (ID: f90fbc04-fa45-44c6-8d55-0d858c99f6e4) - Score: 1.0000
6. Bakmi Jawa (ID: 89ad25b2-63b9-4007-a16c-7d27cc3fa7e7) - Score: 1.0000
7. Pizza Margherita (ID: 651b4b12-3eab-4f13-905f-3409b991e29e) - Score: 1.0000
8. BALI IRI (ID: ef54a31b-dbb4-4056-a50b-f555a5f26080) - Score: 1.0000
9. SET LUNCH B (ID: 89e0e536-c9ba-4b19-97df-78bafadd653b) - Score: 1.0000
10. lemon tea hot (ID: 715dd588-eea2-46e6-9ac5-4765997b3dc7) - Score: 1.0000
    
```

Figure 4. CF Cold Start Simulation Results

The results show that the model still produces recommendations, but the entire prediction score has a maximum value (1.0000). This is due to models that rely on global estimates and item bias in the absence of user data. From this simulation, it can be concluded that the SVD-based Collaborative Filtering approach has not been able to handle cold start optimally. Therefore, a hybrid approach is needed that can leverage product content information to provide more accurate and personalized initial recommendations.

3.3 Evaluation of Hybrid Filtering-Based Recommendation System

The evaluation of the Hybrid Filtering model aims to measure the extent to which the system is able to produce recommendations that are not only accurate and relevant, but also diverse for users. The hybrid approach in this study combines the two main techniques of Content-Based Filtering (CBF) and Collaborative Filtering (CF) by utilizing the α (alpha) parameter as a weight-control of each method's contribution to the final score.

- Evaluate with Precision@10, Recall@10, F1@10, and ILD@10

Four main metrics are used to assess the performance of the system, Precision@10 i.e. the proportion of relevant items from the top 10 recommendations, Recall@10 i.e. the proportion of relevant items that are successfully recommended from all relevant items, F1@10 i.e. the harmonic average of precision and recall which describes the balance of the two, ILD@10 (Intra-List Diversity) i.e. the average of the disparity between items in the list of recommendations, calculated using the cosine distance of the TF-IDF vector.

Evaluations were carried out on various alpha values from 0.1 to 0.9. A high alpha value gives greater weight to CBF, while a low alpha value gives greater weight to CF, as defined in Equation (1). The results of the evaluation are shown in the following Table 6:

Table 6. Hybrid Filtering Evaluation Results with Alpha Variations

Alpha	Precision@10	Recall@10	F1@10	ILD@10
0.1	0.6033	0.8037	0.6538	0.8008
0.3	0.6103	0.8134	0.6616	0.8017
0.5	0.6114	0.8146	0.6628	0.8025
0.7	0.6114	0.8146	0.6628	0.8029
0.9	0.6121	0.8153	0.6634	0.8030

Based on the table above, $\alpha = 0.9$ was selected as the optimal parameter because it achieves the highest F1@10 (0.6634) among all tested configurations, reflecting the best overall balance between precision and recall. Notably, $\alpha = 0.9$ also yields the highest ILD@10 (0.8030) among the tested values, indicating that this configuration does not trade off diversity for accuracy. This means that for this dataset, the

increase of CBF weight does not influence the diversity of the recommendations, resulting in an accurate and diverse system.

- Simulasi Cold Start

To test the system's resilience to cold start user conditions, a simulation was carried out by adding new users to the interaction dataset, without having a purchase history or rating. These users are then asked to enter their initial preferences in the form of text such as "geprek chicken, milk coffee, noodles", which reflect their culinary interests.

Table 7. Example of Cold Start Hybrid Filtering Simulation Results

Rank	Product Name	CBF Score	CF Score	Hybrid Score
1	Sweet Spicy Noodles	0.7642	0.0000	0.6878
2	Chocolate Milkshake	0.7300	0.0000	0.6570
3	Hot Milk Coffee	0.6900	0.0000	0.6210
4	Iced Palm Sugar Coffee	0.6590	0.0000	0.5931
5	Spicy Crispy Chicken (Level 5)	0.6420	0.0000	0.5778

The results are shown in table 7, even if the CF score is zero, the model can still give meaningful recommendations, since the hybrid score becomes $\alpha \times \text{score_CBF}$ under this cold-start condition. The top products are the ones with the most CBF similarity to the user's stated preferences, such as Sweet Spicy Noodles and Chocolate Milkshake. This demonstrates that the hybrid model can produce reasonable initial recommendations with content-based information in the absence of a collaborative signal. This is not a full cold-start evaluation but a simplified simulation.

