Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Food Recipe Recommendation System with Content-Based Filtering and Collaborative Filtering Methods

Ni Putu Triska Widiantari[1)*], Dr. Eng. I Made Agus Dwi Suarjaya[2)], Ni Kadek Dwi Rusjayanthi[3)]
[1,2,3)]Universitas Udayana, Bali, Indonesia
[1)]triskawidiantari@student.unud.ac.id, [2)]agussuarjaya@it.unud.ac.id, [3)]dwi.rusjayanthi@unud.ac.id

Submitted :May 16, 2025 | **Accepted** : Jun 27, 2025 | **Published** : Aug 5, 2025

Abstract: Cooking food at home is a good step to reduce the consumption of unhealthy fast food. The diversity of recipe information available on the internet makes it difficult to choose recipes that match user preferences. Mobile technology can facilitate the transition to a healthy diet by recommending recipes based on user preferences. Therefore, in this study, a recommendation system was developed that can recommend recipes based on the preferences of Android users. This study fills the gap in existing research by integrating implicit feedback from user comments and food recipe information into an Android-based recommender system, which is rarely explored in previous studies that only focus on content-based or collaborative filtering. Using cosine similarity, a content-based recommendation system identifies the proximity between a recipe for food and its related context. The history of user comments on recipes serves as implicit feedback for the collaborative recommendation algorithm. This recommendation system generates recommendations in the form of the top ten food recipes with an evaluation matrix, referred to as NDCG@k and MAP@k. The tests revealed that a content-based filtering technique may produce helpful recommendations, with the highest similarity score of 0.41 for the entry "chocolate cake that you can easily make at home." Meanwhile, in the collaborative filtering method using the Neural Collaborative Filtering (NCF) approach, the system shows consistent performance improvements, with the MAP@10 value increasing from 0,6468 to 0,8402 and the NDCG@10 from 0,7371 to 0,8813 after 50 training epochs.

Keywords: Recommendation systems; content-based filtering; neural collaborative filtering; cosine similarity; implicit feedback

INTRODUCTION

The convenience offered by fast food restaurants increasingly facilitates people not to have to cook their own food (Nugroho et al., 2023). According to Donny K. Mulyantoro (Itsnaini & Alexander, 2024) Senior Researcher at the BRIN Public Health and Nutrition Research Center, explained that consuming fast food that is low in nutrition and high in calories, fat, sugar, and salt is one of the causes of children and adolescents becoming susceptible to serious diseases such as obesity, hypertension, and diabetes. With an obesity rate of 6.9% of its 277.53 million people as of 2023, Indonesia ranked 18th among 104 nations worldwide, according to the database (Muhamad, 2023). Cooking at home is healthier because it produces meals with lower concentrations of energy, fat, and sugar compared to those who cook at home less often (Alpaugh et al., 2020; Starke et al., 2021).

The number of culinary recipes currently available online is enormous. For example, the site Allrecipes.com, presented as the largest culinary social network in the world, has 40 million users and access to more than 3 billion recipes per year in 24 countries (Trattner et al., 2017). Their study registered more than 240,000 recipes and 17 million favorites. This abundance of data complicates the search for recipes that match one's preferences; therefore, an adaptive and contextual recommendation system is necessary.

Technological developments greatly influence various aspects of human life, from how to communicate to how humans search for and access information, including accessing food recipes for cooking in everyday life. Difficulty in choosing recipes on the internet can be caused by the large number of recipes available, which often do not reflect the user's eating habits. The food recipe recommendation system is designed to help users find recipes that match their preferences, thus saving time in choosing recipes. Previous research has used a content-based

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

filtering strategy that makes recipe recommendations based on the user-inputted ingredients by using TF-IDF and cosine similarity algorithms (Sari et al., 2024).

The two main recommendation strategies used in this study are content-based filtering and collaborative filtering. Content-based filtering provides recommendations based on the similarity of recipe features to users' preferences, while collaborative filtering uses the interaction patterns and preferences of other users with similar tastes. In this study, implicit feedback data, in the form of historical user responses to recipes, is used as the information source to build a recommendation model with collaborative filtering, eliminating the need for explicit judgments such as ratings. While previous studies have often relied solely on explicit ratings or single-item methods, this study fills this gap by combining content-based filtering with collaborative filtering that leverages implicit feedback, including user responses, to improve personalization and address data sparsity. This approach aims to indirectly identify similarities in preferences among users, allowing for more personalized and relevant recommendations.

LITERATURE REVIEW

The recommendation system by searching for document similarities is also seen in the research of (Widianto & Pebriyanto, 2024)which uses TF-IDF to represent documents in high-dimensional vectors and cosine similarity to calculate similarities. The results of this study showed an accuracy of 93.6% and the highest cosine similarity score of 0.598 for relevant documents.

The content-based filtering method was also used in a study by (Ramadhan & Musdholifah, 2021), which uses a sophisticated video recommendation system to help students select educational videos that align with the curriculum. This system uses YouTube video data, which is then analyzed using cosine similarity and TF-IDF techniques. Evaluation of the system, conducted through a survey of computer science students, yielded a Rata-Rata-Kinerja score of 81.13% across a range of metrics, including relevance and keberagaman, suggesting that the system is effective in promoting adventurous learning.

