Segmentation of Tourist Interest on Tourism Object Categories by Comparing PSO K-Means and DBSCAN Method

Meri Chindyana*, Lili Ayu Wulandhari

Computer Science Department, BINUS Graduate Program-Master of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia

Corresponding Author Email: meri.chindyana@binus.ac.id

ABSTRACT

Marketing in travel companies will usually offer promos or recommendations regarding various categories of random tourist objects to their customers. The promo or recommendation contains categories of tourist objects that are frequently visited and had good ratings from many customers. However, because companies do not really know and understand the characteristics or interests of each customer, sometimes some promos do not match their interests so that they are not interested in taking the promos that are offered. There are already several papers that discuss tourism recommendations, but they only focus on 1 tourist spot or tourism object category. Based on these problems, this thesis is made to discuss the segmentation of tourist interest in tourism object categories by comparing the PSO K-Means method and the DBSCAN method, which is about recommendations for more specific tour packages according to rating. Characteristics or similar interests between 1 tourist and other tourists will be grouped into 1 cluster. From each cluster that is formed, it can make it easier for companies to know what categories of tourist objects each customer is interested in or like and be able to offer promos or recommendations for tour packages according to tourist interests.

1. INTRODUCTION

According to an article [1], traveling to a place is one of the most popular hobbies for many people. Traveling to a place is eagerly awaited when the holidays arrive because they can spend free time together with family or friends. Traveling to a place can be done overseas or within the country depending on the vacation plans that we have made. According to an article [2], there are 10 benefits that we can get when traveling to a place, namely increasing knowledge about a place, meeting new people, learning a new culture, reducing stress, learning to adapt, becoming a more independent person in all respects, more confident, more patient, practice communication skills and learn a new language. Currently, many means of transportation can be used to travel to a place, making it easier for tourists to travel.

According to an article [3], the marketing agency in a tour company will usually offer promotions or recommendations regarding various categories of tourist objects at random to their customers. The promo or recommendation contains categories of tourist objects that are frequently visited and have good ratings from many customers. However, because the company does not really know and understand the characteristics or interests of each customer, sometimes there are some promos that do not match their interests, so that they are not interested in taking the promos offered.

There have been several papers that discuss tourism recommendations, but only focus on 1 tourist spot or tourist attraction category: Restaurants [4, 5], Museums [6-10], Spa Resorts [11-13], Beaches [14], Parks [15], Nature-based Activities [16, 17], Hotels [18, 19], Movies [20, 21].

Based on research that has been done by previous researchers, it is known that recommendations consist of only one destination, usually, tourists visit several places in 1 session of visiting tourist attractions. Therefore, the author plan to segment tourist interest in the tourist attraction category by comparing the K-Means method, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), and applying the Particle Swarm Optimization (PSO) method to optimize the clustering method.

This segmentation will discuss more specific tour package recommendations according to rating. Characters or similar interests between 1 tourist and other tourists will be grouped into 1 cluster. From each cluster that can be formed, it can make it easier for companies to see which categories of tourist objects are of interest or. Each tourist can offer promos or travel package recommendations according to their interests. For example, in cluster 0, 15 tourists like beaches, restaurants, and malls. So, the promo or tour package recommendation given by the company to the 15 tourists is more directed at the 3 tourism object categories above and other tourist objects that don't need to be offered.

The objective to be achieved is to develop a clustering model of tourist interest using the K-Means and DBSCAN methods based on the review of tourist destination rating data, analyze the character similarities of each cluster obtained against the recommended tourist destinations, and determine the optimization method to be applied to the K-Means method in the case of the tourist interest segmentation.

The benefit to be achieved is to see and understand the interest in the tourism object category of each tourist so that it makes it easier for the marketing party in a travel company to

offer a tour or vacation package appropriately, it can increase the profits of a company because it increases the interest of tourists in using their services, adds insight regarding the methods used for clustering, determines the appropriate optimization method for the K-Means method, and can analyze the similarities of each formed cluster.

The problem that will be discussed in this paper is the methods used for clustering are Partitional and Density-Based methods. The Partitional and Density-Based methods used are the K-Means and DBSCAN methods, the method used for optimization is the PSO (Particle Swarm Optimization) method, this data only contains information about tourism object categories and is not equipped with the names of tourist objects in European countries so that the rankings are only general in nature, for example, the overall restaurant in European countries is given an average rating of 3.42 by tourist A, and this data only discusses the European region.

Therefore, the author using the clustering method to determine the criteria/interests of each tourist to be able to offer promos or travel package recommendations according to tourist interests. Characters or similar interests between 1 tourist and other tourists will be grouped into 1 cluster. Each of the clusters that can be built can make it easier for companies to see what categories of tourist objects each tourist is interested in or made. The author read several papers and found that the clustering methods used for the segmentation technique are K-Means and DBSCAN. Therefore, the author comparing the comparisons of the two methods in the case of segmentation of tourist interest to the categories of tourist objects seen by which are the best. From several papers that the author has read, each optimization technique has different uses. In the case of segmentation of tourist interest in this tourist attraction category, the author chooses to use the PSO technique.

2. RELATED WORKS

There are several papers that discuss segmentation techniques. The following is a summary of each paper:

Using a design science research approach, this study aims to design and evaluate a 'big data analytics' method to support strategic decision-making in tourism destination management. The grouping technique, named P-DBSCAN, is applied to geographical data to identify popular areas of interest. This take into an account the number of photos and the number of tourists, which ensures that the location identified actually has a lot of tourists who have been visited for certain interests. The advantages of P-DBSCAN have been demonstrated in recent research to identify locations visited by tourists. K-means clustering is applied to build visual vocabulary words. Visual words are defined as the center of the group, and the value of k determines the number of available visual words [22].

A multinational comparative study highlighting students' travel motivations and touristic trends. Data were analyzed using two main component analyzes (PCA), a combination of two grouping methods: Ward Method, and the optimal solution method, K-Means method. Seven clusters based on tourist attractions / activities emerged namely explorers, soft explorers, tourists, novelty seekers, evaders, function seekers, and tourism lovers. Findings from this study indicate that perceptions of tourist attractions / activities differ by country although some similarities do exist [23].

Big data in tourism research: A literature review. Regarding centroid-based methods, K-Means is a popular algorithm in tourism research. Such a centroid-based partitioning method has one obvious disadvantage that requires in-depth search to find the best cluster centers, which are inefficient in a data-rich environment; conversely, density-based grouping may be better suited for large data photos, which requires minimum domain knowledge and effectively filters outlaws even in the presence of noise points. Therefore, the clustering of density-based spatial clustering of applications with noise (DBSCAN) and its variants has been developed and widely applied to the grouping of tourism photographs. To build cluster hierarchies, connectivity-based clusters (also known as Hierarchical Clustering), flexible and fast algorithms based on inequality matrices apart from metric spaces have also been introduced to detect tourist attractions and group cities visited by the same tourists [24].

Market analysis of value-minded tourists: Nature-based tourism in the Arctic. Cluster analysis was carried out to classify the participating tourists using hierarchical grouping with Euclidean distance squared and Ward linkage. This reveals a three-cluster solution for the best interpretation, and the K-means cluster algorithm procedure is used to classify three tourism segments based on perceived value dimensions [17].

Network approach to tourist segmentation via user generated content. To reduce dimensions, they join keywords that tend to appear along with hierarchical cluster analysis. The average method of grouping hierarchical hierarchies in groups is done using the agglomerative (bottom-up) and phi linkage approaches (Pearson analog correlation for binary data) [25].

Travelers' use of social media: A clustering approach. Cluster analysis includes two phases. First, three hierarchical algorithms are applied, namely, complete linkage, average linkage and Ward's method [26].

Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. K-means is one of the most popular algorithms used in cluster analysis, mostly in the field of data mining, and statistical data analysis. To conduct cluster analysis in this study, cluster numbers must be explored using SOM for later use in the K-means algorithm [27].

There are several papers that discuss optimization methods. The following is a summary of each paper:

Optimal cluster analysis using hybrid K-Means and Ant Lion Optimizer. K-Means is a popular cluster analysis method that aims to partition a number of data points into a K cluster. This has been successfully applied to a number of problems. However, the efficiency of K-Means depends on the initialization of the cluster center. Different swarm intelligence techniques are applied to grouping problems to improve performance. In this work a hybrid clustering approach based on K-means and Ant Lion Optimization has been considered for optimal cluster analysis. Ant Lion Optimization (ALO) is a stochastic global optimization model [28].

A Modified Bee Colony Optimization (MBCO) and it's hybridization with k-means for an application to data clustering. In this paper, the Modified BCO (MBCO) approach is proposed for data grouping. In the proposed MBCO, the remission characteristics of bees and provide fair opportunities for trustworthy and untrustworthy bees are being treated. To validate the proposed algorithm, seven standard data sets are considered. From the calculation of the percentage of

misclassification, it is observed that the proposed algorithm performs better than some existing algorithms. The simulation results conclude that the proposed algorithm can be used efficiently for grouping data [29].

Feature Selection using K-Means Genetic Algorithm for Multi-objective Optimization. In the previous paper, multi-objective optimization on learning environments using the k-means genetic algorithm (NLMOGA), was proposed and applied to several sets of real life in data. In NLMOGA, solutions are chosen from global population repositories and then learning environments are created to promote the evolution of each goal for the chosen solution. The effectiveness of this approach is evaluated with various real life in benchmark gene expression data sets [30].

Clustering Using a Combination of Particle Swarm Optimization and K-means. This paper applies a combination of particle crowd optimization and K-means for grouping data. The proposed approach tries to improve the performance of traditional partition grouping techniques such as K-means by avoiding the initial requirements of the number of clusters or centroids for grouping. The proposed approach is evaluated using a variety of primary and real-world datasets. In addition, this paper also presents a comparison of the results produced by the proposed approach and with K-means based on the validity steps of the grouping such as inter and cluster distance, quantization error, silhouette index, and Dunn index. Comparison of results shows that as the size of the dataset increases, the proposed approach results in a significant increase in the K-means partition grouping technique [31].