3.4 Comparison of Recommendation Model Performance

As part of the system evaluation process, a comprehensive comparison was conducted among the three recommendation approaches developed in this study, namely Content-Based Filtering (CBF), Collaborative Filtering (CF) using the Singular Value Decomposition (SVD) method, and Weighted Hybrid Filtering. The purpose of this comparison was to analyze the strengths and limitations of each approach

based on standard evaluation metrics and to identify the most effective method for real-world recommendation scenarios. The evaluation results for each recommendation model are presented in Table 8.

Table 8. Performance Comparison Results of the Three Recommendation Models

Recommendation Model	Precision @10	Recall @10	F1@10	ILD@10
Content-Based Filtering	0.8187	0.7597	0.7246	0.2819
Collaborative Filtering	0.7974	0.8021	0.7824	0.9215
Weighted Hybrid Filtering ($\alpha=0.9$)	0.6121	0.8153	0.6634	0.8030

Table 8 demonstrates that each recommendation approach exhibits distinct strengths and characteristics:

- The Content-Based Filtering model showed the highest precision (0.8187), indicating good relevance based on the similarity of content. However, the low diversity (ILD@10 = 0.2819) indicates that the recommended items are very similar to each other.
- The Collaborative Filtering model has achieved the best F1@10 (0.7824) and the best diversity (ILD@10 = 0.9215). However, it is still vulnerable to the cold-start problem for new users or new items.
- The Weighted Hybrid Filtering model ($\alpha = 0.9$) does not get the best precision, F1@10 and ILD@10 of the three approaches. However, it has the best Recall@10 (0.8153) which is better than both CBF (0.7597) and CF (0.8021). The hybrid model is not better than the individual models in terms of precision and F1@10, but it provides much higher diversity than CBF alone (ILD@10 = 0.8030 vs. 0.2819) and integrates content-based information to solve the cold-start problem of CF. The results show that the hybrid model is not always better than the individual methods, but provides a practical trade-off especially in recall and diversity compared to CBF, in the cases with limited interaction history.

CONCLUSION

In this research, Weighted Hybrid Filtering was implemented and evaluated to combine linearly Content Based Filtering (CBF) and Collaborative Filtering (CF) for culinary recommendation. Based on empirical evaluation using a real-world transaction dataset comprising more than 300,000 user-item interactions with a sparsity of 99.93%, the Weighted Hybrid Filtering model with $\alpha = 0.9$ achieved an F1@10 of 0.6634 and an ILD@10 of 0.8030, representing a specific trade-off relative to the individual CBF and CF models, as summarized below.

The hybrid model first outperforms the other two models in terms of the Recall@10 metric (0.8153 for the hybrid model, 0.7597 for CBF, and 0.8021 for CF with relative improvements of 7.32% and 1.65%, respectively), indicating that hybrid models can retrieve a larger set of relevant items than any individual model. However, it is inferior to CBF (F1@10 = 0.7246, Precision@10 = 0.8187) and CF (F1@10 = 0.7824) on F1@10 (0.6634) and Precision@10 (0.6121) resulting in trade-off rather than overall accuracy improvement.

Second, the hybrid model gets a much higher recommendation diversity (ILD@10=0.8030) than CBF (ILD@10=0.2819) with the relative improvement of about 184.9%, but it is still lower than the diversity in CF only (ILD@10=0.9215). In the simulated cold-start scenario, the CF returned undifferentiated prediction scores collapsing to uniform maximum value 1.0000 for new users.

Third, as illustrated in Table 7, the hybrid model could make contextually relevant recommendations by its CBF component even when the CF scores were zeros under the simulated cold-start condition. For example, products containing user preference keywords (e.g. Sweet Spicy Noodles for a user who states an interest in “noodles”) were assigned meaningful hybrid scores from content similarity. As this evaluation was performed on a single simulated user scenario, this result shows the potential of the hybrid model to fall back to content-based information in cold-start conditions, but cannot be considered as conclusive evidence for the full resolution of the cold-start problem.

