SISRES: Web-Based Culinary Recommendation with Collaborative Filtering

Qurrotul Aini1, Fadly Hakim Muhammad2, Eri Rustamaji3, Yamin Thwe4

Abstract—Due to its advantageous location on the border of Jakarta and Tangerang, Tangerang Selatan is a highly developed city. In this instance, the Tangerang Selatan Department of Culture and Tourism (Dispar Tangsel) is required to use the application to assist the general public in realizing the availability of information in the Tangerang Selatan area. Twenty to twenty-five restaurants submit applications each year to the Dispar Tangsel Tourism Business Registry (TDUP). Dispar Tangsel must choose and decide on TDUP licensing priorities from among TDUP requests in order to open a restaurant. The purpose of the research is to offer recommendations to the community on food choices. Rapid application development (RAD) was used in the system's development. Additionally, the collaborative filtering technique has been employed by the decision support system to determine the amount of criteria or weight for restaurants using the weighted sum algorithm and for restaurants using cosine-based similarity algorithms. Additionally, the system design tool made use of MySQL as a database, PHP, the Codeigniter framework, and the unified modeling language (UML). The result demonstrates that the system is capable of displaying the output in accordance with the user's expectations during black box testing, which evaluates the functionality of the system based on the specifications. Collaborative filtering in SISRES can yield a significant improvement in recommendation accuracy. By collectively analyzing user preferences and behaviors, the algorithm can provide more relevant and personalized recommendations.

Index Terms—Collaborative filtering, pearson correlation similarity, culinary, MySQL, rapid application development.

I. INTRODUCTION

Tangerang Selatan is a potential city with the prospect of rapid development because its strategic area borders Jakarta and Tangerang. In Tangerang Selatan, many new shops or restaurants have sprung up and offered a variety of dishes or foods with a new taste or a distinctive taste of the restaurant. Also, street vendors on the side of the road have sprung up and increased; therefore, Tangerang Selatan gets a new tourism sector, i.e., culinary. Based on data from the Culture and Tourism Office of South Tangerang City, the growth of restaurants in Tangerang Selatan continued to increase from 2010 to 2015, then culinary in Tangerang Selatan overgrew. Figure 1 shows the increasing number of restaurants from 2010-2015 in South Tangerang. Every year, the average number of restaurants in South Tangerang increases by 30% [1].

The trend of global tourism is always open to new technology, even more so on the development of web application technology, which has led to an increase in interest in the field of electronic tourism or also commonly called e-tourism. At the same time, tourists take a more active role in the process of producing tourism content. They systematically convey tourism information by utilizing Web 2.0 technology such as social networks, blogs, wikis, and google maps, as well as updating dynamic content.

Fig. 1. Growth number of restaurants in Tangerang Selatan [1]

Disparr Tangerang Selatan experienced the impact of developments in science and technology. Therefore, various interested parties must easily access the required data and

Received: 15 July 2023; Revised: 11 August 2023; Accepted: 19 August 2023

1) Q. Aini, UIN Syarif Hidayatullah Jakarta, Indonesia (e-mail: qurrotul.aini@uinjkt.ac.id).
2) F. H. Muhammad, UIN Syarif Hidayatullah Jakarta, Indonesia (e-mail: fadlyhakim195@gmail.com).
3) E. Rustamaji, UIN Syarif Hidayatullah Jakarta, Indonesia (e-mail: eri.rustamaji@uinjkt.ac.id).
4) Y. Thwe, Rajamangala University of Technology Thanyaburi (RMUTT), Thailand (e-mail: yamin.th@mail.rmutt.ac.th).

http://journal.uinjkt.ac.id/index.php/aism
information effectively and efficiently. Therefore, the government needs a system to assist with new restaurant registration in the city, which will help restaurant owners to make it easier to do TDUP. Also, the system may display the restaurant location geographically till users can determine by their wishes.

