

# IMPLEMENTATION OF THE K-NN ALGORITHM TO DETERMINE PATTERNS OF TOURIST DESTINATION RECOMMENDATIONS

Alfhataya Pratiwi<sup>1</sup>, Reyni Salim Hasibuan<sup>2</sup>, Resia Maranatha Saragih<sup>3</sup>  
STMIK Pelita Nusantara<sup>1,2,3</sup>

**Abstract.** In this digital era, the recommendation system is an important tool in helping users to make decisions, including choosing tourist destinations. This study aims to implement the K-Nearest Neighbors (k-NN) algorithm in determining the recommendation pattern for tourist destinations. The k-NN algorithm was chosen because of its ability to classify based on the closest neighbor of a point, so that it can obtain recommendation patterns that are more personal and relevant to user preferences. This implementation method involves data collection, data preparation, model building, classification process, model evaluation, and finally implementation of the model in a recommendation system. The results of this study indicate that a recommendation system based on the k-NN algorithm is able to provide suggestions for tourist destinations that are more in line with user preferences, thus potentially increasing user satisfaction and making a positive contribution to the tourism industry. However, the effectiveness of this model is highly dependent on the quality and quantity of data used. Therefore, good data collection and preparation is very important in the implementation of this algorithm. This research opens opportunities for further research and development in the field of tourist destination recommendation systems using other algorithms or combining several algorithms to obtain more optimal results. The effectiveness of this model is highly dependent on the quality and quantity of data used. Therefore, good data collection and preparation is essential in the implementation of this algorithm. This research opens up opportunities for further research and development in the field of tourist destination recommendation systems by using other algorithms or combining several algorithms to get more optimal results. The effectiveness of this model is highly dependent on the quality and quantity of data used. Therefore, good data collection and preparation is very important in the implementation of this algorithm. This research opens up opportunities for further research and development in the field of tourist destination recommendation systems by using other algorithms or combining several algorithms to get more optimal results.

Keywords. *Information Correspondance*  
K-NN Algorithm, *Alfhataya Pratiwi et.al*  
Tourism, *STMIK Pelita Nusantara*  
Recommendation [Alfhataya Pratiwi@gmail.com](mailto:Alfhataya Pratiwi@gmail.com)  
System, Tourist  
Destinations, Machine  
Learning.

## INTRODUCTION

Tourism is one sector that has a significant economic impact on many countries and regions[1]. In today's digital era, tourists have a tendency to seek recommendations for tourist destinations based on various criteria, such as personal interests, previous reviews, and popular trends. Information technology, especially machine learning techniques, offers solutions to provide the right recommendations according to user preferences[2]. Tourist destination recommendations are one of the services that are increasingly important in the tourism industry and related fields[3][4]. With so many tourist destinations available, it is very important for tourists to get recommendations for destinations that best suit their preferences. Therefore, an effective and efficient method in providing these recommendations is very important. The use of machine learning technology and methods in the

recommendation process is becoming increasingly popular[5][6]. One of the most popular and efficient algorithms in recommendation systems is K-Nearest Neighbors (K-NN). Through this algorithm, a tourist destination can be recommended based on the similarity of characteristics with other destinations that have been visited or liked by users.

The implementation of the K-NN algorithm can be used to determine patterns for tourist destination recommendations. The K-NN algorithm is a method used to classify data based on the shortest distance to data objects. The K-NN algorithm can also be used to provide recommendations in recommendation systems. Implementation of the k-NN (K-Nearest Neighbors) algorithm to determine patterns in tourist destination recommendations involves several steps. The k-NN algorithm is a popular machine learning method for classification and regression.

Several studies have been carried out by applying the K-NN algorithm to tourist destination recommendation systems[7]. For example, a study uses the K-NN algorithm and the Rapidminer application to provide recommendations for tourist attractions in Labuan. In this research, an accuracy of 83.33% was found with a value of  $K=5$ [8]. Apart from that, the K-NN algorithm can also be used to provide tourist route recommendations. A study uses the K-NN algorithm and the Traveling Salesman Problem method to provide recommendations for tourist routes in North Toraja[9].