The implicit data used in the research of (Mahardhika et al., 2023) is in terms of the quantity of transactions for the kind of goods bought and the quantity of transactions for each user at a particular store. For the two anticipated outcomes—store prediction and category prediction—the average absolute error (MAE) and the root mean squared error (RMSE) were used to evaluate the recommendation system. The average of the two metrics is 0.865 for RMSE and 0.508 for MAE, this value represents poor model performance.

Implicit feedback data is also used in (Faizin & Surjandari, 2020) research with the Neural Collaborative Filtering (NCF) approach which aims to overcome the limitations of online marketplaces in providing a personalized purchasing experience. The NCF approach combines linear and non-linear modeling techniques to improve recommendation accuracy. The results show that NCF significantly outperforms traditional collaborative filtering techniques, especially with higher predictive factors, which confirms its effectiveness in improving recommendation systems for e-commerce platforms in Indonesia.

We conducted a literature review to ensure a clear and structured research methodology. This review focuses on some of the key components commonly used in recommendation and text processing systems, including TF-IDF, Cosine Similarity, and Neural Collaborative Filtering (NCF). These ideas serve as the theoretical foundation for integrating collaborative and content-based filtering techniques.

TF-IDF (Term Frequency-Inverse Document Frequency)

TF-IDF (term frequency-inverse document frequency) is a technique widely used in natural language processing to determine the importance of a word in a document relative to the entire document corpus. The two main components of this strategy are term frequency (TF), which counts how many times a word appears in a document, and inverse document frequency (IDF), which reduces the weight of terms that appear frequently in multiple documents (Juni Permana & Agung Toto Wibowo, 2023). TThis technique highlights words that occur frequently in one document but rarely in other documents in the corpus by multiplying the TF and IDF values to obtain a meaningful numerical representation of the text (Cahyani & Patasik, 2021).

Cosine Similarity

The cosine of the angle between the matching vectors is used to determine the cosine similarity between two publications. This method focuses on the vector's direction rather than its magnitude, making it suitable for normalized data such as TF-IDF vectors (Singh et al., 2020) Smaller angles (closer to 0) indicate higher similarity, and cosine similarity values range between 0 and 1, where 1 signifies identical directions or maximum similarity (Park et al., 2020). Although TF-IDF and cosine similarity are quite effective in handling text representation and matching, these techniques have limitations in capturing more complex semantic contexts.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Neural collaborative filtering (NCF)

NCF)is a neural network based framework that models user-item interactions in a nonlinear manner, offering greater flexibility than traditional matrix factorization techniques. The two primary components of NCF are the multilayer perceptron (MLP), which models nonlinear interactions across several hidden layers, and the generalized matrix factorization (GMF), which captures linear interactions. (He et al., 2017; Faizin and Surjandari, 2020). These components are combined into a final layer called NeuMF, which produces an interaction score reflecting the predicted relevance of an item to a given user. The NCF method, although superior in capturing non-linear patterns between users and items, has challenges in the need for large data and the potential for overfitting if not equipped with proper regularization.

METHOD

The development process for this recommendation system begins with data collection and a literature review to build a solid foundation in relevant methods, especially content-based filtering and collaborative filtering. Following this is data preprocessing stage to clean the data and get it ready for modeling.

Once preprocessing is complete, the workflow branches into two main paths. The first method uses content-based filtering, which makes recommendations based on characteristics of the item and the user profile. The second path leverages collaborative filtering, specifically neural collaborative filtering (NCF), to model user-item interactions based on historical user behavior.

After applying both approaches, we evaluate each recommendation result to assess the performance of both models using appropriate metrics. The results of each evaluation will guide conclusions about the effectiveness of both approaches in the recommendation task.

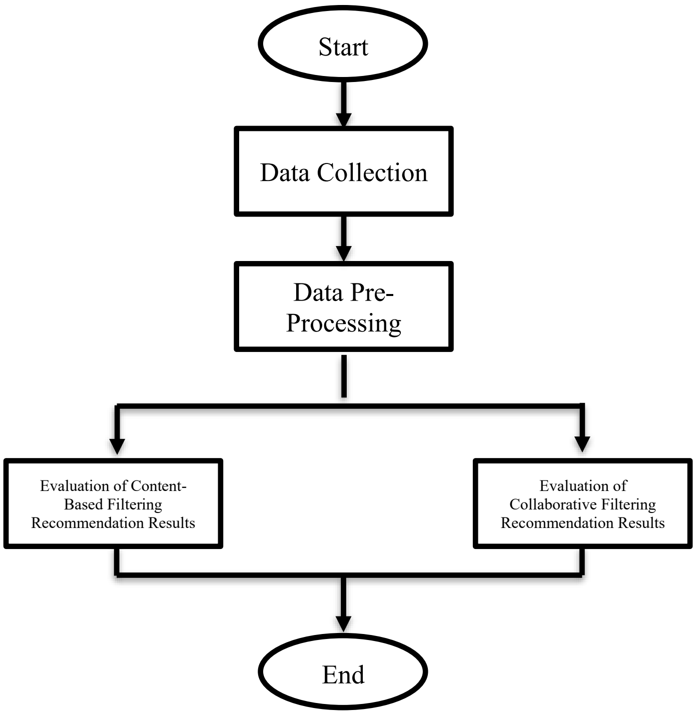

Fig. 1 Research Flow

Figure 1 presents an overview of the research flow of this study. Each research stream is detailed below.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Collection

Data collection is the initial step in this study. The data collection method for this study is described below.