Hybrid K-Means and Improved Self-Adaptive Particle Swarm Optimization for Data Clustering. The K-Means algorithm is currently one of the most popular grouping techniques, because of its simplicity and scalability. However, K-Means performance is strongly influenced by the choice of the initial cluster center, which can lead to suboptimal solutions. In this paper, a novel hybrid clustering grouping algorithm is proposed, named IDKPSOCk, based on an automatically adjustable increase in Particle Swarm Optimization (PSO) and K-Means, which uses a crossover operator to increase the PSO's ability to escape from the local minimum point from the problem room. To evaluate the performance of the proposed approach, experiments were carried out on sixteen benchmark data sets obtained from the UCI Machine Learning Repository. The experimental evaluation, which was carried out using the Friedman hypothesis test in relation to four clustering metrics, has shown the effectiveness of the proposed model in relation to the comparison algorithm [32].

Hybrid Particle Swarm Optimization and K-Means Analysis for Bridge Clustering Based on National Bridge Inventory Data. In this paper, an optimization approach based on a hybrid of metaheuristic particle crowd optimization and k-means method (KPSO) is presented in grouping the data. The aim is to group bridges with similar structural deficiency attributes by minimizing the number of quadratic errors associated with assigning data points to each cluster and determining the most appropriate number of clusters. The approach presented is compared with the basic version of swarm particle optimization (PSO) and the traditional k-means clustering method. The algorithm was tested using the National Bridge Inventory (NBI) database. The results show that KPSO provides better results in terms of objective functions and shows the opportunity to apply optimization techniques for data analysis in civil infrastructure systems [33].

The following conclusions from the literature study that have been carried out:

The author does not use a collaborative filtering method because this method requires profiles from other tourists, meaning that it recommends an item based on the suitability of one tourist profile with another tourist profile. For example, most tourists who like product A also like product B, so if other tourists like product A, we can recommend product B. While the content-based filtering method requires a profile of an item and a history of its activities on these items such as ratings or like or dislike. Requiring profile means requiring features or characteristics of an item such as theme, genre, author, year published, location, or place name. However, in the case of the segmentation of tourist interest in this tourist attraction category, there is no match for the profile of one tourist with another and there are no features or characteristics as above. Therefore, the author using the clustering method to determine the characteristics/interests of each tourist to be able to offer promos or travel package recommendations according to tourist interests. Characteristics or similar interests between 1 tourist and other tourists will be grouped into 1 cluster. From each of the clusters that are formed, it can make it easier for companies to know what categories of tourist objects each tourist is interested in or likes.

The author reads several papers and finds the clustering methods used for the segmentation technique are K-Means and DBSCAN. Therefore, the author conducted a comparison of the two methods in handling the case of segmentation of tourist interest in the category of tourist objects to find out which method was the best.

From several papers that the author has read, each optimization technique has different uses. In the case of segmentation of tourist interest in this tourist attraction category, the author chooses to use the PSO technique.

3. RESEARCH METHODS

The author needs to use the K-Means method for the case of Tourist Interest Segmentation in the Tourism Object Category because it is the best clustering method according to several papers. After that, the author conducted a comparison to determine which is the best clustering method between DBSCAN and K-Means for this case.

The research stages or steps are taken while compiling this thesis are to identify problems, collect data, process data, study literature, choose the clustering method, perform clustering using the K-Means and DBSCAN methods, analyze the clusters formed, evaluate the performance of the method used, select and apply the optimization method.

3.1 Development of the clustering model

In the process of developing the DBSCAN clustering model, the author took several steps based on Figure 1 and Figure 2.

In the process of developing the K-Means clustering model, the author took several steps based on Figure 3 and Figure 4.

The Euclidean Distance formula used by the DBSCAN and the K-Means clustering method based on Figure 2 and Figure 4 is like Eq. (1):

$$d(x_i, \mu_j) = \sqrt{\sum (\mu_j - x_i)^2} \qquad (1)$$

where, d=distance, x_i=the ith point of density reachable to 1 centroid point, i=index of point, μ_j=centroid in jth cluster, j=index of cluster.

3.2 K-Means optimization with PSO

The rationale for using PSO to optimize K-means was to make the performance of the DBI Coefficient from this method better than DBCSAN. Because the higher value performance of the DBI and Silhouette Coefficient shows the better result for clustering. The advantage of this PSO technique is can overcome the weaknesses of the traditional clustering methods, namely the selection of the initial cluster centers. In this stage, what author want to optimize with PSO is the initial centroid point in the K-Means method because one of the weaknesses of the K-Means method is the initial centroid point chosen randomly will give poor clustering results. With the optimization technique through PSO, the initial centroid point for the K-Means method is no longer determined randomly.
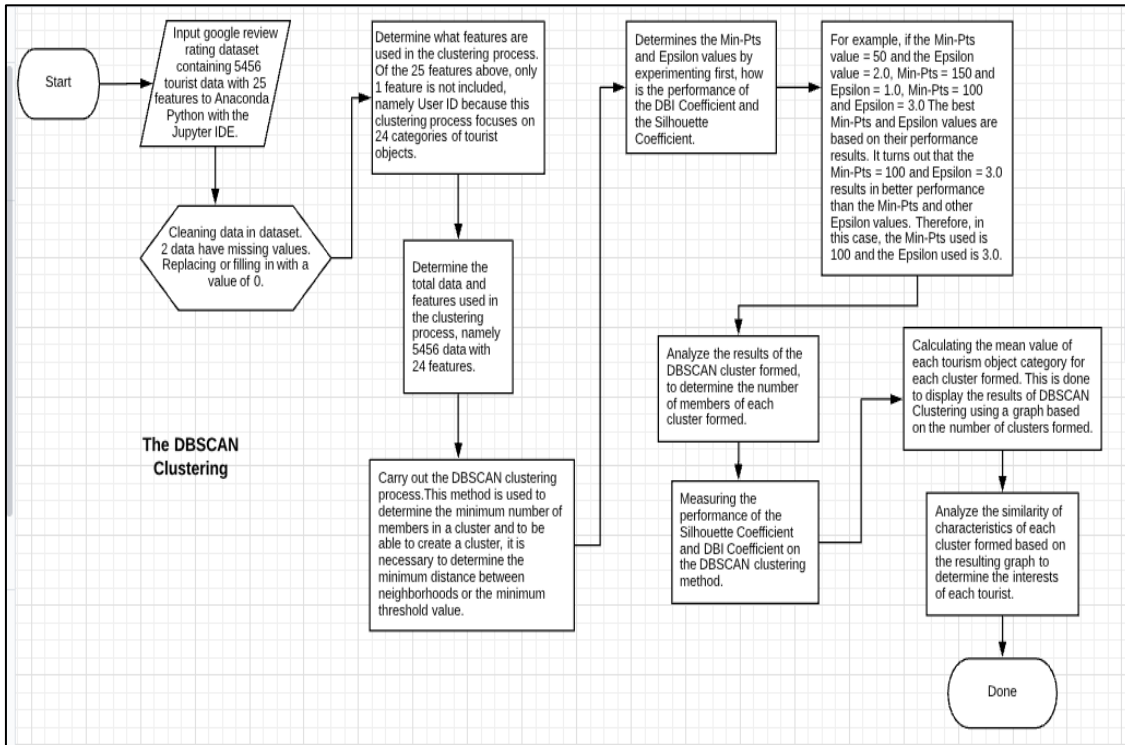

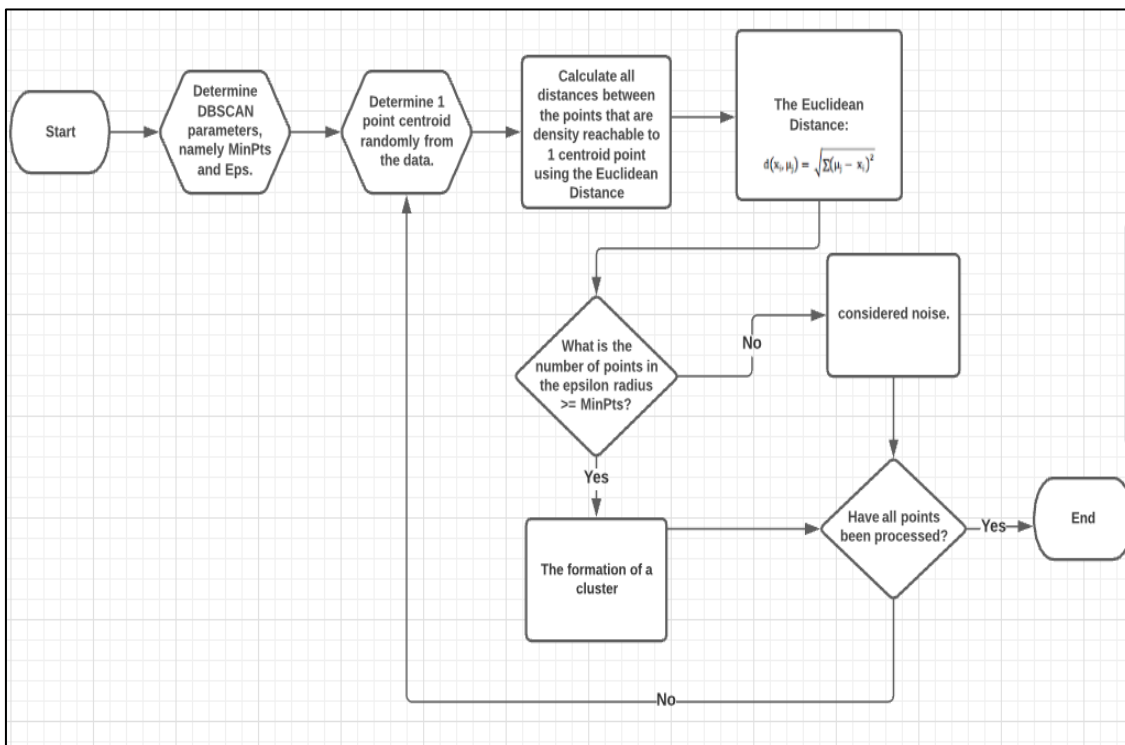**Figure 1.** The DBSCAN clustering method