However, keep in mind that the K-NN algorithm does not work well on high-dimensional data because with a large number of dimensions, it becomes difficult for the algorithm to calculate the distance in each dimension. Therefore, it is necessary to perform feature scaling (standardization and normalization) before applying the K-NN algorithm to the data set. The aim of this research is to implement the k-NN algorithm in determining patterns for tourist destination recommendations. It is hoped that with the implementation of this algorithm, the tourist destination recommendation system can become more personalized and relevant to tourist preferences, thereby increasing tourist satisfaction and ultimately contributing to the growth of the tourism industry.[10]. Implementation of the K-NN algorithm in a tourism recommendation system can help the tourism industry understand and serve tourists better[11]. With more accurate recommendations, tourists will have a more satisfying experience, while tourism service providers can improve promotion and marketing efficiency.

## METHOD

The following are the steps for implementing the k-NN algorithm for tourist destination recommendations[12] [13]:

1. Data collection:

Collect tourist destination data along with relevant features that are used as a reference in recommending destinations. Examples of these features could include location, cost, popularity, ratings, activity, and so on.

This data must also have a tourist destination label or category that corresponds to each data, for example, "Beach", "Mountain", "City", and so on.

2. Data Preparation:

Perform data preprocessing such as filling in missing values (if any), deleting irrelevant data, and normalizing features that require uniform scalability.

3. Data Sharing:

Separate the data into two parts: training data and test data. Training data is used to train the model, while test data is used to test the model's performance.

4. Model Building:

Create a k-NN model using training data. In this step, we can determine the K value, namely the number of nearest neighbors that will be used to classify tourist destinations.

5. Classification Process:

Use the k-NN model that has been created to classify tourist destinations on the test data based on the features they have.

6. Model Evaluation:

Evaluate the k-NN model using appropriate evaluation metrics, such as accuracy, precision, recall, or F1-score. This metric will give an idea of how well the model can correctly recommend tourist destinations[14].

7. Recommended Destination:

Once the model is deemed good, use the model to recommend tourist destinations based on the features of the user who wants to get recommendations.

8. Model Completion:

If the recommendation results are not satisfactory, you can consider improving the quality of the model by changing the K value, changing the evaluation metrics, or even using other machine learning algorithms that are more suitable for this problem.

9. In-App Implementation:

Integrate the k-NN model that has been trained in applications or platforms that require a tourist destination recommendation system.

Note: Implementation of the k-NN algorithm in this context will be more effective if the data used includes many examples of recommended tourist destinations with various features and categories. The more relevant data, the more accurate the k-NN model can recommend tourist destinations.

### **K-Nearest Neighbor (KNN) Algorithm**

The K-Nearest Neighbor (KNN) algorithm is a classification method for a set of data based on learning data that has been previously classified.[10]. The KNN algorithm divides the data into two parts, the first is training data and the second is test data. Training data is used when making basic predictions, while test data consists of values predicted by. Then the training data is converted into a vector and a distance that is calculated using several methods, for example euclidean distance or cosine similarity[9][8].

### **Data Mining**

Data mining carries out the process of using statistical, mathematical, artificial intelligence and machine learning techniques to extract and identify information for knowledge related to various large databases stored in repositories or storage places.[15].

### **Rapid Miner**

Is software used for data processing and becomes a solution for analyzing data mining, text mining and predictive analysis[16]. This application makes it easy for users to calculate large amounts of data by using operators whose function is to modify data. Data is connected through nodes in the operator and then connected to

the results node to display the final results so that the results can be displayed visually with graphs[17].

Case data that we will test is regarding the needs and desires of tourists' types of tourism and then compared with existing training data. The data set used is 18 data as shown in the image below.

## RESULTS AND DISCUSSION

### Process

For the initial stage of the process, we prepare the Excel data format as shown in the image below.