In general, these results demonstrate that the Weighted Hybrid Filtering model ($\alpha = 0.9$)

does not always outperform the individual CBF and CF models. Instead it provides a particular combination of the best Recall@10 of the three methods, diversity much higher than CBF and the possibility to fall back on content-based information in low-interaction scenarios, at the expense of lower precision and F1@10 compared to both single methods.

There are several limitations to this study. The evaluation was conducted on a single real-world dataset from a single point-of-sale platform, and the results reported may not generalize to other culinary platforms or user populations. The cold-start evaluation for both CF and hybrid model is based on few simulated scenarios and not a systematic large-scale cold-start benchmark and should therefore be interpreted as illustrative rather than conclusive. Furthermore, a grid search over a small set of candidate values (0.1-0.9 with step 0.2) was used to select the best value of α without statistical significance testing among configurations; and the evaluation did not consider comparison with alternative hybridization strategies or other baseline recommender systems beyond CBF and CF alone.

Future work could address these limitations by testing the approach on other food datasets so as to investigate generalizability, by running more extensive cold-start experiments with more simulated users and items, using statistical significance testing while comparing configurations, and by exploring dynamic or learned alpha weighting strategies. and the use of contextual features such as time-of-day or seasonal preferences to further improve recommendation quality.

DECLARATIONS

Author Contributions: The first author conceptualized the study, developed the methodology, and prepared the original draft. The second author conducted the data curation, model implementation, and formal analysis. The third author performed the visualization and validation of the experimental results. The fourth author supervised the overall research progress, guided the execution of the experiments, and reviewed the manuscript. The fifth author provided project administration, reviewed and edited the manuscript, and approved the final version.

Funding: This research received no external funding.

Conflict of Interest: The authors declare no conflict of interest regarding the publication of this manuscript.

Generative Artificial Intelligence Use Statement: Generative AI tools were used for grammar and improving text readability during the preparation of the manuscript. The authors remain fully responsible for the originality, scientific content, analysis, interpretation, and conclusions of this work.

Ethics Approval and Informed Consent: This study utilized anonymized transaction data from Tantri Point of Sale (POS) and did not involve direct interaction with human participants. Informed consent was obtained for the usability evaluation phase.

Data Availability: The transaction datasets, evaluation results, and source code supporting the findings of this study are available from the corresponding author upon reasonable request, subject to permission from the data owner.

REFERENCES

- [1] A. Hakim, "The Digital Transformation of Traditional Culinary Businesses Towards Online Success," *Involv. Int. J. Bus.*, vol. 2, no. 1, pp. 11–25, Jan. 2025, doi: 10.62569/iijb.v2i1.105.
- [2] J. J. Yun, X. Zhao, G. Del Gaudio, V. Della Corte, and Y. Sadoi, "Leveraging business model innovation through the dynamics of open innovation: a multi-country investigation in the restaurant industry," *Eur. J. Innov. Manag.*, vol. 28, no. 4, pp. 1342–1365, Apr. 2025, doi: 10.1108/EJIM-07-2023-0607.
- [3] A. Dhaliwal, "Deconstructing contemporary consumer decision-making towards culinary consumption patterns: A multidimensional analysis of dine-in services in the age of cloud-based food service ecosystems," *Food Humanity*, vol. 6, p. 100996, May 2026, doi: 10.1016/j.foohum.2026.100996.
- [4] O. Alshareet and A. Awasthi, "Collaborative filtering in the age of AI: foundations, innovations, and emerging trends," *Computing*, vol. 107, no. 11, p.

- 216, Nov. 2025, doi: 10.1007/s00607-025-01564-2.
- [5] D. B. Rajesh and A. Kumar, "Collaborative filtering models an experimental and detailed comparative study," *Sci. Rep.*, vol. 15, no. 1, p. 31667, Aug. 2025, doi: 10.1038/s41598-025-15096-4.
- [6] M. Muflih, S. Ratna, and G. Mahalisa, "Development of A Collaborative Recommendation System Based on Singular Value Decomposition (SVD) on E-Commerce Data," *J. Appl. Inform. Comput. JAIC*, vol. 9, no. 6, Dec. 2025.
