Implementation of Data Mining Analysis to Determine the Tuna Fishing Zone Using DBSCAN Algorithm

Muhammad Ramadhani and Devi Fitrianah

***Abstract*—The aim of this study is to map the tuna fishing zones based on the daily fish catch data from the Hindian Ocean. With the study, it is expected to deliver a potential tuna fishing zones mapping, where it is based on the number of catch along with its spatial data. The study utilized a data mining approach with DBSCAN algorithm as the method to cluster the data. The study yields information that the Bigeye tuna is dominated the catch in the west monsoon, while Yellowfin tuna dominated the catch in the east monsoon. Based on the trial using the DBSCAN algorithm, we know that the optimal Eps and MinPts value are 1.5 and 5 respectively to generate a convergence cluster.**

***Index Terms*—Data mining, DBSCAN algorithm, spatial analysis, clustering, rapidminer.**

I. Introduction

Capture fisheries as a system that has an important role in the provision of food, employment opportunities, trade and welfare and recreation for some Indonesians need to be managed that are long-term oriented (sustainability management) [1]. The potential of fisheries which reaches 6.4 million tons / year, coastal strength reaching 81,000 km, the potential of islands with 17,500 islands, the potential of coral reef resources reaching 85,000 km2, the potential of cultivation areas 24,528,178 ha in experience has not been able to improve the fishermen's economy [2]. There are still few studies on tuna fishing areas in Indonesian waters, so the authors took the initiative to conduct research on Tuna Fishing Areas. The algorithm that is suitable for conducting this study is the DBSCAN algorithm. DBSCAN groups together points that are close to each other based on a distance measurement (usually Euclidean distance) and a minimum number of points. It also marks as outliers the points that are in low-density regions [3].

The author will see where the areas will become Tuna Fishing Areas regard to time and place, spatial on tuna fish can be grouped based on proximity to other tuna fish points. In the previous study, it was said that the DBSCAN Method can answer the need for information which areas have high levels of density of sufferers of dengue fever indicated by the presence of clusters and which areas have a high density of low-grade fever outbreaks indicated by noise [4].

This indicates that the clustering process using the DBSCAN algorithm can be categorized as good, especially for tuna fishing areas that have several points from the existing coordinates. In this study, the dataset of tuna fish catches will be processed to determine spatial data on tuna fishing areas using DBSCAN algorithm and using rapidminer as a tool for program execution. The contribution of this study will be used for further study in many ways, one is to deliver a potential tuna fishing zones mapping.

Manuscript received March 12, 2019; revised July 6, 2019. This work was supported in part by Universitas Mercu Buana, Indonesia.

The authors are with Computer Science of Universitas Mercu Buana, Jakarta, Indonesia (e-mail: 41515010087@student.mercubuana.ac.id, devi.fitrianah@mercubuana.ac.id).

II. Literature Review

A. Data Mining

Data mining is a branch of science that combines databases, statistics, artificial intelligence and machine learning [5]. Examples of cases in data mining such as, search for names that are most commonly used in the US state or grouping documents from search results with search engines based on the context [6]. The ultimate goal of data mining is to obtain important information from raw data. The first stage of data extraction is data input, then proceed to the second stage, namely preprocessing, which includes the process of feature selection, dimensionality reduction, and normalization. The purpose of preprocessing is to prepare input data before the data mining process. Then in the third stage there is a data mining process in which there are four core, namely predictive modeling, association analysis, cluster analysis, and anomaly detection. At the last stage there is postprocessing which is the result of data mining [7].

B. Clustering

Grouping a number of data or objects into clusters (groups) so that each in the cluster will contain data that is as close as possible and different from objects in other clusters [8]. The grouping of tuna catch data that we can process according to what we want is that it can determine the tuna catchment area itself. The cluster itself is Tuna so that in a broad shell with tuna data, we can determine the catch area with the Clustering method.

C. Spatial Analysis

Spatial analysis is an analysis that is limited by several factors, such as space, communication and transformation, spatial data shows the position, size and possible topological relationships (shape and layout) of objects on earth [9]. Spatial analysis is also finding in relationships and characteristics that may exist implicitly in spatial databases [10].