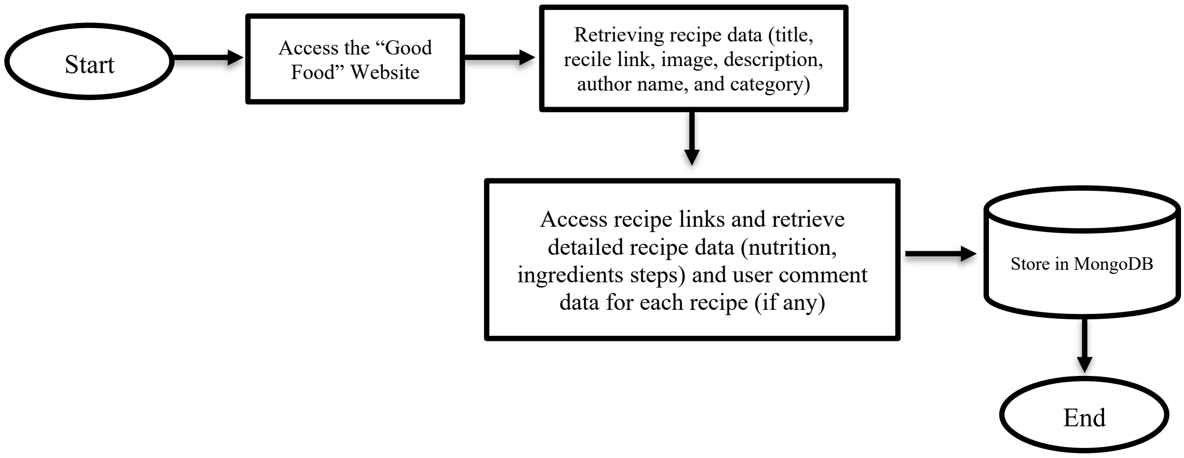

Fig. 2 Data Collection Flow

The data collection stage consists of collecting information about food recipes on the "Good Food" website, as well as comments on each food recipe. Data collection uses the Beautiful Soup framework and MongoDB as a database to store the collected data. The recipe dataset used in this study consists of 15,312 entries, containing information such as title, recipe ID, rating value, number of ratings, URL, image URL, description, publication date, author name, and category. User comments are collected as a form of interaction for the collaborative filtering method. A total of 222,121 comments were collected from 111,121 users, spread across 9,504 recipes, using Selenium-based scraping techniques. All data is stored in a MongoDB database.

Data Preprocessing

Data preprocessing is the step following successful data collection. This stage consists of several steps, the details of which can be seen in the following image.