- [7] C. A. M. Zulaeka *et al.*, "Recommendation System for Solaria Food Menu Based on Customer Ratings Using Singular Value Decomposition," *J. Ris. Dan Apl. Mat. JRAM*, vol. 9, no. 1, Apr. 2025.
- [8] M. A. Dzulkarnain, "Application of item-based collaborative filtering method for culinary tourism recommendations in Bandung City," UIN Sunan Gunung Djati Bandung, 2019. [Online]. Available: <https://dspace.uinsgd.ac.id/items/1b2ef8e7-7a54-40cd-8566-6aa852a4aaf5>
- [9] R. Garapati and M. Chakraborty, "Recommender systems in the digital age: a comprehensive review of methods, challenges, and applications," *Knowl. Inf. Syst.*, vol. 67, no. 8, pp. 6367–6411, Aug. 2025, doi: 10.1007/s10115-025-02453-y.
- [10] B. Akkaya, "Current Trends in Recommender Systems: A Survey of Approaches and Future Directions," *J. Comput. Sci.*, vol. 10, no. 1, Jun. 2025, doi: <https://doi.org/10.53070/bbd.1652022>.
- [11] Z. Xia, A. Sun, J. Xu, Y. Peng, R. Ma, and M. Cheng, "Contemporary Recommendation Systems on Big Data and Their Applications: A Survey," *IEEE Access*, vol. 12, pp. 196914–196928, 2024, doi: 10.1109/ACCESS.2024.3517492.
- [12] D. K. Panda and S. Ray, "Approaches and algorithms to mitigate cold start problems in recommender systems: a systematic literature review," *J. Intell. Inf. Syst.*, vol. 59, no. 2, pp. 341–366, Oct. 2022, doi: 10.1007/s10844-022-00698-5.
- [13] S. Shinagam, "Comparison of Collaborative Filtering and Content-Based Filtering for Food Recommender Systems," Universiteit Utrecht, 2025.
- [14] M. R. Firdaus, "Integration of user context and diversity in content-based filtering for food recipe recommendations," UIN Sunan Gunung Djati Bandung, 2025. [Online]. Available: <https://dspace.uinsgd.ac.id/items/6dfc1a17-d000-4c3f-8496-6636b5372e80>
- [15] Q. Zhou *et al.*, "Hybrid collaborative filtering model for consumer dynamic service recommendation based on mobile cloud information system," *Inf. Process. Manag.*, vol. 59, no. 2, p. 102871, Mar. 2022, doi: 10.1016/j.ipm.2022.102871.
- [16] K. Peng, M. Raghavan, E. Pierson, J. Kleinberg, and N. Garg, "Reconciling the Accuracy-Diversity Trade-off in Recommendations," in *Proceedings of the ACM Web Conference 2024*, Singapore Singapore: ACM, May 2024, pp. 1318–1329. doi: 10.1145/3589334.3645625.
- [17] A. Chaudhari, A. A. Hitham Seddig, A. Sarlan, and R. Raut, "A Hybrid Recommendation System: A Review," *IEEE Access*, vol. 12, pp. 157107–157126, 2024, doi: 10.1109/ACCESS.2024.3480693.
- [18] C. S. Rajpoot, V. Tiwari, and S. K. Vishwakarma, "An enhanced hybrid recommender system using an adaptive optimization approach," Jun. 2026, doi: <https://doi.org/10.1007/s44163-026-00999-6>.
- [19] R. Widayanti, M. H. R. Chakim, C. Lukita, U. Rahardja, and N. Lutfiana, "Improving Recommender Systems using Hybrid Techniques of Collaborative Filtering and Content-Based Filtering," vol. 4, no. 3, Sep. 2023.
- [20] J. Karaman and Triyanna Widiyaningtyas, "Ensemble Hybrid Recommender System (CBF, CF, KNN, NBC) With Multi-View TF-IDF for Robust Preliminary Medical Diagnosis," *J. Tek. Inform.*, vol. 19, no. 1, pp. 62–72, Apr. 2026, doi: 10.15408/jti.v19i1.48705.
- [21] P. Kouki, J. Schaffer, J. Pujara, J. O'Donovan, and L. Getoor, "Generating and Understanding Personalized Explanations in Hybrid Recommender Systems," *ACM Trans. Interact. Intell. Syst.*, vol. 10, no. 4, pp. 1–40, Dec. 2020, doi: 10.1145/3365843.