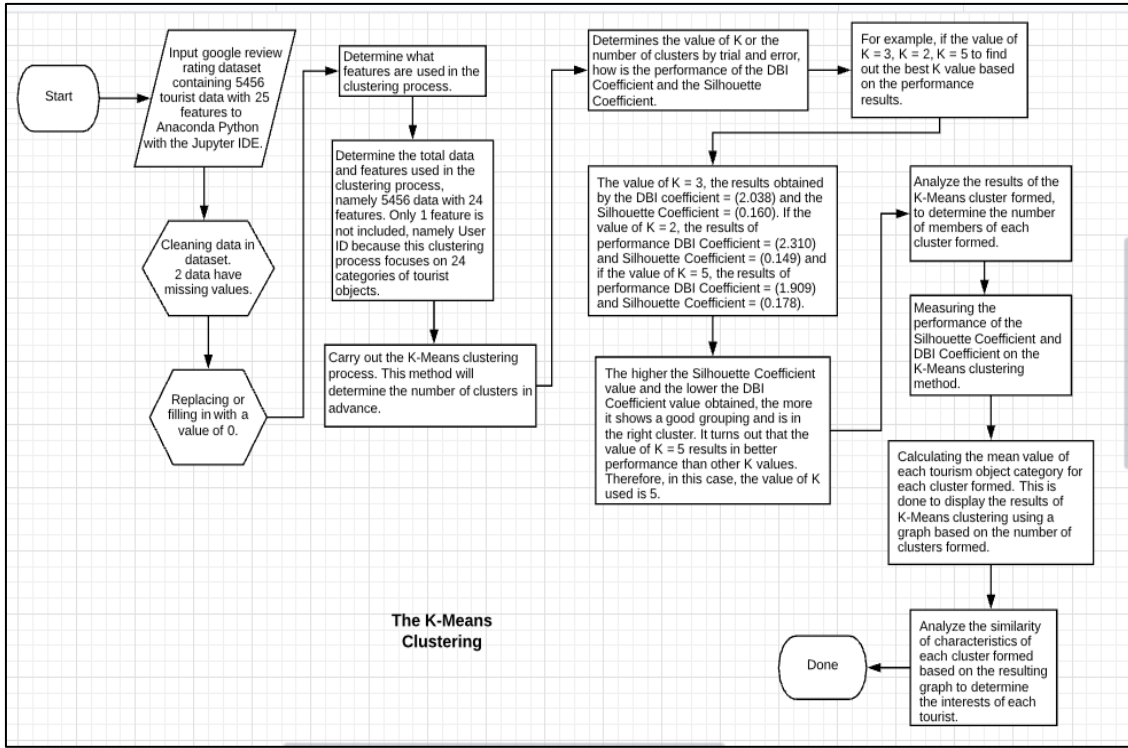**Figure 2.** Flowchart the DBSCAN clustering method
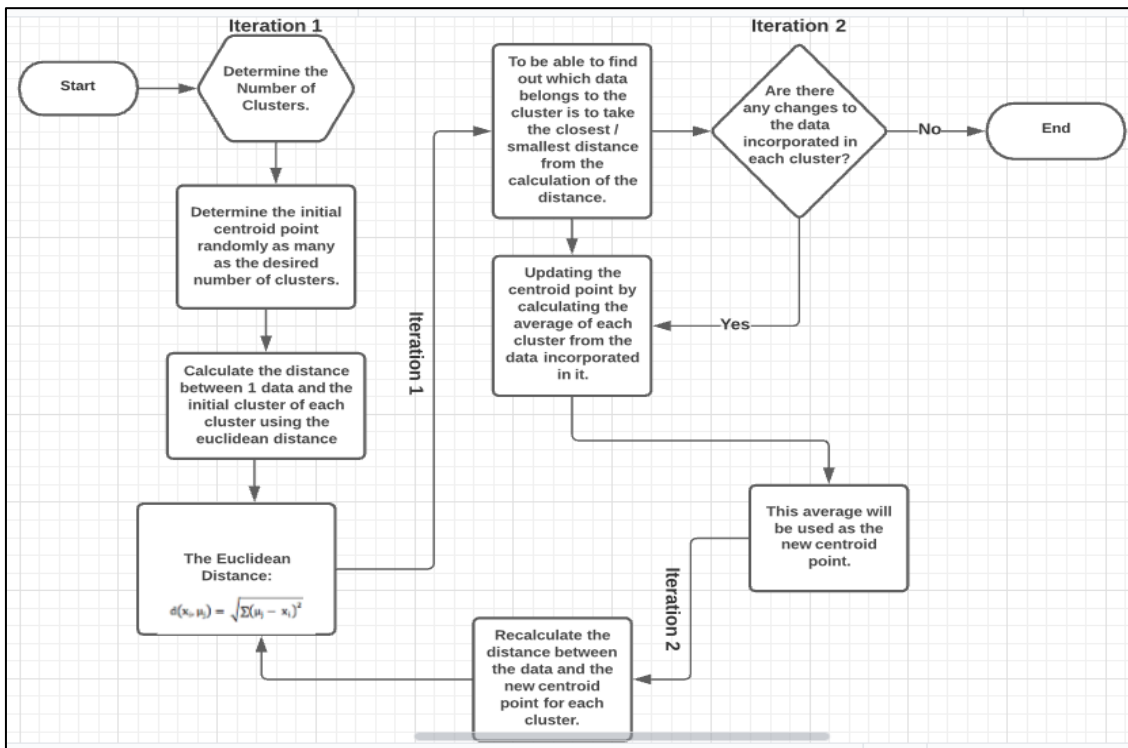
Figure 3. The K-Means clustering method

Figure 4. Flowchart the K-Means clustering method

PSO works to find a candidate solution, where the candidate solution is a solution for centroid, therefore the particle parameter in PSO contains centroid. These solution candidates will be the starting centroid points at K-Means. The n-cluster on K-Means is 5, so there will be 5 centroids. The author determines the learning rate c1= 0.5 and c2= 0.3, the weight of inertia (w) = 0.9, and the maximum iteration = 300. The steps for K-Means optimization with PSO are as follows based on Figure 5:

The formula used to update the speed and position in the PSO K-Means clustering method based on Figure 5 is like Eqns. (2), (3):

$$v_i(t) = wv_i(t-1) + c1r1(x_{pi} - x_i) + c2r2(x_{gi} - x_i) \quad (2)$$

$$x_i(t) = x_i(t-1) + v_i(t) \quad (3)$$

where, i=particle index, t=iteration, w=inertia weight, v_i=velocity of the ith particle, x_i=position of the ith particle, x_{pi}=best position of the i-th particle, x_{gi}=best position of all particles, $c1$, $c2$=constant learning rate, the value is between 0-1, $r1$, $r2$=random numbers between 0-1.

According to the study [24] How to determine the best parameters in the PSO method is:

(1) Number of Particles: For better results, it is recommended that the number of particles used is 10. The ranges are between 20-40.

(2) Max velocity (max. Velocity): this parameter is set for the displacement of the particles. If the speed of vi is between -10 and 10, the max speed is 20.

(3) Learning Rates: Generally, the values of c1 and c2 are 2. Different problems, different values. The range is between 0-4.

(4) Inertia Weight: The range is between 0.8-1.2

(5) Stop condition: This can happen when the maximum number of iterations of the PSO method and the minimum error requirement has been reached.

The author can determine this hyperparameter from the results of previous experiments if the learning rates are c1 and c2=1.49, the weight of inertia (w)=0.72, and the maximum iteration = 300. From the experimental results, it turns out that the hyperparameters that produce the best clustering results are in the following hyperparameters: learning rate c1=0.5 and c2=0.3, inertia weight (w)=0.9, and maximum iteration=300.