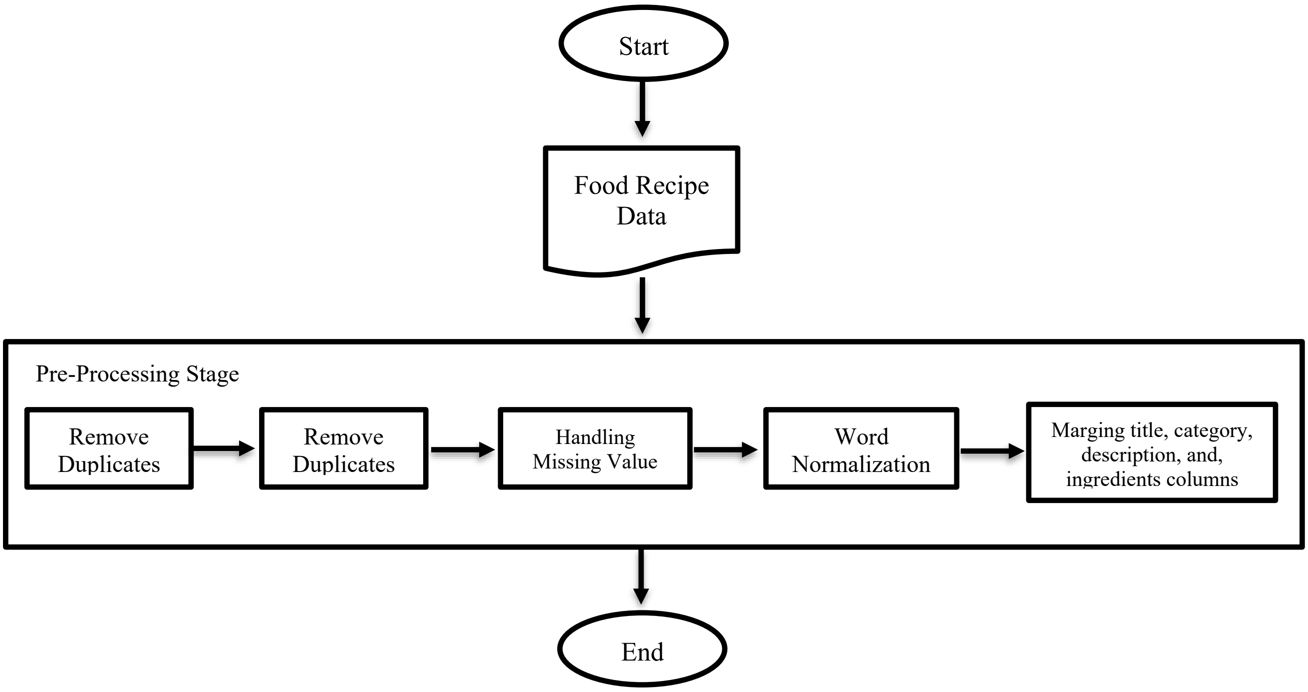

Fig. 3 Preprocessing Flow

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Figure 3 shows the preprocessing flow, which consists of several steps. The collected data is processed through duplicate removal, case correction, missing value management, and word normalization. Duplicate removal is performed to avoid data duplication, case correction ensures consistent letter formatting, missing value management is performed to avoid impacting model performance, and word normalization simplifies words to their standard format.

The title, category, description, and ingredient columns are then combined as the main elements of the content-based filtering method, as they represent important information about the recipe that will tailored to the user's preferences.

Content-based filtering

After the recipe data has gone through the pre-processing stage, the model for the recommendation system is created using a content-based filtering approach. The approach used to build a recommendation system model using content-based filtering is shown in the figure below.

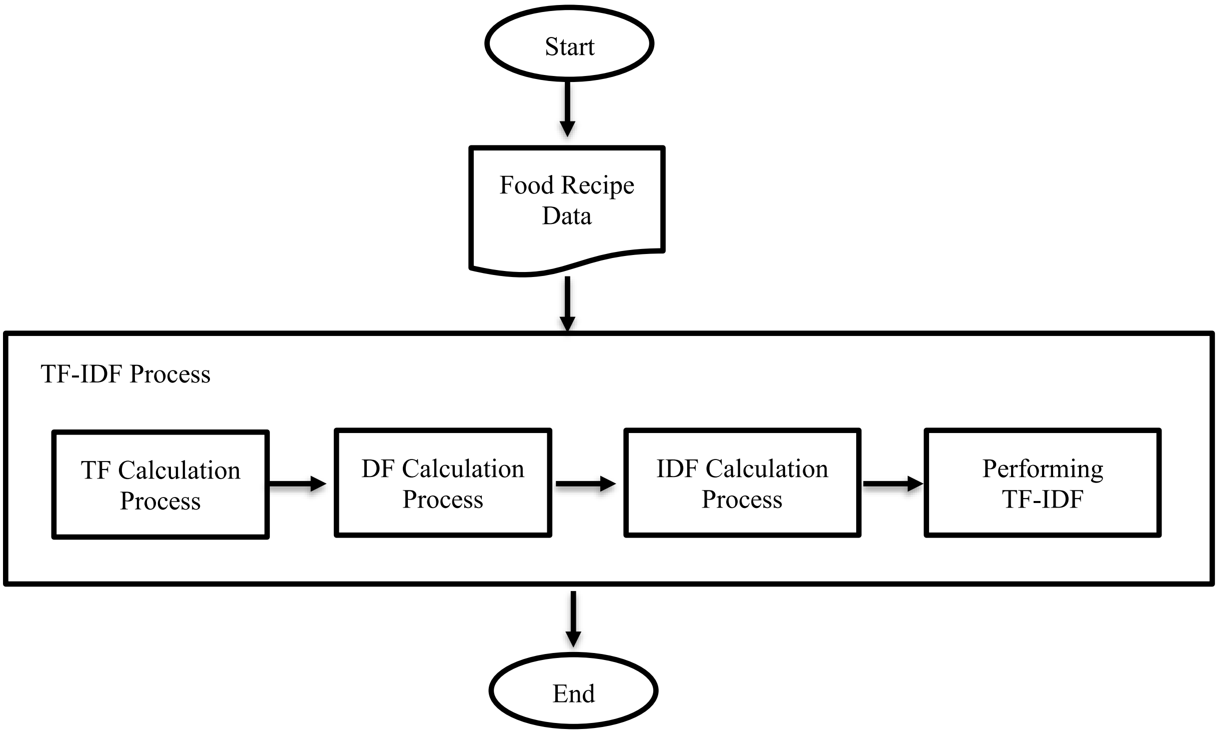

Fig. 4 Content-based filtering model development flow

Figure 4 shows the flow of creating recommendation model using content-based filtering method. This process starts with preprocessed recipe data. Then, the TF-IDF method is applied to weight the words based on the frequency of their occurrence in each recipe. The resulting vector is calculated for similarity using cosine similarity. Below is an illustration of the TF-IDF calculation in this study.