1				
2	Aktivitas	Jenis Wisata	Jenis Harga	Destinasi Rekomendasi
3	Oversea	Refreshing	Murah	Pulau-Pulau Kecil
4	Oversea	Adventure	Sedang	Pulau Rinca dan Pulau-Pulau lain
5	Oversea	Honeymoon	Mahal	Kawasan Komodo dan Pulau lain
6	Overland	Adventure	Murah	Goa Rangko
7	Overland	Adventure	Sedang	Waerebo
8	Overland	Refreshing	Mahal	Keliling Flores
9	Overland	Refreshing	Murah	Goa Rangko
10	Oversea	Adventure	Mahal	Kawasan Komodo dan Pulau Lain
11	Overland	Adventure	Murah	Goa Rangko
12	Overland	Refreshing	Sedang	Ende-Kelimutu Kelimutu
13	Overland	Adventure	Mahal	Bajawa-Ende-Kelimutu
14	Oversea	Honeymoon	Sedang	Pulau-Pulau Kecil
15	Oversea	Refreshing	Mahal	Kawasan Komodo dan Pulau lain
16	Oversea	Adventure	Murah	Pulau-Pulau Kecil
17	Oversea	Honeymoon	Murah	Pulau-Pulau Kecil
18	Overland	Honeymoon	Mahal	Labuan Bajo
19	Overland	Honeymoon	Murah	Goa Rangko
20	Oversea	Refreshing	Sedang	Pulau-Pulau Kecil

Figure 1. Initial data on tourist attractions

Next, we open the Rapidminer application and click Import Data to add the Excel data format that we will test.

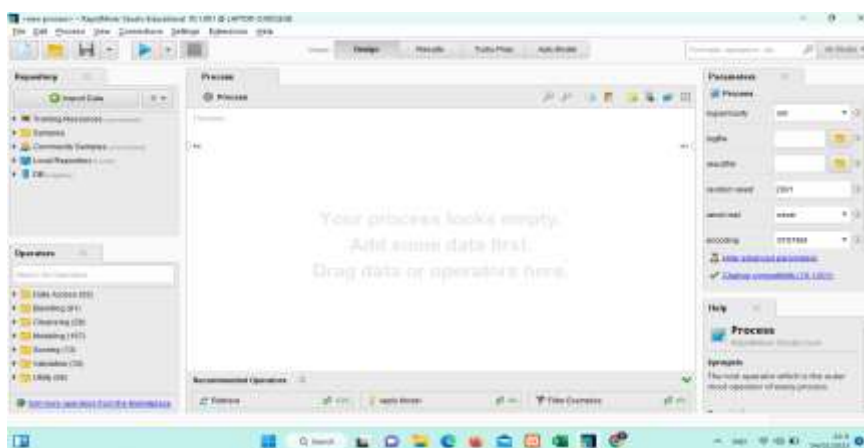


Figure 2. Initial appearance of RapidMiner



**Figure 3.** Select Data Location

After we add data to this application, we click next to continue with the next step