Therefore, the authors apply these parameters: learning rate c1=0.5 and c2=0.3, inertia weight (w)=0.9, and maximum iteration=300. The number of particles to get a better optimization result for the PSO initialization is 10. Each particle contains 5 centroids because there are 5 clusters. This is what is called a candidate solution. There are 10 candidate solutions in the case of segmentation of tourist interest.
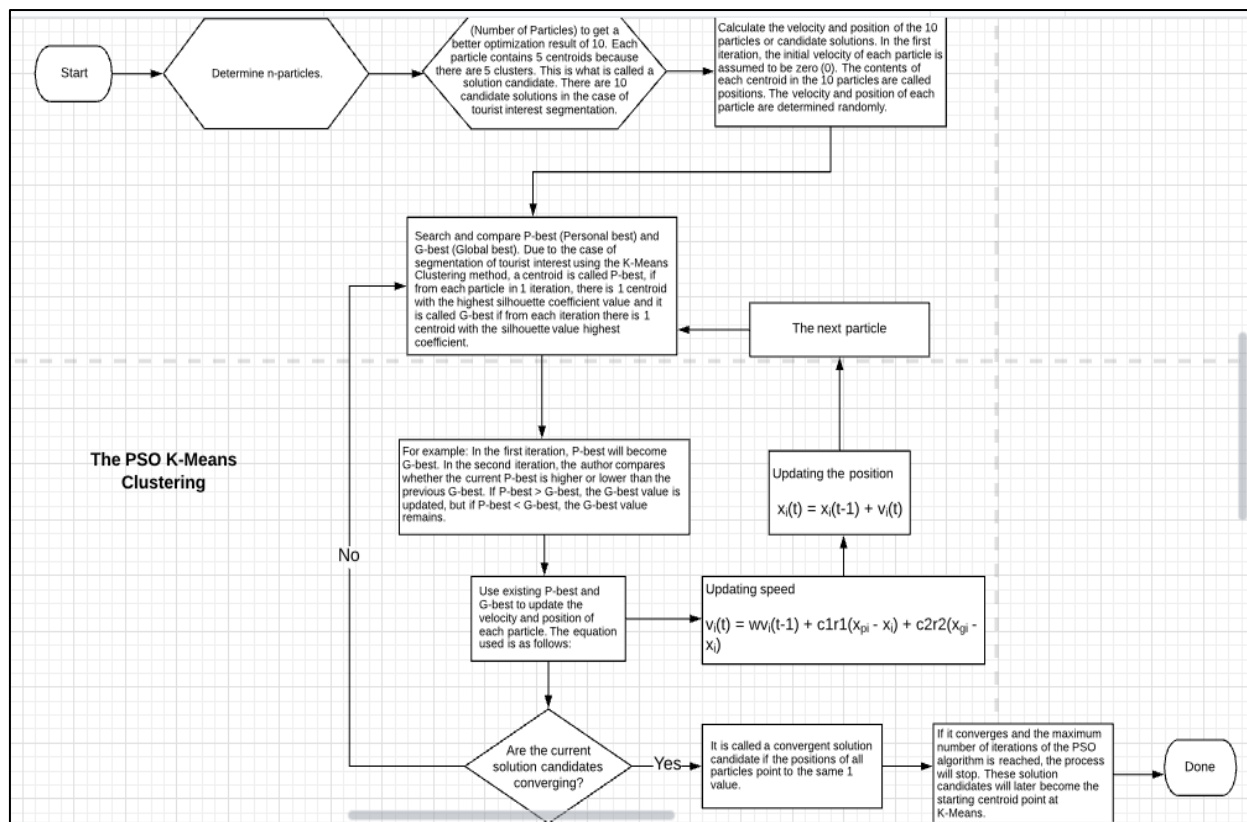

Figure 5. The PSO K-Means clustering method

4. RESULTS AND DISCUSSION

In this stage tell about discussing the performance of the K-Means and DBSCAN methods, then compare the 2 methods to find out which method has better performance, and also discussing the performance of the K-Means PSO method. The experimental results show that the performance of the Silhouette Coefficient from the K-Means Clustering method is better than DBSCAN, but the DBI Coefficient performance from the DBSCAN Clustering method is better than K-Means.

Because of that, the author optimizes the K-Means method with the PSO optimization technique and then compares it again with the DBSCAN method to find out which is the best clustering method between these two methods. The result is the PSO K-Means method shows better results toward both the

Silhouette Coefficient and the DBI Coefficient than DBSCAN due to the PSO technique can optimize the performance of the K-Means method.

4.1 Performance results of the K-Means method

In the K-Means Clustering method, performance results are:
(1) Silhouette Coefficient (0.178).
(2) DBI Coefficient (1,909).

4.2 Performance results of the DBSCAN method

In the DBSCAN Clustering method, performance results are:
(1) Silhouette Coefficient (-0.054).
(2) DBI Coefficient (1,408).

4.3 Results of the clusters formed in the K-Means method and PSO K-Means

The results in K-Means and PSO K-Means Clustering are 5 clusters with the following number of tourists:

Table 1. K-Means and PSO K-Means cluster results

Formed Clusters	Number of Tourists
Cluster 0	1183
Cluster 1	892
Cluster 2	748
Cluster 3	1690
Cluster 4	943
Total= 5Cluster	

From the results of the cluster division, Table 1 show that there are 1183 tourists included in cluster 0, 892 tourists are included in cluster 1, 748 tourists are included in cluster 2, 1690 tourists are included in cluster 3, 943 tourists are included in cluster 4.

4.4 Results of the cluster formed in the DBSCAN method

The results obtained in DBSCAN Clustering are the formation of 5 clusters with the following number of a tourist.

From the results of the cluster division, Table 2 show that there are 4790 tourists included in cluster -1, 124 tourists are included in cluster 0, 270 tourists are included in cluster 1, 170 tourists are included in cluster 2, 102 tourists are included in cluster 3.

Table 2. DBSCAN cluster results

Formed Clusters	Number of Tourists
Cluster -1	4790
Cluster 0	124
Cluster 1	270
Cluster 2	170
Cluster 3	102
Total=5 Cluster	

4.5 Comparative analysis of performance 2 clustering methods

The comparison seen in terms of performance for the methods, namely K-Means and DBSCAN, is as follows:

Table 3. Comparison performance of 2 clustering methods

Clustering Method	Performance Without Optimization	
	Silhouette Coefficient	DBI Coefficient
DBSCAN	- 0.054	1.408
K-Means	0.178	1.909

Table 3 shows that the K-Means method produces Silhouette Coefficient better than the DBSCAN method. The DBSCAN method produces DBI Coefficient better than the K-Means method. The clustering method was showed better performance if it had the highest value of Silhouette Coefficient and the smallest value of DBI Coefficient. The suitable clustering method for the segmentation of tourist interest is the K-Means method.

4.6 Results of the similarity of character analysis K-means and PSO K-means method

Based on Figure 6 and Figure 7, the analysis results of similar characters in the K-Means and PSO K-Means methods are as follows:
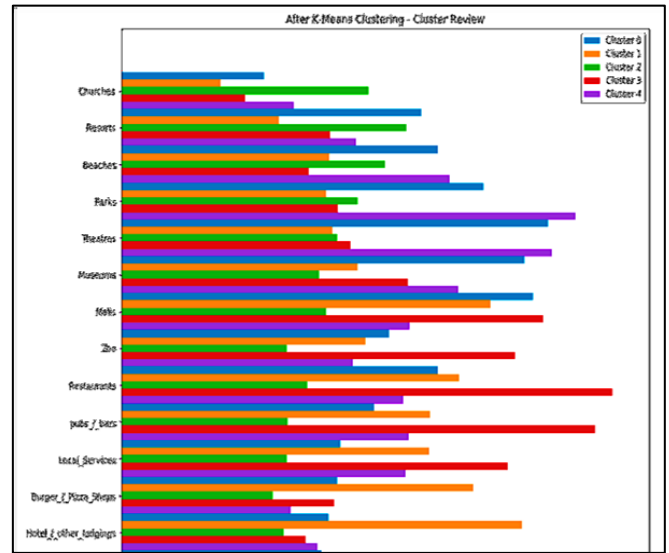

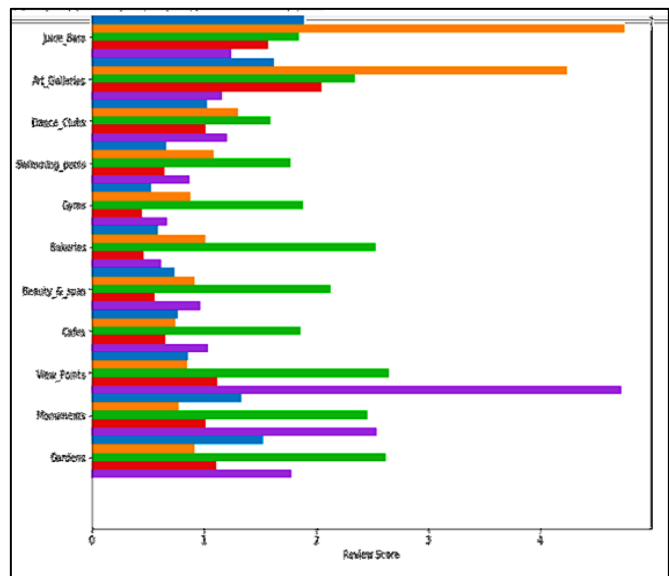**Figure 6.** K-Means and PSO K-Means cluster review graph

Figure 7. Continued K-means and PSO K-means cluster review graph

The way to analyze the clusters that were formed based on these 2 graphs above is to find out which one tourist is giving the best rating in the tourist object category. After that, need to know which one suitable cluster for each tourist object category as in Table 4: in Figure 6 and Figure 7, categories of church tourist objects, have 5 colors, namely blue, orange, green, red, and purple. The green color is more dominant than other colors, the green color shows the color for cluster 2, which means that tourists in cluster 2 prefer and give the best rating for churches, therefore churches are more suitable to be included in cluster 2. This example also applies to other the tourist attraction category.

Table 4. The results of the analysis of K-Means and PSO K-Means cluster grouping for 24 categories of tourist objects

No.	Features	Cluster
1.	Churches	2
2.	Resorts	0
3.	Beaches	4
4.	Parks	4
5.	Theatres	4
6.	Museums	0
7.	Malls	3
8.	Zoo	3
9.	Restaurants	3
10.	Pubs / Bars	3
11.	Local Services	3
12.	Burger / Pizza shops	1
13.	Hotel / Other Lodgings	1
14.	Juice Bars	1
15.	Art Galleries	1
16.	Dance Clubs	2
17.	Swimming Pools	2
18.	Gyms	2
19.	Bakeries	2
20.	Beauty and Spas	2
21.	Cafes	2
22.	View-Points	4
23.	Monuments	4
24.	Gardens	2

After it is known that all the categories of tourist objects are included in which cluster, the author analyze the similarities of the characters from the cluster formed. For example, in cluster 0 it is found that tourists in the cluster prefer resorts and museums. According to the author's opinion, tourists in cluster 0 have characters who like to take pictures, like to get unique experiences from the facilities provided by European tourist objects, like to go to tourist objects at affordable prices.