Collaborative filtering (CF) is one of the knowledge discovery methods that is often used in recommender systems, where the memory is able to accentuate the current actual user data for estimating new things for the certain user [2]. The recommender system (RS) is a system that employs the beliefs of a group of users to facilitate entities in groups effectively seeking interest in new things from many possible alternatives. A decision support system deals with determining a decision for a user who has benefits and is appropriate from several or many choices. RS can also be specified as software that helps users get valuable and exciting items from an extensive collection of items in a personalized way [3]. Many studies on CF have been carried out since 2004. CF application has been collaborated with other algorithms in solving problems and increasing accuracy. Most of the research are in the entertainment domain is in the form of movie recommendations and e-commerce. While the datasets used is mainly from MovieLens 100K2 [4].

Another previous study declared that by applying CF, the built recommendation system could make it easier for users to determine beauty salons based on the rating given by customers. Furthermore, the system is supported by the functional feasibility test with the Guttman scale by five beauty shop administrators. The result revealed that the value of the feasibility test is close to 1, which means it is feasible [5]. Using this algorithm, researchers have been able to investigate the modified CF based on user confidence and time context. The approach method in similarity calculation is an integrated algorithm, i.e., adjusted cosine-based similarity (ACOS), CF user confidence, and time context [6]. This study investigated the performance of algorithms similar to the calculation of mean absolute error (MAE), root-mean-square deviation (RMSE), precision, recall, and F1-measure. The simulation results showed that user confidence and time context (UCTC_User) is more effective in improving the performance of the recommendation system. Therefore, the current research aims to build a decision support system that uses collaborative filtering, the level of criteria/weights of restaurants with the weighted sum algorithm and determines restaurants with cosine-based similarity [6]. Another study found that using CF and an SVM, it was possible to select a user's first Netflix movie genre based on what people were saying about it on Twitter. RMSE for Collaborative Filtering was used to calculate performance. Additionally, the SVM's precision and recall performance were calculated [7]. The advantage of this study over the previous one is the distribution of restaurants in South Tangerang and the recommendation of desired restaurants with high ratings at the nearest location.

The current research aims to design and create a distribution system using collaborative filtering. While the contribution of this study provides a restaurant registration system, makes decisions, and provides suggestions for choosing a place to eat according to preferences. The recommendation system includes designing a culinary distribution system in Tangerang Selatan and information on the distribution of restaurant locations that can make it easier for the Dispar Tangsel, restaurant owners, and the community.

II. RESEARCH METHOD

This section sums up the research development that begin with the preliminary study. The preliminary study is conducted by observations and interviews at the Tangsel Dispar office to obtain information flow and running systems that support information on the distribution of restaurant locations. Figure 2 depicts the research framework, from data collection through system development.

A. Data Collection

This research collects spatial data in the form of an administrative map of Tangerang Selatan in *.shp format, which has been processed from the Bappeda Tangerang Selatan using geographic projections (longitude and latitude), with a scale of 1:35,000. Meanwhile, non-spatial data is in the form of restaurant data at the Tangsel Tourism Office.

The digital map displayed is owned by Google by accessing the Google Maps API, where there are four types of map model options provided:
- Roadmap to display a 2-dimensional ordinary map;
- Satellite, to display satellite photos;
- Terrain, to show the physical relief of the earth's surface and show how high a location is;
- Hybrid, showing satellite photos on which what appears on the roadmap (streets and city names) is illustrated.

http://journal.uinjkt.ac.id/index.php/aism
B. Rapid Application Development (RAD)

RAD is a design and prototyping-focused approach to development that aims to gain immediate user feedback. RAD suggests greater flexibility compared to traditional development approaches, which involve initial planning and subsequent execution. Quick incremental improvements and ongoing user feedback iterations help to ultimately produce superior results [8], [9]. The RAD steps, which start with requirement planning and end with implementation, are depicted in Fig. 3.

1) Requirement planning
to develop the system, requirements planning is needed to identify goal, requirement information, and analyze the problem of the running system. Then, create the proposed system and identify the drawback and advantages of them. After that, conduct item-based CF computations by calculating the similarity of two culinary objects with the Pearson correlation [10].