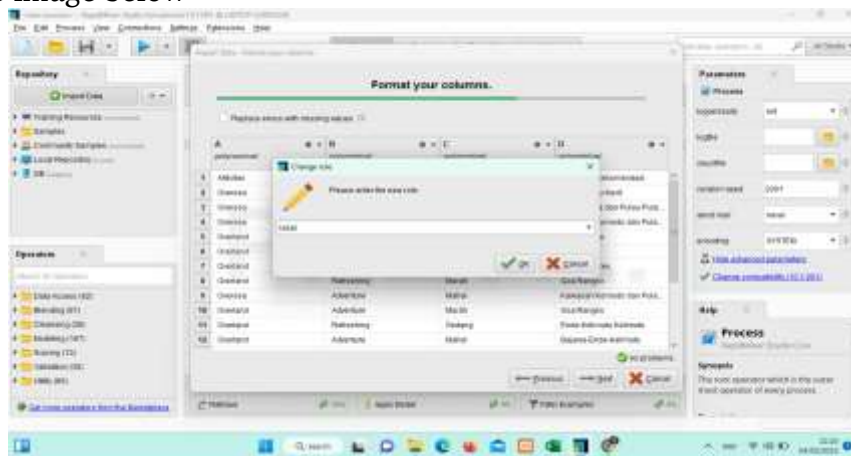
**Figure 4.** Select Data

After we next a data table will appear like the picture above indicating the data we are using is correct with no errors, then we next come back.



**Figure 5.** Import Data

In part D, we change it to change role, then we create a label, OK, then it will look like the image below



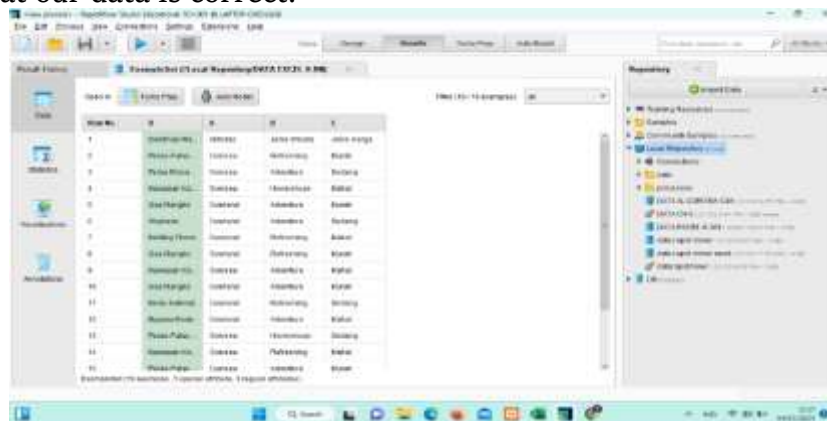
**Figure 6.** Add Role

After that the result will be like the image below



**Figure 7.** Roles

Then we go next and after we finish, the data will appear again as below, indicating that our data is correct.