Table 1. TF-IDF Calculation Overview

Input:

Doc 1: "*healthy chicken curry quick and wholesome family meal with rice skinless chicken breasts*"

Doc 2: "*authentic quick indian one-pot with asian spices coriander chicken thighs skinned boneless*"

TF-IDF Calculation

Word	TF Doc 1	TF Doc 2	DF	IDF	TF-IDF Doc 1	TF-IDF Doc 2
asian	0.00	0.08	1	1.40	0.00	0.11
boneless	0.00	0.08	1	1.40	0.00	0.11
breasts	0.07	0.00	1	1.40	0.10	0.00
chicken	0.15	0.08	2	1.00	0.15	0.08
…	...	...	...	...	...	...

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Calculating the Term Frequency (TF) is the first step, which is the frequency of words in document. We look at the document frequency (DF), this indicates the frequency of the term's occurrence in a document. The calculation of the Inverse Document Frequency (IDF) follows, which gives more weight to words that occur infrequently.

The results of these three steps are used in the TF-IDF weighting, which represents the importance of a word to the entire document. This value is used to calculate similarity using cosine similarity. Below is an example of the similarity calculation using cosine similarity in this study.

Table 2. Cosine Similarity Example

Input: ***healthy chicken curry quick and wholesome family meal with rice skinless chicken breasts***				
Cosine Similarity To Overall Food Recipe Data				
No	**idRecipe**	***title***	***description***	**Simmilarity Scores**
1.	236622	chinese chicken curry	*"Cook our easy, healthy chicken curry in just 15 minutes. Serve this replica of your favourite takeaway dish with fluffy rice for a wholesome family meal"*	0.63794893
2.	234510	tangy chicken breasts	*"An easy meal in under half an hour that will impress the whole family"*	0.43000642
3.	264914	summer traybake chicken	*"A great dish for dinner parties and family meals in - quick, easy and very tasty"*	0.42451263

An illustration of how to calculate the cosine similarity between user input and all recipe data is shown in Table 2. The next step is to determine the similarity between documents after the word weights have been calculated using TF-IDF. This process generates a score that represents the similarity of a recipe to the input provided. The highest score indicates the most relevant recipe, as shown in the first row of the table, since its description is closest to the user input.

Neural Collaborative Filtering (NCF)

The NCF model is built from user feedback data as a representation of the interaction with recipes. Comments are considered as implicit feedback and in this study, it is assumed that comment data represents the user's interest without any direct evaluation in the form of ratings (Deng et al., 2022). Recipes with comments indicate that users have interacted with them. Figure 5 illustrates the flow of creating recommender system model using NCF.

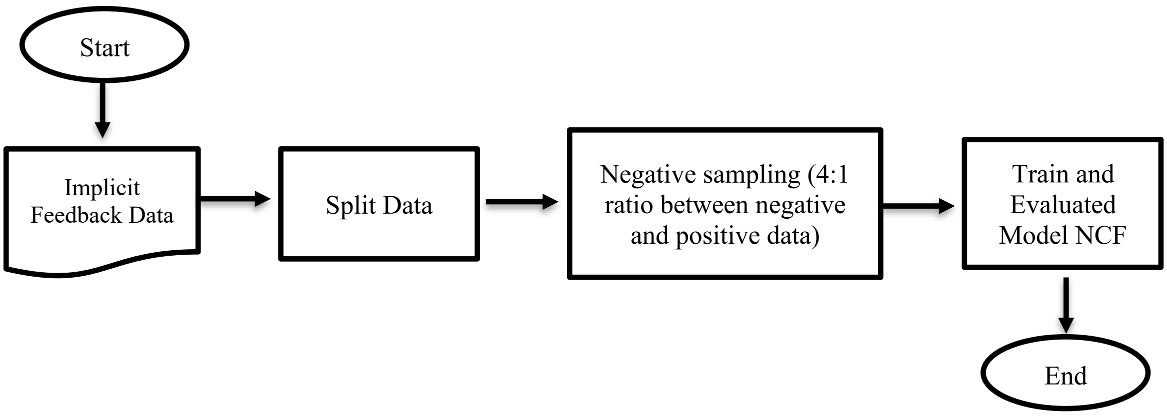

Fig. 5 NCF model development flow

Figure 5 shows the flow of creating a recommender system using the collaborative filtering method based on Neural Collaborative Filtering (NCF). Interaction between users and recipes is represented by the pairs idUser and idRecipe and is displayed chronologically. Data separation is performed using a time-based "leave-one-out" method, where each user's most recent interaction is used as testing data and the previous interaction as training data. This approach reflects a realistic scenario, where the recommendation system must predict the next interaction based on the user's history.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Furthermore, the negative sampling process is performed at a 4:1 ratio to provide enough examples of negative interactions for the model to learn to distinguish between recipes users like and dislike. Negative interactions indicate that the user no longer feels connected to the recipe, while positive interactions indicate engagement.