Then, the author checks through the article whether it is true that tourists with characters like the ones above enjoy visiting resorts and museums. This article is external data that support the author's opinion regarding the character of tourists in cluster 0. After the author know the characteristics or interests of each tourist, the author finding out the names of places and cities for each category of tourist objects that are preferred and following the characteristics of tourists from each cluster through articles with the aim that the names of places and cities can be used as recommendations for tourist attractions which will later be offered to tourists. Based on the above analysis, it is found that:

1. Cluster 0 --> blue color is called the MuRe Cluster
   a. Resorts
   b. Museums - Europe's most loved tourist attraction
   The analysis obtained from the MuRe cluster is:
   a. The characteristics of tourists who like to take pictures or commonly referred to as photography and prefer to go to European tourist objects that have interesting photo spots.
   b. Like to get a unique experience from the facilities provided by European tourist attractions. For example, in a resort you can experience swimming in a swimming pool that directly leads to the beach or mountain, usually called an infinity pool. While the museums can add insight and information about the history, art, archaeological discoveries, museums building architecture that is truly luxurious and magnificent.

   c. Like to go to European tourist attractions at affordable, comfortable prices, the trip to the station is not too far and has satisfying facilities and services so that it can eliminate fatigue from daily activities.

According to an article [34], tourists in cluster 0 have a classic type of personality, that is, they have a regular personality, even for their holidays. They weren't the kind of explorer who went on vacation somewhere without a plan. They like to get to know local beliefs and culture, as well as visit historical places.

According to an article [35], a tourist who likes photography indicates that he has a philosophical side in him because it is not easy to express something, but he always pays attention to aesthetics. They are not messy and have a strategic way of thinking.

According to the article [36], museums in Europe that suit the characteristics of tourists who like to take pictures and have unique experiences are

1. State Hermitage Museum - Russia
2. Musee d'Orsay - Paris
3. Prado National Museum - Madrid, Spain
4. Vasa Museum - Stockholm, Sweden
5. Louvre Museum - France

According to an article [37], resorts and museums in Europe that suit the characteristics of tourists who like to go to tourist attractions at affordable prices are

1. British Museum - London
2. Luxembourg City Youth Hostel - Luxembourg City
3. Hotel du Parlement - Brussel
4. Ibis Amsterdam Center Stopera - Amsterdam
5. Christopher's Inn Gare du Nord - Paris

According to an article [38], resorts in Europe that suit the characteristics of tourists who like to have unique experiences such as infinity pools are

1. Hotel Villa Honegg - Switzerland
2. Lefay Resort - Italy
3. Dionysos Hotel - Turkey
4. Katikies Hotel - Greece
5. Perivolas Luxury Hotel – Greece

2. Cluster 1 --> orange color is called the AJHPB Cluster.
   a. Juice Bars - Europe's most preferred tourist attraction
   b. Art Galleries
   c. Hotels / Other lodgings
   d. Burger / Pizza Shops
   The analysis obtained from the AJHPB cluster is
   a. The characteristics of tourists who like to go to European tourist objects that provide food and drinks that are delicious and healthy for the body, such as juices, smoothies, burgers, pizza, vegetables.
   b. Love to go to European tourist attractions that can spend hours with family and friends such as reunions, birthdays, meetings with family and friends with a relaxed atmosphere equipped with facilities such as wi-fi, air conditioning, attractive and modern views, clean environment.

According to an article [34], tourists in cluster 1 have an artistic type of personality and a harmonious type, namely, they really like all things about art and love peace. In their vacation, tourists want to find peace or peace of mind by going to a place that can pamper them, enjoying time alone in peace. The mindset of artistic types is always out of the box and their perspective is different from most people. This artistic type really enjoys an unconventional, unconventional lifestyle to maximize his vacation. They like to have a vacation to a place

that is inhabited by many artists who can also use it to exchange ideas.

According to an article (googlereview), Juice Bars in Europe that match the characteristics of tourists above are
1. Sama Sama Crepe and Juice Bars - Lisbon, Portugal
2. JOIN Juice Bars - Athens, Greece
3. BOB's Juice Bars - Paris, France
4. Big Juice Ltd - Bristol, UK
5. The Cold Pressed Juicery - Amsterdam, London

According to an article [39], the Burger / Pizza Shops in Europe that match the characteristics of the tourists above are
1. Burger and Beyond - London, England
2. Gasoline Grill - Copenhagen, Denmark
3. Patty and Bun - London, England
4. Barrels Burgers & Beer - Stockholm, Sweden
5. Burger Bar - Amsterdam, London

According to an article [38], Hotels / Other lodgings in Europe that match the characteristics of tourists above are
1. Elounda Gulf Villas and Suites - Greece
2. Cavo Tagoo - Greece
3. Hotel Caruso - Italy
4. The Cambrian Adelboden - Switzerland

3. Cluster 2 --> green color is called the 2B2G2CSD Cluster.
   a. Churches
   b. Dance Clubs
   c. Swimming Pools
   d. Gyms
   e. Bakeries
   f. Beauty and Spas
   g. Cafes
   h. Gardens - Europe's most preferred tourist attraction

The analysis obtained from the 2B2G2CSD cluster is
   a. The characteristics of tourists who like to go to European tourist objects that have many useful activities, learn interesting things and can channel their respective hobbies during the trip, such as diligent worship, sports, swimming, culinary bread and sitting relaxed in cafes while doing assignments, clubbing, caring skin and body, hang out with a beautiful view, nice and attractive like seeing beautiful flowers in Europe.
   b. Like hanging out with anyone, especially new people, so that he can make many friends.

According to an article [34], tourists in cluster 2 have a party and socialite type personality, that is, they are very happy to party and socialize. They are more comfortable when they are outside than at home. This type has an extrovert nature, which is happier to meet new people, then hang out, and dance all night long.

According to an article [40], Gardens in Europe that are following the characteristics of tourists whose hobbies are hanging out with beautiful, nice and interesting views such as seeing beautiful flowers is
1. Keukenhof Gardens - London
2. Hampton Court - UK
3. Hallerbos - Belgium
4. Mainau - Germany

According to an article [41], dance clubs in Europe are in line with the characteristics of the traveler whose hobby is clubbing
1. Fabric - London
2. Berghain Panorama Bar - Berlin
3. Air - Amsterdam
4. Concrete - Paris

5. CDLC (Carpe Diem Lounge Club) - Barcelona

According to an article [42], Cafes in Europe which is following the characteristics of tourists whose hobbies are sitting relaxed in cafes while doing assignments is
1. Wanda Cafe Optimista - Madrid, Spain
2. Coffee & - Ukrainian
3. Fudisiac - Bucharest, Romania
4. Family Cafe - Wroclaw, Poland
5. Coffee Cake - Nizhniy Novgorod, Russia

According to an article [43], Beauty and Spas, Swimming Pools in Europe are following the characteristics of tourists whose hobbies are swimming, caring for skin and body.
1. Alpentherme - Central Austria
2. Hotel Heliopic - Chamonix, France
3. Eagles Palace - Halkidiki, Greece
4. The Chedi - Montenegro

According to an article [44], Bakeries in Europe are following the characteristics of tourists whose culinary hobby is bread food.
1. Bakery House - Roma Italia Konditorei
2. Confiserie Sprungli - Zurich Switzerland
3. Kaffee Schneller - Munich Germany
4. Choux A La Creme Popelini - Paris France
5. My Little Patisserie - Amsterdam

According to an article [45], Churches in Europe are following the characteristics of tourists whose hobby is diligent worship
1. St. Basil's Cathedral - Moscow, Russia
2. Sagrada Familia - Barcelona, Spain
3. Kölner Dom - Cologne, Germany
4. Duomo - Milan, Italy

According to an article [46], Gyms in Europe are following the characteristics of tourists whose hobby is sports
1. La Belle Équipe - Paris
2. Das Gym - Vienna
3. Third Space - London

4. Cluster 3 --> red color is called the RPMZL Cluster.
   a. Restaurants - Europe's most preferred tourist attraction
   b. Pubs/bars
   c. Malls
   d. Zoo
   e. Local Services

The analysis obtained from the RPMZL cluster is characteristics of tourists who like to shop and are willing to spend more money at the end of the month for sightseeing, refreshing, night entertainment, culinary food and drinks, and love animals and want to see the most diverse collection of animals in the world.

According to an article [34], tourists in cluster 3 have a luxurious type personality, namely, they already know what kind of trip they like and want to taste the taste of a luxurious vacation.

According to an article [47], tourists who like to go to zoos are people who love animals. They are warm and love to share with others. They also get the same love from those around them.

According to an article [48], restaurants in Europe that suit the characteristics of tourists who like culinary food and drinks are
1. Restaurant Frantzén - Sweden
2. Schloss Schauenstein - Switzerland
3. Asador Etxebarri - Spain
4. Le Calandre - Italy

5. L'Arpège - France

According to an article [49], pubs/bars in Europe that are following the characteristics of tourists who like nightlife are
1. Bodega Bellver - Palma in Mallorca, Spain
2. Oldenburg - Madrid, Spain
3. Brauerei Neder - Forchheim, Germany
4. The Mayflower - London, England

According to an article [50], Malls and Local Services in Europe which are following the characteristics of tourists who like shopping, sightseeing and refreshing are
1. Galleries Lafayette - Paris, France
2. Galleria Vittorio Emanuele II - Milan, Italy
3. Harrods - London, England
4. Department store Quartier 206 - Berlin, Germany

According to an article [51], Zoos in Europe that suit the characteristics of tourists who like animals and want to see the most diverse collection of animals in the world are
1. Basel Zoo - Switzerland
2. Edinburgh Zoo - Scotland
3. Vienna Zoo - Austria
4. Berlin Zoo - Germany
5. Prague Zoo - Czech Republic
6. Cluster 4 --> purple color is called the TVBcPaMo Cluster.
   a. Beaches
   b. Parks
   c. Theaters
   d. View-Points - Europe's most loved tourist attraction
   e. Monuments

The analysis obtained from the TVBcPaMo cluster is

a. The characteristics of tourists who like to go to European tourist objects that have many visitors and have showing of interesting attractions.

b. Love to go to European tourist attractions with cool and nice views.

c. Like to go to tourist objects that can keep memories or beautiful memories with family and loved ones.

d. Likes to do activities that challenge adrenaline such as hiking and paragliding to be able to see a good view of a tourist attraction.