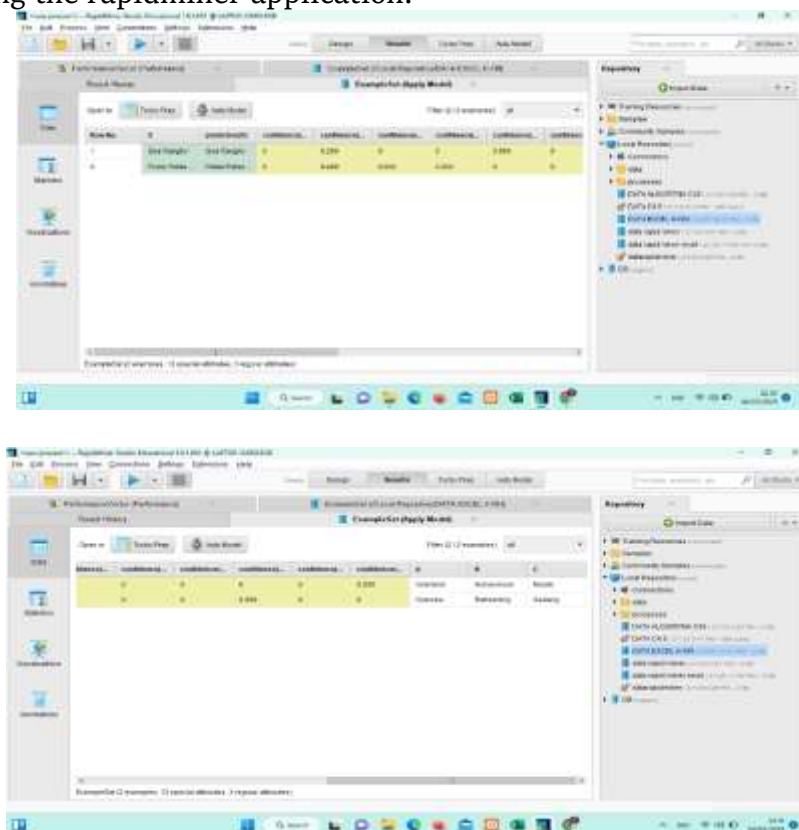
**Figure 8.** The data is correct

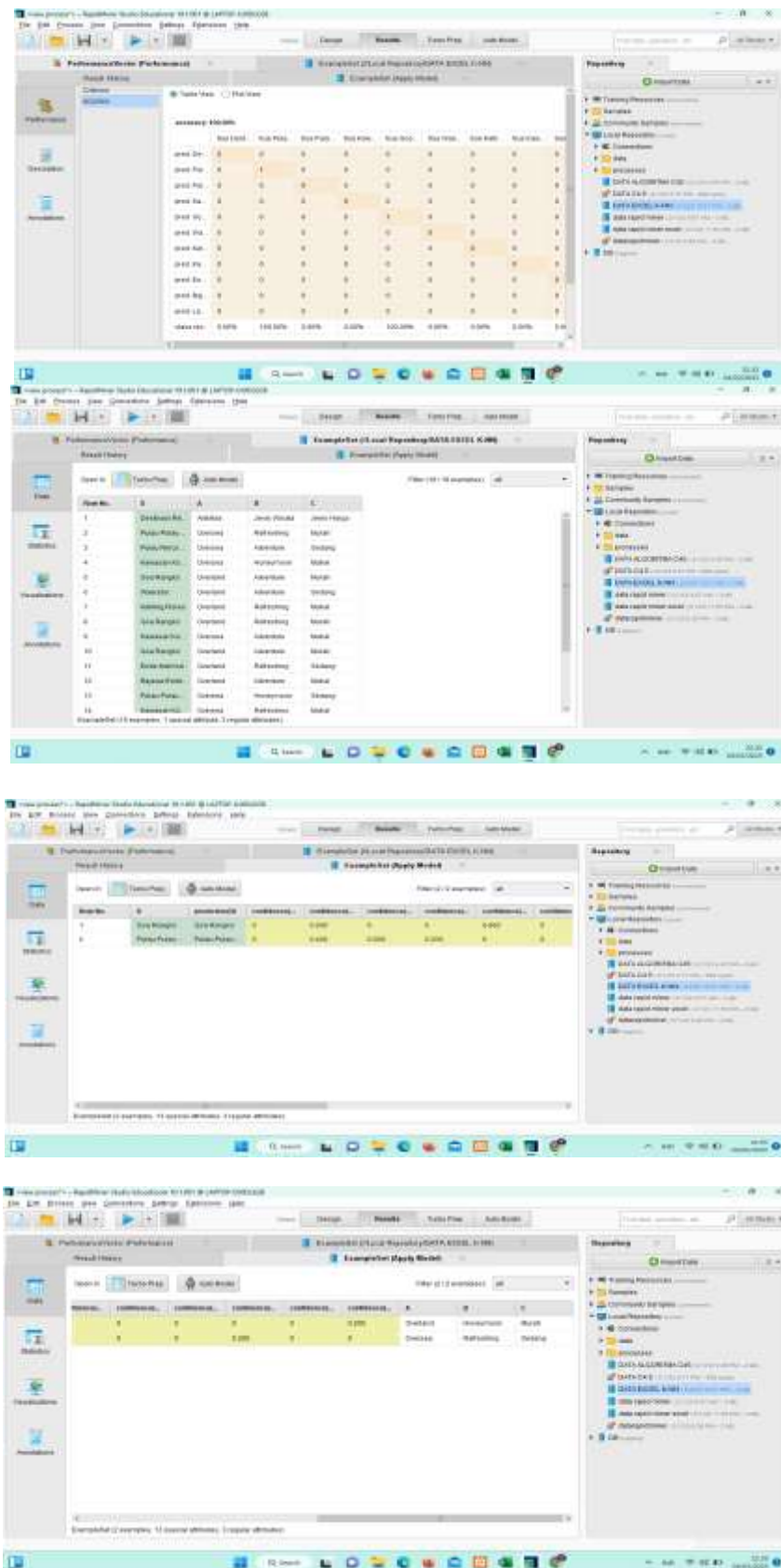
For the next stage, we go to design, just click design on Rapidminer and look for the data that we have processed then we drag it, then we select the split data operator, we drag it then we set the sampling type, change it to linear sampling, after we do that Move to edit, click Add Entry, set the ratio to 0.9 and set Add Entry again to 0.1, then click OK. We click on the operator and look for the algorithm that we use, namely the K-NN Algorithm, then we drag it back to the process column, after that we look for it back to the apply model operator, we drag it back to the process column, after everything we have done, and finally we connect all the data, then The result will be as below.