The NCF model is trained using the training data, using a deep learning neural network to recognize interaction patterns based on latent representations of users and recipes. After training is completed, the model is evaluated using the Hit Ratio@k and NDCG@k metrics to measure its performance. The model creation process ends when the evaluation results show satisfactory performance.

RESULT

This research results in a food recipe recommendation system, built in several main phases. The first phase includes the process of data collection, recipes, and data preprocessing, leading to the development of a model for a recommendation system. To suggest recipes based on content similarity, content-based filtering employing TF-IDF and cosine similarity is used throughout the model construction phase. Using comment data as a means of user involvement, the Neural Collaborative Filtering (NCF) technique was used to produce collaborative filtering.

Data Collection Results

The data collected includes recipe information such as recipe title, description, ingredients, cooking steps, category, and user comments. These user comments are used as implicit feedback that reflects the interaction between users and recipes. This data forms the basis for developing a recommendation system model in this study.

Data Preprocessing Results

The preprocessing phase occurs after data collection has been successfully completed. The goal is to make certain that the data is clean and appropriate for model building.

Table 3. Case Folding Result

Before *Case Folding*	After *Case Folding*
”Chinese chicken curry”	“chinese chicken curry”
“Cook our easy, healthy chicken curry in just 15 minutes. Serve this replica of your favourite takeaway dish with fluffy rice for a wholesome family meal”	“cook our easy healthy chicken curry in just 15 minutes serve this replica of your favourite takeaway dish with fluffy rice for a wholesome family meal”
Barney Desmazery	barney desmazery
"1 hr 10 mins", "Easy", "Healthy", "Gluten-free"	"1 hr 10 mins", "easy", "healthy", "gluten-free"

Table 3 shows the results of implementing the case folding step on the recipe data. All text has been converted to lower case. Symbols other than letters and numbers, such as commas and periods, have been removed from the description column. However, these symbols are still present in other columns.

Table 4. Normalization Result

Data Before Word Normalization	Data After Word Normalization
1 hr 5 mins	65
["50g butter", "85g smoked bacon lardons", "1 large onion, halved and finely chopped", "250g chestnut mushrooms, thickly sliced", "300g arborio risotto rice", "150ml dry white wine", "1.4l hot chicken stock", "140g cooked chicken, chopped", "50g grated parmesan, plus extra to serve (optional)", "½ small pack of flat-leaf parsley, chopped"]	“butter, smoke bacon lardons, large onion, chestnut mushroom, arborio risotto rice, dry white wine, hot chicken stock, cook chicken, grate parmesan, parsley”

Table 4 shows results of word normalization in the time and ingredient columns. After normalization, the time data becomes more uniform in minutes. Ingredient data used to be more detailed and varied. These data have now been reordered so that each column contains only one ingredient.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Recommendation Results Using the Content-Based Filtering Method

The cosine similarity score, which is computed using the similarity of the content, and the accuracy of the recommendation in relation to the input are the factors used to determine the outcome of the content-based filtering method. These recommendations are generated based on the context of the food or ingredients the user is consuming. Recommendations are also made based on the chosen recipe. The recommendation results using the content-based filtering method are shown below.

Table 5. Food Recipe Recommendation Results Using Content-Based Filtering

Input:* *chocolate cake that is easy to make at home			
No	**Judul Resep Makanan**	**Id Resep**	***Similarty Score***
1	Fudgy dark chocolate cake	228270	0.4056631504642899
2	Chocolate rice crispy cakes	238728	0.39080369304392865
3	Reindeer cake pops	236937	0.3888464036602488
4	Hedgehog cake	236760	0.3758649812758478
5	Double-the-love chocolate cake	222801	0.36689264734609434
6	Light & fluffy chocolate mocha cake	227571	0.3588736153301697
7	Double chocolate loaf cake	265773	0.35773120079106013
8	Chocolate ice cream	701196	0.3560370688116682
9	Triple chocolate & peanut butter layer cake	240065	0.34918310614553477
10	Halloween cake pops	222356	0.34483702211321465

Table 5 shows 10 recommendation results based on the content-based filtering method. The input used is input_user = ['*chocolate cake that is easy to make at home*'], so the system makes recommendations based on chocolate-based cakes. The results have a modest level of relevance, as indicated by the greatest similarity score of 0.4056. The system can identify the most common keywords in the input data and suggest words or sentences that are highly comparable. Because the TF-IDF approach just takes word frequency into account and ignores semantic context, this similarity score is appropriate even though it is not very high.

Recommendation Results with Neural Collaborative Filtering

Recipe recommendations made using the collaborative filtering method were evaluated using the NDCG@k and MAP@k metrics to measure the relevance of the recommendation results. To evaluate the model's capacity to learn user-recipe interactions, training performance was also examined using the loss value for every epoch.