According to an article [34], tourists in cluster 4 have a personality type full of challenges and lovers of nature, namely motivated by their curiosity about new things and really love all things related to nature. They really like challenges and explore every area even if the terrain is tough. They also have a very sensitive and conscientious nature. They are the type of people who really care about other people and everything around them including the beauty of nature.

According to an article [52], Beaches in Europe that match the characteristics of the tourists above are
1. Our Lady Beach - Porquerolles, France
2. Navagio Beach - Zakynthos, Greece
3. Zlatni Rat Beach - Brac, Dalmatia, Croatia
4. Touts des Moro - Mallorca, Spain
5. Oludeniz Beach - Turkey

According to an article [53, 54], Parks in Europe that match the characteristics of tourists above are
1. Disneyland Paris - France
2. Europa Park - Germany
3. Efteling - London
4. Parque de Maria Luisa in Seville - Spain
5. Lycabettus Hill in Athens - Greece

According to an article [55], theaters in Europe that match the characteristics of tourists above are
1. The Volkstheater - Vienna, Austria

2. National Theater - Prague, Czech Republic
3. Juliusz Slowacki Theater - Krakow, Poland
4. The Globe Theater - London, England
5. Bastille Opera House - Paris, France

According to an article [56], View-Points in Europe that correspond to the above characteristics are
1. Beachy Head, East Sussex - England
2. Pedra Serpal In Galicia - Spain
3. The Eiffel Tower In Paris - France
4. View from Mt Snowdon - Wales, England
5. Pico Ruivo, Madeira Islands - Portugal

According to an article [57], Monuments in Europe that match the above characteristics are
1. The Leaning Tower of Pisa - Italy
2. Big Ben - London, England
3. The Schönbrunn Palace - Vienna, Austria
4. Notre-Dame - Paris, France
5. Neuschwanstein Castle – Germany

4.7 Results of the similarity character analysis in the DBSCAN method

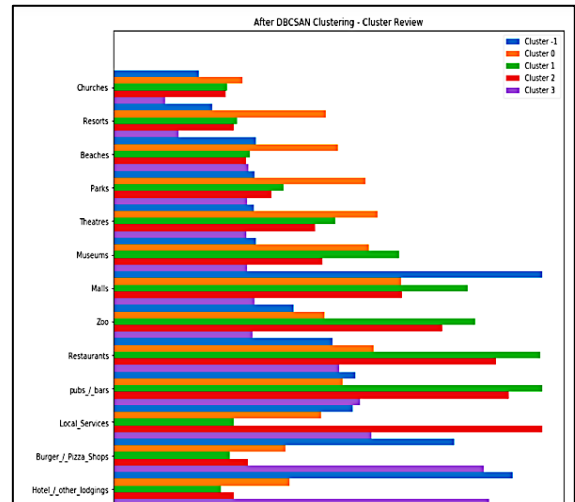Based on Figure 8 and Figure 9, the analysis of character similarities in the DBSCAN method is as follows:

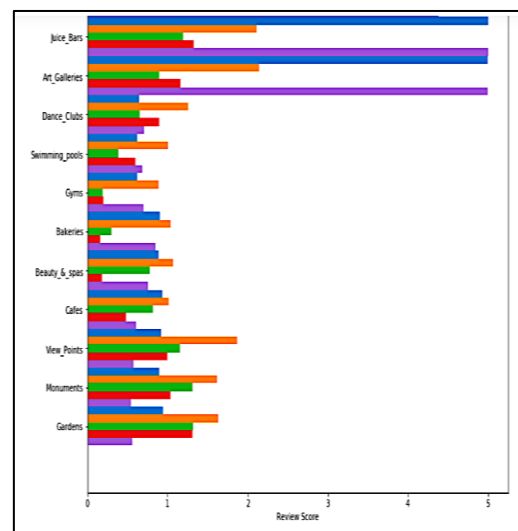
Figure 8. DBSCAN cluster review graph

Figure 9. DBSCAN cluster review graph continued

The way to analyze the clusters that were formed based on these 2 graphs above is to find out which one tourist is giving the best rating in the tourist object category. After that, need to know which one suitable cluster for each tourist object category as in Table 5: in Figure 8 and Figure 9, a category of church tourist objects, has 5 colors, namely blue, orange, green, red, and purple. The orange color is more dominant than the other colors, the orange color shows the color for cluster 0, meaning that tourists in cluster 0 prefer and give the best rating for churches, therefore churches are more suitable to be included in cluster 0. This example also applies to the tourist attraction category.

After it is known that all the categories of tourist objects are included in which cluster, the author analyzed the similarities of the characters from the cluster that was formed. For example, in cluster -1, it is found that tourists in the cluster prefer malls, hotels / other lodgings, juice bars, and art galleries.

According to the author's opinion, tourists in cluster -1 have characters who like to go to European tourist attractions that provide food and drinks that are delicious and healthy for the body, such as juices, smoothies, vegetables. Love to go to European tourist attractions that can spend hours with family and friends such as reunions, birthdays, meetings with family and friends with a relaxed atmosphere equipped with facilities such as wi-fi, air conditioning, attractive and modern views, clean environment and likes hanging out with everyone, especially new people so that you can make many friends.

Then the author checks through the article whether it is true that tourists with characters like the one above enjoy visiting malls, hotels / other lodgings, juice bars, and art galleries. This article is an external data that support the author's opinion regarding the character of tourists in cluster -1. Based on the above analysis, it is found that:

Table 5. The results of the DBSCAN method cluster grouping analysis for 24 tourism object categories

No.	Features	Cluster
1.	Churches	0
2.	Resorts	0
3.	Beaches	0
4.	Parks	0
5.	Theatres	0
6.	Museums	1
7.	Malls	-1
8.	Zoo	1
9.	Restaurants	1
10.	Pubs / Bars	1
11.	Local Services	2
12.	Burger / Pizza shops	3
13.	Hotel / Other Lodgings	-1
14.	Juice Bars	-1 & 3
15.	Art Galleries	-1 & 3
16.	Dance Clubs	0
17.	Swimming Pools	0
18.	Gyms	0
19.	Bakeries	0
20.	Beauty and Spas	0
21.	Cafes	0
22.	View-Points	0
23.	Monuments	0
24.	Gardens	0

1. Cluster -1 --> blue color is called the AJHM Cluster.
   a. Malls - Europe's most preferred tourist attraction

   b. Hotels / Other lodgings
   c. Juice Bars
   d. Art Galleries
The analysis obtained from the AJHM cluster is
   a. The characteristics of tourists who like to go to European tourist objects that provide food and drinks that are delicious and healthy for the body, such as juices, smoothies, vegetables.
   b. Love to go to European tourist attractions that can spend hours with family and friends such as reunions, birthdays, meetings with family and friends with a relaxed atmosphere equipped with facilities such as wi-fi, air conditioning, attractive and modern views, clean environment.
   c. Like hanging out with anyone, especially new people, so that he can make many friends.
2. Cluster 0 --> orange color is called the 2B2G2CSDR-TVBcPaMo Cluster.
   a. Churches
   b. Resorts
   c. Beaches
   d. Parks
   e. Theaters - Europe's most preferred tourist attraction
   f. Dance Clubs
   g. Swimming pools
   h. Gyms
   i. Bakeries
   j. Beauty and Spas
   k. Cafes
   l. View-Points
   m. Monuments
   n. Gardens
The analysis obtained from the 2B2G2CSDR-TVBcPaMo cluster is
   1. Characteristics of tourists who like to go to European tourist attractions that have many useful activities, learn interesting things and can channel their respective hobbies during a trip such as dancing, singing, sports, swimming, culinary bread food, caring for the skin, hang out with nice views and interesting and get to know plants in Europe and how to care for them.
   2. Likes to go to European tourist attractions that have many visitors and have interesting attractions.
   3. Like to go to European attractions with cool and nice views.
   4. Likes to go to tourist objects that can keep memories or beautiful memories with family and loved ones.
   5. Likes to take pictures and prefers to go to European tourist objects that have interesting photo spots.
   6. Like to get a unique experience from the facilities provided by European tourist attractions. For example, in a resort you can experience swimming in a swimming pool that directly leads to the beach or mountain, usually called an infinity pool.
3. Cluster 1 --> green color is called the RPZM Cluster.
   a. Museums
   b. Zoo
   c. Restaurants
   d. Pubs / Bars - Europe's most popular tourist attraction
The analysis obtained from the RPZM cluster is
   a. The characteristics of tourists who like to shop and are willing to spend more money at the end of the month for sightseeing, refreshing, culinary food and

drinks, and seeing the most diverse animal collections in the world.
   b. Like to take pictures and prefer to go to European tourist objects that have interesting photo spots.
   c. Like to get a unique experience from the facilities provided by European tourist attractions. For example, the museums can add insight and information about the history, art, archaeological discoveries, museums building architecture that is truly luxurious and magnificent.

4. Cluster 2 --> red color is called the Ls Cluster.
   Local Services - Europe's most preferred tourist attraction

The analysis obtained from the Ls cluster is
The characteristics of tourists who like to go to European tourist objects at affordable, comfortable, travel to the station are not too far away and have satisfying facilities and services so that they can eliminate fatigue from daily activities.