- [22] S. Suriati, M. Dwiastuti, and T. Tulus, "Weighted hybrid technique for recommender system," *J. Phys. Conf. Ser.*, vol. 930, p. 012050, Dec. 2017, doi: 10.1088/1742-6596/930/1/012050.
- [23] M. F. Aljunid and M. D. Huchaiah, "An efficient hybrid recommendation model based on collaborative filtering recommender systems," *CAAI Trans. Intell. Technol.*, vol. 6, no. 4, pp. 480–492, Dec. 2021, doi: 10.1049/cit2.12048.
- [24] A. Sonule, H. Jagtap, and V. Mendhe, "Weighted Hybrid Recommendation System," *Int. J. Res. Anal. Rev.*, vol. 11, no. 1, 2024, doi: 10.56975/ijrar.v11i1.284204.
- [25] M. E. Çakır, Z. Çetinkaya, F. Horasan, and A. H. Yurttakal, "Adaptive weighting-based hybrid recommender system: self-learning adaptive recommendation (SL-ARec)," *PeerJ Comput. Sci.*, vol. 12, p. e3856, May 2026, doi: 10.7717/peerj-cs.3856.
- [26] Muhammad Azhari Desky, Defry Hamdhana, and Munirul Ula, "Diet Food Recommendation System with Hybrid Content-Based and Collaborative Filtering Approach," *Rabit J. Teknol. Dan Sist. Inf. Univrab*, vol. 11, no. 1, pp. 111–121, Jan. 2026, doi: 10.36341/rabit.v11i1.6647.
- [27] S. Suwitno, A. H. Wijaya, and W. Wiyono, "Integration Content-Based and Collaborative Filtering in AI-Based Culinary Recommendation For Community of Tangerang City," *Bit-Tech*, vol. 8, no. 2, pp. 1983–1994, Dec. 2025, doi: 10.32877/bt.v8i2.3182.
- [28] V. Alfianti, "Identifikasi kelengkapan bacaan Al-Qur'an menggunakan transfer learning wav2vec2," UIN Sunan Gunung Djati Bandung, 2024. [Online]. Available: <https://dspace.uinsgd.ac.id/items/1b2ef8e7-7a54-40cd-8566-6aa852a4aaf5>
- [29] F. I. Fajar, D. S. Maylawati, W. B. Zulfikar, A. Wahana, N. Lukman, and M. A. Ramdhani, "Random Forest, K-Nearest Neighbors, and Adaboost Regressor in Diamond Price Forecasting," in *2024 12th International Conference on Cyber and IT Service Management (CITSM)*, Batam, Indonesia: IEEE, Oct. 2024, pp. 1–7. doi: 10.1109/CITSM64103.2024.10775959.
- [30] R. Rijki, Y. A. Gerhana, G. Sandi, M. D. Firdaus, and E. Nurlatifah, "Comparison of Long Short-Term Memory and Recurrent Neural Network For Stock Market Price Movement Classification in Islamic Bank Finance," *CoreID J.*, vol. 4, no. 1, pp. 14–21, Apr. 2026, doi: 10.60005/coreid.v4i1.152.
- [31] K. Aarif, J. Deepika, M. A. Kumar, and P. Sanjana, "Deep neural collaborative filtering model for personalized travel recommendation," *Sci. Rep.*, vol. 16, no. 1, p. 4602, Jan. 2026, doi: 10.1038/s41598-025-34585-0.
- [32] S. Dewi, "Classification of Non-Civil Servant Performance Appraisal Using Naïve Bayes Classifier Algorithm," *CoreID J.*, vol. 1, no. 2, pp. 75–82, Jul. 2023, doi: 10.60005/coreid.v1i2.10.
- [33] M. Mouhiha, O. A. Oualhaj, and A. Mabrouk, "Combining Collaborative Filtering and Content Based Filtering for Recommendation Systems," in *2024 11th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Leeds, United Kingdom: IEEE, Jul. 2024, pp. 1–6. doi: 10.1109/WINCOM62286.2024.10658169.
- [34] H. Liu, W. Wang, Y. Zhang, R. Gu, and Y. Hao, "Neural Matrix Factorization Recommendation for User Preference Prediction Based on Explicit and Implicit Feedback," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–12, Jan. 2022, doi: 10.1155/2022/9593957.