2) Workshop Design
The use case diagrams are the first step in this stage's description of the proposed system design. The use case diagram includes actor and use case identification as well as use case narratives, activities, and flow diagrams for the activities. Database and interface design are also covered [11], [12], [13].

3) Implementation
There are two different processing, i.e., programming and testing. In programming, the Graphic User Interface (GUI) results are entered into a programming language at this stage. Therefore, it can be run as an interface display application for user interaction with web-based systems. Also in testing part, to determine whether the users can accept the proposed system, the authors conducted an acceptance test system using black box testing. Black box testing is a type of testing that treats software as internal performance is unknown. Therefore, the testers view the software as a "black box," which is not essential to see the contents, but it is enough to be subjected to the testing process on the outside [14].

C. Collaborative Filtering

Presently, the most common recommendation algorithms can be broken down into content-based, collaborative filtering (CF), hybrid, and other algorithms. Content-based recommendation employs a series of discrete features of items, e.g., the genres, directors, and movie actors, to generate recommendations [15]. CF recommendation aims to calculate a list of exciting items to target users based on the preferences of their like-minded neighborhood. These two approaches are often combined to make a hybrid recommendation [16].

Commonly, the CF algorithm is divided into the memory-based CF
recommendation [17]. The memory-based CF recommendation finds comparable products by making full use of prior data. The item-based CF suggestion and the user-based CF recommendation are two ways to separate the memory-based CF recommendation. Based on the similarity between the items, the item-based CF recommendation identifies an item set comparable to the target project. The user-based CF recommendation bases its forecasts and suggestions on neighborhood data from active users [18].

The CF filters data based on user behavior to provide new information to other users. Collaborative filtering is the process of filtering or evaluating items using the opinions of others [3]. Figure 4 shows a schematic of the collaborative filtering process. The CF algorithm represents all users’ ratings as a matrix where each entry is the rating value of the user for each item. The active user (u) in this scheme is a user who will be searched for items using the CF algorithm [10], [19]. Each user expresses his opinion about his list of items. The set of opinions is called the rating of user u and is denoted by Iu. Once the system has determined the nearest neighbors, it will represent the items that the user may like in two forms. As seen in Fig. 4, CF employs two procedures given the rating table as an input, where u1, u2, ..., um are system users and I1, I2, ..., In is a list of items to be predicted:

- Prediction is a numerical value Pui expressing the predicted likelihood of item i that does not belong to Iu. This predicted value is on the same scale as opinion values provided by user u.
- Recommendation lists N items that the active user will like the most. The recommended list must include items not purchased by the active user. This CF algorithm interface is referred to as Top-N recommendation.

Item-based collaborative filtering seeks out trends in user rating behavior to forecast how users will rate future goods or food. Akhmad, for instance, like the food at Rumah Makan Padang "RMP" and Rumah Makan Sunda "RMS," but he has not eaten food at Rumah Makan Betawi "RMB." Then Akhmad notices that Rahmat and Harry like food at "RMB" with the same score as the previous two restaurants. Thus, Akhmad concludes that he will also enjoy the food "RMB," as shown in Fig. 5.

To find users with similar interests, it is necessary to find a similarity value describing how similar a user is to another user. User similarity is only calculated from objects that users have rated. One of the formulas used is Correlation-based Similarity [10], i.e., Pearson Correlation Similarity:

\[ \text{sim}(i,j) = \frac{\sum_{u \in U} (R_{ui} - \bar{R})(R_{uj} - \bar{R})}{\sqrt{\sum_{u \in U} (R_{ui} - \bar{R})^2} \sqrt{\sum_{u \in U} (R_{uj} - \bar{R})^2}} \]

(1)

The similarity between two items i and j are measured by computing the Pearson-r correlation (corr(i,j)). To make the correlation computation accurate, we must first isolate the co-rated cases (i.e., cases where the users rated both i and j). Let U stand for the group of people who rated both i and j. One of the techniques considered to see the rating of target users is the weighted sum, as per the following formula [10]:

\[ P_{ui} = \frac{\sum_{\text{all similar items}, N} (S_{ui} \cdot N^{R_{ui}, N})}{\sum_{\text{all similar items}, N} (S_{ui} \cdot N)} \]

(2)

This method calculates the prediction item i for a user u by computing the sum of the ratings given by the user on items similar to i. Each rating is weighted by the corresponding similarity Sui between items i and j. Usually, it can denote the prediction Pui as shown in (2).