5. Cluster 3 --> purple color is called the AJBP cluster.
   a. Burger / Pizza Shops
   b. Juice Bars - Europe's most preferred tourist attraction
   c. Art Galleries - Europe's most loved tourist attraction

The analysis obtained from the AJBP cluster is
   a. The characteristics of tourists who like to go to European tourist objects that provide food and drinks that are delicious and healthy for the body, such as juices, smoothies, vegetables.
   b. Love to go to European tourist attractions that can spend hours with family and friends such as reunions, birthdays, meetings with family and friends with a relaxed atmosphere equipped with facilities such as wi-fi, air conditioning, attractive and modern views, clean environment.
   c. Like hanging out with anyone, especially new people, so that he can make many friends.

4.8 Results of K-Means optimization performance with PSO

In the K-Means PSO method, the following performance results are obtained:
   (1) Silhouette Coefficient (0.189).
   (2) DBI Coefficient (1,892).

4.9 Comparative analysis of performance optimization

The comparison seen in terms of performance for the methods, namely K-Means, DBSCAN, and PSO K-Means is as follows:

Table 6. Comparison of optimization performance

Clustering Method	Performance Without Optimization	
	Silhouette Coefficient	DBI Coefficient
DBSCAN	- 0.054	1.408
K-Means	0.178	1.909
Performance After Optimization		
PSO K-Means	0.189	1.892

Table 6 show that the DBI Coefficient value generated by the K-Means method without optimization is not good, so the author optimizing the method using PSO.

Table 6 show that the PSO K-Means method produces better and increased Silhouette Coefficient and DBI Coefficient performance compared to the K-Means method before optimization in the case of tourist interest segmentation, where the Silhouette Coefficient value increases by 0.011 and the DBI value. The coefficient decreased by 0.017.

The author also read articles about the reasons why many tourists make European countries their vacation destinations. This reason will be used as a supporting factor for which clustering method is more suitable, appropriate, and influences the world of tourism in real-time.
According to articles [58-61] the following reasons:
   1. Consists of many countries.
   2. Offers a variety of experiences and adventures for both sea and mountain lovers.
   3. Has nature with beautiful, cool, and friendly climate.
   4. Full of historical buildings with beautiful and nice architecture.
   5. Has a diverse culture and a UNESCO world in a heritage site?
   6. Has a friendly environment for the elderly and disabilities people?
   7. Have a portion of food that fits most people's tongues.
   8. Have a lively nightlife.
   9. Can be guaranteed the safety and cleanliness of the city so that tourists feel comfortable and always want to have a vacation to a European country.
   10. Easy access and transportation for tourists.
   11. Easy visa application.

Table 7. Categories of the most desirable tourist objects from each cluster

K-Means and PSO K-Means		DBSCAN	
Cluster 0	Museums	Cluster -1	Malls (The number of members in this cluster is more than the number of members in other clusters)
Cluster 1	Juice Bars	Cluster 0	Theatres
Cluster 2	Gardens	Cluster 1	Pubs / Bars
Cluster 3	Restaurants	Cluster 2	Local Services
Cluster 4	View-Points	Cluster 3	Juice Bars, Art Galleries

Based on the 11 reasons above and Table 7 show that the grouping of the characteristics of each tourist against the categories of tourist objects in European countries is more suitable, appropriate, and easier to understand or predict in real-time using the K-Means and PSO K-Means method.

5. CONCLUSION AND SUGGESTIONS

To find out and determine the segmentation of tourist interest, two clustering methods are used, namely the K-Means and DBSCAN methods. The conclusion obtained from this research is that the K-Means method is carried out by first determining the number of clusters. In the case of segmentation of tourist interest, the author determines the K

value is 5 and the DBSCAN method is carried out by determining the minimum number of members in 1 cluster and to be able to create 1 cluster it is necessary to have the minimum distance between neighborhoods or what is the minimum threshold value. For example, we make a threshold of 0.5 means that if > 0.5 is not considered as a Neighborhood, whereas if < 0.5 is considered as 1 Neighborhood, then we recalculate the distance from the center point to the next point. Another example, cluster 0 will be formed when the minimum number of members is met, for example, 15 people.

The minimum number of members in DBSCAN is often referred to as Min-Pts (Minimum Points) and the minimum distance between neighborhoods or minimum thresholds is often referred to as Epsilon. All objects that do not fit into any cluster are considered noise. In the case of segmentation of tourist interest, the author determines the epsilon value is 3.0 and the Min-Pts value is 100. The analysis of the character similarities of each cluster that is formed utilizing similar characteristics or interests between 1 tourist and other tourists will be grouped into 1 cluster, where the clusters are formed will be given an appropriate name and describe the characters in 1 cluster, making it easier for the marketing party in a travel company to know and understand the interests of each tourist object category. This can make tourists more interested in using their services because the vacation or tour package promos offered are according to their interests.

In terms of performance, the K-Means method produces better Silhouette Coefficient performance than the DBSCAN method because the higher the silhouette coefficient value, the better the clustering method is used and the DBI Coefficient performance produced by the DBSCAN method is better than the K-Means method because the smallest value of the results Davies-Bouldin index calculations imply values for better clustering algorithms. In the K-Means Clustering method, obtained a Silhouette Coefficient of (0.193) and a DBI Coefficient of (1,740). In the DBSCAN Clustering method, obtained a Silhouette Coefficient of (-0.054) and a DBI Coefficient of (1.408).

The author does not use a collaborative filtering method because this method requires profiles from other tourists, meaning that it recommends an item based on the match between the profiles of one traveler and the profiles of other tourists. For example, most tourists who like product A also like product B, so if other tourists like product A, we can recommend product B. While the content-based filtering method requires a profile of an item and a history of its activity on these items such as a rating or like or dislike. Requiring a profile means requiring features or characteristics of an item such as theme, genre, author, year published, location, or place name. However, in the case of the segmentation of tourist interest in this tourist attraction category, there is no match for the profile of one tourist with another and there are no features or characteristics as above. Therefore, the author using the clustering method to determine the characteristics/interests of each tourist to be able to offer promos or travel package recommendations according to tourist interests.

Characteristics or similar interests between 1 tourist and other tourists will be grouped into 1 cluster. From each of the clusters that are formed, it can make it easier for companies to know what categories of tourist objects each tourist is interested in or likes. From several papers that the author has read, each optimization technique has different uses. In the case of segmentation of tourist interest towards this tourism object category, it is found that the suitable optimization

technique for the K-Means Method is the PSO (Particle Swarm Optimization) technique because what you want to optimize in the K-Means Method is the initial centroid point. Usually, the initial centroid point is determined manually and randomly, with the optimization technique, the author can find out what is the best starting centroid point in this case.

This PSO technique has helped the author to find the best starting centroid point by the system, not manually. The benefits and uses of the PSO method are that the PSO method can overcome the weaknesses of the traditional clustering method, namely the selection of the initial cluster center. The impact of the PSO Method on the K-Means Method is that it can improve the performance of the K-Means method. This is because the DBI Coefficient value of the K-Means method is not good when compared to the DBSCAN method. Based on research results from the PSO K-Means method, it is found that there is a better increase in the value of the Silhouette Coefficient and DBI Coefficient in this method compared to the K-Means method before optimization. The resulting Silhouette Coefficient value is (0.189) and the DBI Coefficient is (1,892). The results of cluster grouping and character similarity analysis in the K-Means method are the same as the results of cluster grouping and character similarity analysis in the PSO K-Means method.

The results of the analysis of cluster characteristics formed from the PSO K-Means Clustering method are more suitable with external data regarding the reasons why many tourists make European countries one of their holiday destinations rather than the DBSCAN method. For the further development of this research, the author suggest that we can further optimize the characteristics of tourists so that they do not only know their rating of the tourist attraction category but also know the genre, age, and hobbies of each tourist. This allows a marketing company to get to know its customers better. Can reach tourist characteristics of tourist attraction categories in other countries, not only in European countries.

REFERENCES

[1] https://www.kompasiana.com/wilmaximus/5ac7f59fab1 2ae469843b7a2/traveling, accessed on 20 March 2020.

[2] https://www.idntimes.com/travel/journal/ricky-ismail/10-manfaat-traveling-yang-harus-kamu-tahu-c1c2/full, accessed on 20 March 2020.

[3] https://pemasaranpariwisata.com/2017/11/13/karakterist ik-pemasaran-destinasi-pariwisata/, accessed on 20 March 2020.

[4] Gao, Y., Yu, W., Chao, P., Zhang, R., Zhou, A., Yang, X. (2015). A restaurant recommendation system by analyzing ratings and aspects in reviews. In International Conference on Database Systems for Advanced Applications, pp. 526-530. https://doi.org/10.1007/978-3-319-18123-3_33

[5] Chu, W.T., Tsai, Y.L. (2017). A hybrid recommendation system considering visual information for predicting favorite restaurants. World Wide Web, 20(6): 1313-1331. https://doi.org/10.1007/s11280-017-0437-1

[6] Aroyo, L., Stash, N., Wang, Y., Gorgels, P., Rutledge, L. (2007). CHIP demonstrator: Semantics-driven recommendations and museum tour generation. In the Semantic Web, pp. 879-886. https://doi.org/10.1007/978-3-540-76298-0_64

[7] Benouaret, I., Lenne, D. (2015). Combining semantic

and collaborative recommendations to generate personalized museum tours. In East European Conference on Advances in Databases and Information Systems, pp. 477-487. https://doi.org/10.1007/978-3-319-23201-0_48

[8] Kovalenko, O., Mrabet, Y., Schouten, K., Sejdovic, S. (2015). Linked data in action: Personalized museum tours on mobile devices. In ESWC Developers Workshop, pp. 14-19.

[9] Tavčar, A., Csaba, A., Butila, E.V. (2016). Recommender system for virtual assistant supported museum tours. Informatica, 40(3): 279-284.