Figure 9. Apply Model

The final step is to click play to see the results of the data that we have tested or processed using the rapidminer application.





**Figure 10.** Data test results with RapidMiner



This is the result of our group's work in the process of testing the K-NN Algorithm using the Rapidminer application, if there are deficiencies in writing and errors in the process, please understand and forgive us for that, we from group 3 (three) say thank you.

## CONCLUSION

To find recommendation patterns for tourist destinations, the k-NN algorithm has been used successfully and yielded significant results. Based on certain features of a tourist destination, such as location, cost, popularity, rating, and available activities, this algorithm allows the recommendation system to provide more personalized and relevant destination recommendations. In the tourism sector, this method has great benefits because it provides recommendations for tourist attractions according to visitor preferences. Ultimately, the tourist experience will improve and the tourism industry will thrive as a result. However, keep in mind that the effectiveness of the k-NN algorithm is highly dependent on the quality and amount of data used. To implement this algorithm, Good data collection and preparation is essential. Also, although the k-NN algorithm shows good results, there are many other algorithms that can also be used in recommendation systems, such as collaborative filtering, content-based filtering and deep learning. Each algorithm has advantages and disadvantages, and depending on the situation and needs, may be the right choice. Overall, the k-NN algorithm has great potential for use in tourist destination recommendation systems. This opens up opportunities for further research and development in this area. Each algorithm has advantages and disadvantages, and depending on the situation and needs, may be the right choice. Overall, the k-NN algorithm has great potential for use in tourist destination recommendation systems. This opens up opportunities for further research and development in this area. Each algorithm has advantages and disadvantages, and depending on the situation and needs, may be the right choice. Overall, the k-NN algorithm has great potential for use in tourist destination recommendation systems. This opens up opportunities for further research and development in this area.

## REFERENCE

- [1] MJ Hadi, "Study of the Potential of New Jurit Village to Become an Educational Tourism Destination," *Alamtana J. Devoted. Mass.*, vol. 1, no. 1, 2020.
- [2] ADP SIAHAAN, "Analysis of the Competitiveness of the Tourism Industry in Medan City," *Univ. North Sumatra*, vol. 16, 2020.
- [3] Yusuf Fadlila Rachman, Darmawan Lahru Riatma, M. Masbahah, Trisna Ari Roshinta, Fendi Aji Purnomo, and Putri Nugrahaningsih, "PRIORITY RECOMMENDATION SYSTEM FOR TOURISM DESTINATION DEVELOPMENT IN MADIUN DISTRICT," *NAWASENA J. Ilm. Tourist*, vol. 1, no. 2, 2022, doi: 10.56910/nawasena.v1i2.231.
- [4] Dwi Putra Githa and DP Singgih Putri, "DESIGN OF A TOURISM DESTINATION RECOMMENDATION SYSTEM IN BALI," *SINTECH (Science Inf. Technol. J.)*, vol. 2, no. 2, 2019, doi: 10.31598/sintechjournal.v2i2.388.
- [5] W. Septiandari, S. Hidayatullah, S. Alvianna, A. Rachmadian, and DH Sutanto,