III. RESULT

A. Requirement Planning

Based on observation and an interview with Dispar Tangerang, there is difficulty for the public to find accurate information about restaurant locations that suit their tastes and suggestions for restaurants that suit the people of South Tangerang City and restaurant owners. It is also difficult to conduct TDUP licensing. To obtain a TDUP license, restaurant owners must visit the Dispar Tangerang, which requires a long process.

There are areas for improvement in the current system in Dispar Tangerang, i.e., the TDUP license registration process must come to the office directly; there is data redundancy when processing data, and it needs a web-based information system. The proposed system is called a restaurant spatial information system (SISRES). The shortcomings of the current system are as follows:

- Unavailability of restaurant suggestions according to taste, so people need clarification to determine a suitable restaurant.
- Not updated data about the restaurant's status in real-time, and there is a possibility of data redundancy.
- The duration of TDUP approval for restaurants; therefore, restaurant owners must wait for days.
- The advantages of the SISRES are as follows:
  - This system can manage and provide information about restaurants and distribution locations that the user has determined.
  - Restaurant owners can submit applications through the system, which then the Office can verify and validate the application.
  - Restaurants data can be updated in real-time because it uses a web-based system to manage the data online.

The data taken is the rating data of the five most popular restaurants based on the highest number of visits to culinary objects visited by visitors based on statistical data at Dispar
Tangsel. The dataset is taken from an user as input for item-based CF. Table 1 is the data obtained from users.

<table>
<thead>
<tr>
<th>User</th>
<th>Culinary Object</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMB</td>
</tr>
<tr>
<td>Akhmad</td>
<td>5</td>
</tr>
<tr>
<td>Rahmat</td>
<td>3.5</td>
</tr>
<tr>
<td>Harry</td>
<td>1</td>
</tr>
<tr>
<td>Qosim</td>
<td>4</td>
</tr>
</tbody>
</table>

Item-based collaborative filtering is used to predict empty rating data (grey column). Next, the calculation process determines the similarity value of two culinary objects that users have rated. First, we take two columns in Table 1, RMB and RESTO. Second, calculate the similarity between RMB and RESTO:

Average rating on RMB = \( \frac{5 + 3.5}{5} = 3.17 \)

Average rating on RESTO = \( \frac{4 + 4}{4} = 4.10 \)

The similarity between RMB and RESTO using (1).

Pearson correlation between RMB and RESTO:

\[
\frac{(5-3.17)(5-4.10) + (3.5-3.17)(4-4.10)}{\sqrt{(5-3.17)^2 + (4-3.17)^2} \sqrt{(5-4.10)^2 + (4-4.10)^2}} = \frac{1.62}{1.69} = 0.96
\]

Similar calculations are also carried out among culinary object, as shown in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>RMB</th>
<th>RESTO</th>
<th>RMP</th>
<th>RMS</th>
<th>CAFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMB</td>
<td>1.00</td>
<td>0.96</td>
<td>0.81</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>RESTO</td>
<td>0.96</td>
<td>1.00</td>
<td>-0.07</td>
<td>-0.04</td>
<td>0.58</td>
</tr>
<tr>
<td>RMP</td>
<td>0.81</td>
<td>-0.07</td>
<td>1.00</td>
<td>1.00</td>
<td>0.78</td>
</tr>
<tr>
<td>RMS</td>
<td>0.83</td>
<td>0.58</td>
<td>0.78</td>
<td>0.78</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The same calculation is also carried out the similarity between one restaurant and other restaurants, resulting in a restaurant similarity matrix as in Table 3.