Table 6. Performance of NCF models based on Loss, MAP@10, and NDCG@10 over 50 training epochs

Epoch	***Loss***	**MAP@10**	**NDCG@10**
1	0.6728	0.6468	0.7371
2	0.6394	0.6754	0.7584
3	0.6111	0.6947	0.7728
4	0.5866	0.7078	0.7826
5	0.5658	0.7171	0.7895
6	0.5478	0.7245	0.7950
7	0.5321	0.7318	0.8005
8	0.5183	0.7381	0.8052
9	0.5065	0.7442	0.8098
...	...	...	...
41	0.3801	0.8317	0.8750
42	0.3794	0.8328	0.8758
43	0.3779	0.8339	0.8766
44	0.3766	0.8351	0.8775
45	0.3756	0.8362	0.8783
46	0.3744	0.8371	0.8790
47	0.3731	0.8381	0.8797
48	0.3718	0.8387	0.8802
49	0.3707	0.8393	0.8806
50	0.3702	0.8402	0.8813

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 9, Number 3, July 2025
DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

In this study, the developed model is an extension of Neural Collaborative Filtering (NCF) by adding several additional features, such as user and item embedding, and recipe representation vectors from the content-based filtering method. The addition of these features aims to enrich the representation of user interactions, improve prediction quality, and overcome challenges in implicit data such as comments.

This approach is in line with the concept of Feature-Augmented Collaborative Filtering explained by Kulshreshtha (2017) in the Deep Augmented Hybrid Collaborative Filtering (DAHCF) model. In DAHCF, user and item features are learned through deep learning architectures such as Autoencoders and Restricted Boltzmann Machines (RBM), then combined and fed into a neural network to model interactions and produce more accurate predictions.

The model was trained for 50 epochs and showed stable performance, with a decrease in loss value from 0.6728 to 0.3702. In addition, the main evaluation metrics also increased significantly, namely MAP@10 from 0.6468 to 0.8402 and NDCG@10 from 0.7371 to 0.8813. This shows that the model is not only able to generate relevant recommendations but also arrange them in a high-quality order. The use of additional features has been proven to help the model understand user preferences more deeply, improve system personalization, and provide a more contextual recommendation experience.

DISCUSSIONS

Two primary methods are used in this study to create a recipe recommendation system: collaborative filtering utilizing a neural collaborative filtering (NCF) model supplemented with extra features, and content-based filtering. The evaluation's findings demonstrate that every approach has unique benefits and drawbacks, which are crucial for comprehending how well the system works in various contexts.

The contextual relevance of user input is successfully captured by the content-based filtering method, which makes use of cosine similarity and TF-IDF vectorization. For instance, when the user inputs "easy homemade chocolate cake," the system provides a list of recipes for chocolate desserts that are graded highly. On the other hand, 0.4056, the highest similarity score, is regarded as moderate. Due to TF-IDF's exclusive reliance on lexical frequency and disregard for semantic meaning, this shortcoming is to be expected. Additionally, it might be challenging to create varied recommendations beyond known user preferences, since content-based filtering frequently suggests recipes with identical phrases.

Strong personalization capabilities are demonstrated by the collaborative filtering model with a feature-enhanced NCF architecture, on the other hand. The model is able to learn user preferences through implicit feedback by integrating interaction timelines, recipe vectors from content-based filtering, and latent user and object embeddings. The model's performance gradually improved after over 50 training sessions, lowering the loss from 0.6728 to 0.3702. Evaluation measures like NDCG@10 went from 0.7371 to 0.8813 and MAP@10 went from 0.6468 to 0.8402. According to these findings, the model successfully displays pertinent recipes in the recommendation list in addition to identifying them.

But there are also certain restrictions noted. Because of its deep learning architecture and requirement to comprehend numerous hyperparameters (e.g., embedding dimension, learning rate, batch size), the NCF-based model takes a lot longer to train than TF-IDF. Even if training is done effectively with Google Colab's GPU resources, techniques like pruning or model distillation could be needed to scale this model for real-time applications. Additionally, the model may induce bias into interaction labeling because it believes that user feedback represents favorable preferences, which is not necessarily the case.

Our findings continuously show that NCF is better than conventional collaborative filtering techniques for handling sparse and noisy data when compared to similar studies, including that of (Faizin & Surjandari, 2020; Kulshreshtha, 2017), who documented the efficacy of NCF in e-commerce with implicit feedback. Further supporting the idea of deep enhanced hybrid collaborative filtering (Kulshreshtha, 2017) is the incorporation of extra contextual features into NCF, demonstrating that feature improvement can greatly increase recommendation accuracy.

CONCLUSION

Using two techniques—content-based filtering and collaborative filtering with Neural Collaborative Filtering (NCF)—the study successfully developed a recipe recommendation system. From these results, the following can be concluded, the content-based filtering method provides effective recommendations based on material and textual similarities, which is very useful for users with specific questions or preferences. The NCF-based collaborative filtering model is able to learn user-recipe interaction patterns through implicit feedback (user comments) and consistently improves performance over the training period, as evidenced by the increase in MAP@10 and NDCG@10 scores.