- "The Influence of Image of Tourist Destinations and Service Quality on Visitor Satisfaction at Tiga Warna Beach," *Competence J. Manag. Studs.*, vol. 15, no. 2, 2021, doi: 10.21107/kompetensi.v15i2.12512.
- [6] W. Hadi and A. Yulianto, "Implications of Tourist Attraction, Recommendations for Tourism and Destination Image for Tourist Visits to Sambisari Temple Tourism Object," *Khasanah Ilmu - J. Tourism and Culture*, vol. 13, no. 2, 2022, doi: 10.31294/khi.v13i2.13307.
- [7] RM Wirakusuma, "Character Analysis of International Tourists on Tourism Facilities in the Prawirotaman Area," *J. Tourism*, vol. 1, no. 1, 2014.
- [8] CHP Panjaitan, LJ Pangaribuan, and CI Cahyadi, "Analysis of the K-Nearest Neighbor Method Using Rapid Miner for the Recommendation System for Tourist Attractions in Labuan Bajo," *Remik*, vol. 6, no. 3, 2022, doi: 10.33395/remik.v6i3.11701.
- [9] F. Sologia, R. Aurachman, and PGK Artha, "Tourist Route Recommendations Using the Traveling Salesman Problem Method with the K-Nearest Neighbor Algorithm (Case Study: North Toraja)," *e-Proceedings Eng.*, vol. 7, no. 2, 2020.
- [10] R. Alamsyah, AD Wiranata, and R. Rafie, "Defect Detection in Ceramic Tiles Using the K-Nearest Neighbor Method," *Techno.Com*, vol. 18, no. 3, 2019, doi: 10.33633/tc.v18i3.2459.
- [11] NRA Candra DA, "Branding: Instagramable Videos for Pokdarwis Dharma Jati & the Thinthir Art Community in the Development of 'AB-S' (Nature-Culture-Spiritual) Tourism Villages in Anggrasmanis Village, Karanganyar Regency," *Servant of Arts*, vol. 11, no. 1, 2020, doi: 10.33153/abdiseni.v11i1.3128.
- [12] A. Pradana and R. Ridwansyah, "Classification of Thesis Topics Based on Meaning with Web Semantic Approach," *J. Comput. Apply.*, no. Vol. 7 No. 1 (2021), 2021, doi: 10.35143/jkt.v7i1.4603.
- [13] SF Mentari, AB Kusdinar, and GP Hartawan, "Indications of Forms of Sexual Deviance Using Web-Based K-Nearest Neighbor," *J. Technol. inform. and Comput.*, vol. 7, no. 2, 2021, doi: 10.37012/jtik.v7i2.624.
- [14] MS Fajri, N. Septian, and E. Sanjaya, "Evaluation of the Implementation of the K-Nearest Neighbors (kNN) Machine Learning Algorithm on Low Resolution Gamma Spectroscopy Data," *Al-Fiziya J. Mater. sci. Geophys. Instrum. Theor. Phys.*, vol. 3, no. 1, 2020, doi: 10.15408/fiziya.v3i1.16180.
- [15] SR Durugkar, R. Raja, KK Nagwanshi, and S. Kumar, "Introduction to data mining," *Data Mining and Machine Learning Applications*. 2022, doi: 10.1002/9781119792529.ch1.
- [16] B. Devipriya and Y. Kalpana, "Evaluation of sentiment data using classifier model in rapid miner tool," *int. J.Eng. Adv. Technol.*, vol. 9, no. 1, 2019, doi: 10.35940/ijeat.A1323.109119.
- [17] D. Novianti, "Implementation of the Naïve Bayes Algorithm on Hepatitis Data Sets Using Rapid Miner," *Paradig. - J. Comput. and Inform.*, vol. 21, no. 1, 2019, doi: 10.31294/p.v21i1.4979.