<table>
<thead>
<tr>
<th></th>
<th>RMB</th>
<th>RESTO</th>
<th>RMP</th>
<th>RMS</th>
<th>CAFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMB</td>
<td>1.00</td>
<td>0.96</td>
<td>0.81</td>
<td>0.81</td>
<td>0.83</td>
</tr>
<tr>
<td>RESTO</td>
<td>0.96</td>
<td>1.00</td>
<td>-0.07</td>
<td>-0.04</td>
<td>0.58</td>
</tr>
<tr>
<td>RMP</td>
<td>0.81</td>
<td>-0.07</td>
<td>1.00</td>
<td>1.00</td>
<td>0.78</td>
</tr>
</tbody>
</table>

The rating prediction is then determined using the outcomes of the previously calculated similarity value. Table 4 makes the supposition that Qosim's rating has never been rated; hence, the data are used for testing.

<table>
<thead>
<tr>
<th></th>
<th>RMB</th>
<th>RESTO</th>
<th>RMP</th>
<th>RMS</th>
<th>CAFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qosim</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

In Table 4, it is known that empty culinary objects have been given a rating. However, RMB and RMS data are used as testing data, and Qosim is considered not to have given a rating to the two objects. Nevertheless, in reality, Qosim has given a rating to RMB 3 and RMP 4. Next, determine the number of neighbors for calculation. In this example, we use six neighbors in the prediction calculation. Then determine several conditions in determining the number of neighbors, i.e.:

- Neighboring culinary objects that Qosim has rated.
- Similarity between RMB and its neighbors is above 0.5.

For example, in Table 4, it is known that there are three neighbors whose value is above 0.5, i.e., RESTO, RMS, and Cafe. It is also known that there is the same similarity value, i.e. RMB with RMS and RMP with RMS with a similarity value of 0.81. However, because Qosim has not rated RMP, the similarity used is the similarity of CAFE with RMS. Then calculate the prediction using a weighted sum according to (2):

Qosim's rating prediction of RMB:

\[
\frac{(4 \times 0.96) + (4 \times 0.83) + (4 \times 0.81)}{0.96 + 0.83 + 0.81} = \frac{6.4}{2.6} = 2.4
\]

RMB is expected to receive a grade from Qosim of 2.4 as opposed to its initial rating of 3.

**B. Workshop Design**

The design process SISRES starts with a use case diagram that explains the activities carried out, such as actor identification, use case identification, use case diagram design, use case narrative, activity diagram, and sequence diagram that explains the flow of activities on the SISRES. In the use case diagram in Fig. 6, there are six actors who interact with each other, with an explanation of their respective functions and tasks. Next, create an activity diagram that describes the various activity flows in the SISRES, including how each flow starts, decisions that may occur, and how they end. Activity diagrams can also describe parallel processes that may occur in multiple executions. There are 20 activity diagrams in this system. The activity diagram about ‘View Result of Document Verification and Approval’ is shown in Fig. 7.
The activity diagram (Fig. 7) can be done by the Head of Dispar office, to access it the Head of Dispar must first enter the system, select the application approval menu, then the system will display a list of application data. If the actor chooses ‘detail’, the system will display application data details. While, class diagrams describe the state of a system, as well as services to manipulate the state of methods or functions so that classes have three main areas: names, attributes and methods. The class diagram of SISRES is shown in Fig. 8.

Next, we create the sequence design. The sequence diagram is carried out by the Head of Dispar (Fig. 9). To access it, Head of Dispar must first enter the system, select the document menu, then the system will display a list of application data. If Head of Dispar chooses ‘detail’, the system will display application data details, if the actor chooses ‘Approval’, the system will change the status of the application data and save it to the database. Likewise, if the actor chooses delete, the system will delete the data.

Database design is also made in SISRES. First, mapping the cardinality of the information system designed based on the class diagram to determine the relationship between classes and what keys are used as foreign keys. The mapping can be seen in Fig. 10.
After the coding process is complete, the next stage is system testing using black box testing, by conducting test cases against the application using a test table by entering data into the system and then seeing the output whether it matches the expected results as in Table 5.