[10] Osche, P.E., Castagnos, S., Napoli, A., Naudet, Y. (2016). Walk the line: Toward an efficient user model for recommendations in museums. In 2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), Thessaloniki, Greece, pp. 83-88. https://doi.org/10.1109/SMAP.2016.7753389

[11] Guo, Y., Denizci Guillet, B., Kucukusta, D., Law, R. (2016). Segmenting spa customers based on rate fences using conjoint and cluster analyses. Asia Pacific Journal of Tourism Research, 21(2): 118-136. https://doi.org/10.1080/10941665.2015.1025085

[12] Dryglas, D., Salamaga, M. (2018). Segmentation by push motives in health tourism destinations: A case study of Polish spa resorts. Journal of Destination Marketing & Management, 9: 234-246. https://doi.org/10.1016/j.jdmm.2018.01.008

[13] Ahani, A., Nilashi, M., Ibrahim, O., Sanzogni, L., Weaven, S. (2019). Market segmentation and travel choice prediction in Spa hotels through TripAdvisor's online reviews. International Journal of Hospitality Management, 80: 52-77. https://doi.org/10.1016/j.ijhm.2019.01.003

[14] Torres-Bejarano, F., González-Márquez, L.C., Díaz-Solano, B., Torregroza-Espinosa, A.C., Cantero-Rodelo, R. (2018). Effects of beach tourists on bathing water and sand quality at Puerto Velero, Colombia. Environment, Development and Sustainability, 20(1): 255-269. https://doi.org/10.1007/s10668-016-9880-x

[15] Saayman, M., Dieske, T. (2015). Segmentation by motivation of tourists to the Kgalagadi Transfrontier Park. South African Journal of Business Management, 46(2): 77-87.

[16] Tkaczynski, A., Rundle-Thiele, S.R., Prebensen, N.K. (2015). Segmenting potential nature-based tourists based on temporal factors: The case of Norway. Journal of Travel Research, 54(2): 251-265. https://doi.org/10.1177%2F0047287513514296

[17] Wang, W., Chen, J.S., Prebensen, N.K. (2018). Market analysis of value-minded tourists: Nature-based tourism in the Arctic. Journal of Destination Marketing & Management, 8: 82-89. https://doi.org/10.1016/j.jdmm.2016.12.004

[18] Phillips, P., Zigan, K., Silva, M. M.S., Schegg, R. (2015). The interactive effects of online reviews on the determinants of Swiss hotel performance: A neural network analysis. Tourism Management, 50: 130-141. https://doi.org/10.1016/j.tourman.2015.01.028

[19] Xu, X., Li, Y. (2016). The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach. International Journal of Hospitality Management, 55: 57-69.

https://doi.org/10.1016/j.ijhm.2016.03.003

[20] Colombo-Mendoza, L.O., Valencia-García, R., Rodríguez-González, A., Alor-Hernández, G., Samper-Zapater, J.J. (2015). RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes. Expert Systems with Applications, 42(3): 1202-1222. https://doi.org/10.1016/j.eswa.2014.09.016

[21] Gomez-Uribe, C.A., Hunt, N. (2015). The Netflix recommender system: Algorithms, business value, and innovation. ACM Transactions on Management Information Systems (TMIS), 6(4): 1-19. https://doi.org/10.1145/2843948

[22] Miah, S.J., Vu, H.Q., Gammack, J., McGrath, M. (2017). A big data analytics method for tourist behaviour analysis. Information & Management, 54(6): 771-785. https://doi.org/10.1016/j.im.2016.11.011

[23] Marques, C., Mohsin, A., Lengler, J. (2018). A multinational comparative study highlighting students' travel motivations and touristic trends. Journal of Destination Marketing & Management, 10: 87-100. https://doi.org/10.1016/j.jdmm.2018.06.002

[24] Li, J., Xu, L., Tang, L., Wang, S., Li, L. (2018). Big data in tourism research: A literature review. Tourism Management, 68: 301-323. https://doi.org/10.1016/j.tourman.2018.03.009

[25] Hernandez, J.M., Kirilenko, A.P., Stepchenkova, S. (2018). Network approach to tourist segmentation via user generated content. Annals of Tourism Research, 73: 35-47. https://doi.org/10.1016/j.annals.2018.09.002

[26] Amaro, S., Duarte, P., Henriques, C. (2016). Travelers' use of social media: A clustering approach. Annals of Tourism Research, 59: 1-15. https://doi.org/10.1016/j.annals.2016.03.007

[27] Dursun, A., Caber, M. (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. Tourism Management Perspectives, 18: 153-160. https://doi.org/10.1016/j.tmp.2016.03.001

[28] Majhi, S.K., Biswal, S. (2018). Optimal cluster analysis using hybrid K-Means and Ant Lion Optimizer. Karbala International Journal of Modern Science, 4(4): 347-360. https://doi.org/10.1016/j.kijoms.2018.09.001

[29] Das, P., Das, D.K., Dey, S. (2018). A modified Bee Colony Optimization (MBCO) and its hybridization with k-means for an application to data clustering. Applied Soft Computing, 70: 590-603. https://doi.org/10.1016/j.asoc.2018.05.045

[30] Anusha, M., Sathiaseelan, J.G.R. (2015). Feature selection using k-means genetic algorithm for multi-objective optimization. Procedia Computer Science, 57: 1074-1080. https://doi.org/10.1016/j.procs.2015.07.387

[31] Patel, G.K., Dabhi, V.K., Prajapati, H.B. (2017). Clustering using a combination of particle swarm optimization and K-means. Journal of Intelligent Systems, 26(3): 457-469. https://doi.org/10.1515/jisys-2015-0099

[32] Pacifico, L.D., Ludermir, T.B. (2019). Hybrid k-means and improved self-adaptive particle swarm optimization for data clustering. In 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, pp. 1-7. https://doi.org/10.1109/IJCNN.2019.8851806

[33] Galvan-Nunez, S., Attoh-Okine, N. (2017). Hybrid particle swarm optimization and k-means analysis for bridge clustering based on national bridge inventory data. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, 3(2): F4016001. https://doi.org/10.1061/AJRUA6.0000864

[34] https://seruni.id/seru-destinasi-wisata-sesuai-kepribadian/, accessed on 20 March 2020.

[35] https://www.popbela.com/career/inspiration/titaflorita/membaca-karakter-dari-hobi/full, accessed on 20 March 2020.

[36] https://www.gotravelly.com/blog/museum-terbaik-di-eropa/, accessed on 20 March 2020.

[37] https://giannovie.wordpress.com/tag/penginapan-di-eropa/, accessed on 20 March 2020.

[38] https://www.infinitypools.holiday/the-best-of-europes-hotel-infinity-pools/, accessed on 20 March 2020.

[39] https://bigseventravel.com/2019/02/best-burgers-europe-2/, accessed on 20 March 2020.

[40] https://blog.misteraladin.com/11-destinasi-di-eropa-buat-liat-hamparan-bunga-tercantik-saat-musim-semi/, accessed on 20 March 2020.

[41] https://www.blog.bidroom.com/5-best-dance-clubs-in-europe/, accessed on 20 March 2020.

[42] https://www.his-travel.co.id/blog/article/detail/kafe-yang-unik-dan-menarik-di-eropa, accessed on 20 March 2020.

[43] https://www.theguardian.com/travel/2018/nov/17/best-spa-breaks-europe-great-value, accessed on 20 March 2020.

[44] https://www.gotravelly.com/blog/10-cake-bakery-terenak-di-eropa/, accessed on 20 March 2020.

[45] https://www.europeish.com/beautiful-churches-europe/, accessed on 20 March 2020.

[46] https://www.redbull.com/gb-en/best-gyms-in-the-world, accessed on 20 March 2020.

[47] https://www.hipwee.com/list/destinasi-favorit-saat-liburan-ternyata-bisa-menggambarkan-kepribadian-lho-mana-yang-paling-kamu-banget/, accessed on 20 March 2020.

[48] https://www.prnewswire.com/news-releases/swedens-frantzen-claims-top-spot-on-the-much-anticipated-annual-event-opinionated-about-dinings-top-100-european-restaurants-list-2019-300853520.html, accessed on 20 March 2020.

[49] https://theeuropeanbarguide.com/top-100-bars-in-europe-2019-part-2/, accessed on 20 March 2020.

[50] https://www.europeanbestdestinations.com/top/best-shopping-centers-in-europe/, accessed on 20 March 2020.

[51] https://getbybus.com/en/blog/best-zoos-in-europe/, accessed on 20 March 2020.

[52] https://www.europeanbestdestinations.com/top/best-beaches-in-europe-2015/, accessed on 20 March 2020.

[53] https://www.iamexpat.nl/lifestyle/lifestyle-news/7-best-theme-parks-europe, accessed on 20 March 2020.

[54] https://www.pandotrip.com/city-parks-in-europe-32028/, accessed on 20 March 2020.

[55] https://expatexplore.com/blog/theatres-across-europe/, accessed on 20 March 2020.

[56] https://whereintheworldisnina.com/europes-most-epic-viewpoints/, accessed on 20 March 2020.

[57] https://www.architectureartdesigns.com/16-european-monuments-you-must-see-at-least-once-in-your-life/, accessed on 20 March 2020.

[58] https://kumparan.com/kumparantravel/lima-alasan-menjadikan-jerman-destinasi-liburan-tahun-ini-1533814983435185576, accessed on 20 March 2020.

[59] https://www.pegipegi.com/travel/ini-alasan-perancis-paling-banyak-dikunjungi-wisatawan-dunia/, accessed on 20 March 2020.

[60] https://travel.kompas.com/read/2014/10/14/135100927/Kenapa.Harus.Mengunjungi.Eropa, accessed on 20 March 2020.

[61] https://market.travelbiz.id/blog/liburan-ke-eropa-8-alasan-balkan-harus-jadi-destinasi-wisata-pilihan, accessed on 20 March 2020.