However, this study has several limitations, including the use of implicit feedback may not always reflect true user preferences, as not all users leave comments or ratings. The cold-start problem remains a challenge for collaborative filtering for new users or new recipes without interaction data. For future use, this system can be

*name of corresponding author

DOI : https://doi.org/10.33395/sinkron.v9i3.14778
e-ISSN : 2541-2019
p-ISSN : 2541-044X

enhanced by adding additional data on user behavior, such as click-through rates or time spent reading recipes. The use of a hybrid recommendation system that combines the advantages of content-based and collaborative filtering techniques. Explore more advanced deep learning models or transformer-based architectures to capture richer contextual and semantic relationships in the data.

REFERENCES

Alpaugh, M., Pope, L., Trubek, A., Skelly, J., & Harvey, J. (2020). Cooking as a health behavior: Examining the role of cooking classes in a weight loss intervention. *Nutrients*, *12*(12), 1–13. https://doi.org/10.3390/nu12123669

Cahyani, D. E., & Patasik, I. (2021). Performance comparison of tf-idf and word2vec models for emotion text classification. *Bulletin of Electrical Engineering and Informatics*, *10*(5), 2780–2788. https://doi.org/10.11591/eei.v10i5.3157

Deng, H., Zhai, C., & Zheng, L. (2022). Neural Collaborative Filtering for Chinese Movies Based on Aspect-Aware Implicit Interactions. *IEEE Access*, *10*, 114540–114551. https://doi.org/10.1109/ACCESS.2022.3217911

Faizin, A., & Surjandari, I. (2020). Product recommender system using neural collaborative filtering for marketplace in indonesia. *IOP Conference Series: Materials Science and Engineering*, *909*(1). https://doi.org/10.1088/1757-899X/909/1/012072

Itsnaini, F. M., & Alexander, H. B. (2024, May 30). Banyak Remaja Terkena Obesitas karena Makan 'Junk Food' Berlebihan. *Kompas.Com*.

Juni Permana, A. H. J. P., & Agung Toto Wibowo. (2023). Movie Recommendation System Based on Synopsis Using Content-Based Filtering with TF-IDF and Cosine Similarity. *International Journal on Information and Communication Technology (IJoICT)*, *9*(2), 1–14. https://doi.org/10.21108/ijoict.v9i2.747

Kulshreshtha, D. (2017). Feature Augmented Deep Neural Networks for Collaborative Filtering. *IJCAI*. https://grouplens.org/datasets/movielens/

Mahardhika, M. N., Rahayu, F., & Zuchriadi, A. (2023). Product Recommendation System Using Implicit Feedback Based on Collaborative Filtering in E-Commerce. *Seminar Nasional Teknologi Informasi, Komunikasi Dan Industri (SNTIKI) 15*.

Muhamad, N. (2023, July 18). *10 Negara dengan Prevalensi Obesitas Tertinggi di Dunia*. Databoks.

Nugroho, R. H., Samsudin, A., Dwi, D., Zahrain, M., Ainun, R., Putri, S., Rizma, A., & Ayu, D. (2023). *Analisis Faktor Yang Mempengaruhi Pembelian Konsumen Pada Restoran Cepat Saji*. *4*, 1213.

Park, K., Hong, J. S., & Kim, W. (2020). A Methodology Combining Cosine Similarity with Classifier for Text Classification. *Applied Artificial Intelligence*, *34*(5), 396–411. https://doi.org/10.1080/08839514.2020.1723868

Ramadhan, F., & Musdholifah, A. (2021). Online Learning Video Recommendation System Based on Course and Sylabus Using Content-Based Filtering. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *15*(3), 265. https://doi.org/10.22146/ijccs.65623

Sari, N. N. K., Priskila, R., & Putra, P. B. A. A. (2024). IMPLEMENTASI CONTENT-BASED FILTERING MENGGUNAKAN TF-IDF AND COSINE SIMILARITY UNTUK SISTEM REKOMENDASI RESEP MASAKAN. *Jurnal Keilmuan Dan Aplikasi Bidang Teknik Informatika*.

Singh, R. H., Maurya, S., Tripathi, T., Narula, T., & Srivastav, G. (2020). Movie Recommendation System using Cosine Similarity and KNN. *International Journal of Engineering and Advanced Technology*, *9*(5), 556–559. https://doi.org/10.35940/ijeat.E9666.069520

Starke, A., Willemsen, M., & Trattner, C. (2021). Nudging Healthy Choices in Food Search Through Visual Attractiveness. *Frontiers in Artificial Intelligence*, *4*, 18.

Trattner, C., Parra, D., & Elsweiler, D. (2017). Monitoring obesity prevalence in the United States through bookmarking activities in online food portals. *PLoS ONE*, *12*(6). https://doi.org/10.1371/journal.pone.0179144

Widianto, A., & Pebriyanto, E. (2024). Document Similarity using Term Frequency-Inverse Document Frequency Representation and Cosine Similarity. *Journal of Dinda Data Science, Information Technology, and Data Analytics*, *4*(2), 149–153. http://journal.ittelkom-pwt.ac.id/index.php/dinda

*name of corresponding author