<table>
<thead>
<tr>
<th>Activities</th>
<th>Expected Results</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration</td>
<td>The public and restaurant owners can register and then the restaurant owner can apply for a TDUP license.</td>
<td>Suitable</td>
</tr>
<tr>
<td>Log in</td>
<td>All users can successfully access the system.</td>
<td>Suitable</td>
</tr>
<tr>
<td>Edit Profile</td>
<td>Restaurant owners can edit the restaurant profile.</td>
<td>Suitable</td>
</tr>
<tr>
<td>Upload Document</td>
<td>Restaurant owners able to input the document upload form.</td>
<td>Suitable</td>
</tr>
<tr>
<td>Upload Menu</td>
<td>Restaurant owner can input the restaurant menu form.</td>
<td>Suitable</td>
</tr>
<tr>
<td>View Rating</td>
<td>Restaurant owners can view restaurant ratings.</td>
<td>Suitable</td>
</tr>
<tr>
<td>View and evaluate document application</td>
<td>Staff of Dispar successfully view submission documents and evaluate documents.</td>
<td>Suitable</td>
</tr>
<tr>
<td>View result of evaluation and document verification</td>
<td>Head of Section Dispar successfully viewed the evaluation result and verified the document.</td>
<td>Suitable</td>
</tr>
<tr>
<td>Reject document</td>
<td>Staff and Head of Section Dispar successfully rejected the documents.</td>
<td>Suitable</td>
</tr>
<tr>
<td>View result of document verification and approval</td>
<td>Head of Dispar successfully views the verified document and approves the document.</td>
<td>Suitable</td>
</tr>
</tbody>
</table>

The research findings in the form of research data are further discussed or critically interpreted with particular relevant theoretical approach. State the gap between current research and previous/related research which signifies the significance of your research. Data can also be supported with the presentation of tables, images, etc. Captions for images are placed below the picture, also with providing sequenced numbering. Captions for table is written above it with sequenced numbering so that it can be easily referred to. Some data can be stated in the following sentence.

**D. Discussion**

This study raises the CF capability to be applied in spatial information systems that recommend desired culinary, including location and rating by public. This research is in line with the research of [20], [21], [22]. CF is also applied to restaurant recommendation systems, where their research is focused on computing CF performance, such as MAE and accuracy. Using item-based CF, which produces recommendation results based on the similarity value of culinary objects, to make suggestions is another benefit of the current research. Culinary items with a high similarity score are more popular for the public. However, the present study does not explain CF performance calculation, but only focuses on

http://journal.uinjkt.ac.id/index.php/aism
creating a web-based for culinary recommendations (SISRES) using the RAD development approach.

IV. CONCLUSION
A web-based culinary recommendation system is built to solve problems in the Dispar Tangerang Selatan for culinary object owners and the public in determining which culinary objects to visit. This system is designed using the RAD development method with UML notation, PHP programming, a MySQL database, the Codeigniter framework, ArcGIS, and Mapguide Maestro. According to the results of black box testing, SISRES can operate and respond according to its function. The collaborative filtering technique has been employed by the decision support system to determine the amount of criteria or weight for restaurants using the weighted sum algorithm and for restaurants using cosine-based similarity algorithms. Collaborative filtering in SISRES can yield a significant improvement in recommendation accuracy. By collectively analyzing user preferences and behaviors, the algorithm can provide more relevant and personalized recommendations. Additionally, the system design tool made use of MySQL as a database, PHP, the Codeigniter framework, and the unified modeling language (UML).

The next challenge for CF to be aware of the recommendations is user- and item-based dot products. Thus, if an item is not seen during training, the system cannot create an embedding for it and cannot query the model with this item. From some of the constraints in this research and the challenges of item- and user-based collaborative filtering, future work can be highlighted on developing geographic information systems with applications that use heuristic methods to generate embeddings of fresh items. Also, apply usability testing to evaluate the user experience with moderated and unmoderated methods.

REFERENCES

http://journal.uinjkt.ac.id/index.php/aism