Because of the large amount of spatial data that can be obtained from satellite imagery, medical devices, video cameras, etc., it is expensive and it is often unrealistic for users to examine spatial data in detail. Spatial Analysis aims to automate the process of discovering that knowledge.

Thus, it plays an important role in:

a) Extract in spatial patterns and features

doi: 10.18178/ijmlc.2019.9.5.862

b) Capture the intrinsic relationship between spatial and non-spatial data
c) Presents data order in a concise manner and at a higher conceptual level
d) Helps to rearrange spatial database to accommodate semantic data, and to achieve better performance [11].

The study uses spatial analysis with captured data using satellite imagery and has so much data that it needs very good data management to produce the latest data or information.

D. Density Based Spatial Clustering of Application with Noise (DBSCAN)

DBSCAN algorithm is one of the algorithms used for classification or grouping of data. Each object from a radius area (cluster) must contain at least a minimum number of data. All objects that are not included in the cluster are considered as noise [12]. Whereas in this study it is very appropriate to use DBSCAN, because this study determines the spatial area of fish catches with a fairly wide range. DBSCAN can help record the spatial area of tuna catches because DBSCAN can determine a spatial data using its algorithm. According to Ester in the journal said that DBSCAN is very efficient for large-scale spatial databases [13].

DBSCAN algorithm has 2 parameters;

a) Eps: minimum distance between two points. This means that if the distance between two points is lower or equal to this value (eps), these points are considered neighbors.
b) MinPoints: minimum number of points to form solid regions. For example, if we set the MinPoints parameter as 5, then we need at least 5 points to form a solid region.
c) Density Reachable: An object p is the density reachable of object q with respect to and MinPts in a set of objects D if there is a chain of objects $p1, p2,\ldots,pn$, where $p1 = q$ and $pn = p$ where pi+1 density reachable directly from pi with respect to and MinPts, for 1 £ i £ n, pi member of D [14].

The parameters above are parameters that must exist in determining the spatial area with DBSCAN. The above parameters that will be cluster forming parameters in the data that you want to process [15].

E. Related Works Regarding To Density Based Spatial Clustering of Application With Noise (DBSCAN) and Capture Fisheries

There are several studies discussing DBSCAN implementation, in the A. R. Ajiboye, A. G. Akintola, and A. O. Ameen's journal [16] discussing Anomaly Detection using DBSCAN implementation in RapidMiner Applications, the journal discusses the whole starting from the dataset, modeling and cluster results. The Eps and MinPts values needed to get the best results in the journal use Eps: 1 and MinPts: 5 values and produce 2 solid clusters and 1 cluster noise.

The next study related to the determination of tuna fishing areas is D. Fitrianah, H. Fahmi, A. N. Hidayanto, and A. M. Arymurthy [17] who also discussed the determination of tuna fishing areas, but in the journal discusses about Potential Fishing Zone's and in this study discusses the Spatial Areas of Tuna Fishing. It almost looks similar, but in this study discusses Spatial, while in the journal, the study discusses Temporal although there are several similar stages but the results of this study are very different of each other.

III. Methodology

In this section, the author will explain some parts in the methodology stage. As in Fig. 1:

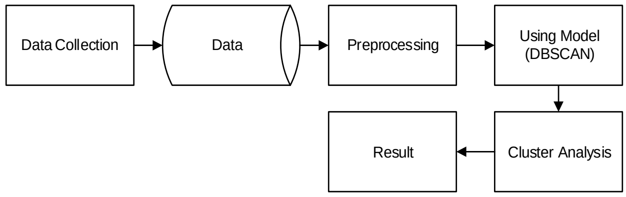

Fig. 1. Block diagram of methodology.

A. Data Collection

The author obtained data on tuna fish catches at sea around 0.05°-21.15°S and 95°-139°E of Bali island with the results of collecting data on tuna catches as much as 35,000 data on tuna catches from 1978-1991. This data is obtained from PT. Perikanan Nasional in Indonesia. The data is also used in study of Tuna Potential Fishing Zones by D. Fitrianah, A. N. Hidayanto, J. L. Gaol, H. Fahmi, and A. M. Arymurthy [18].

TABLE I: Data Collection of Tuna's Fishing

Attribute	*Description*	*Data Type*	*Data*
1	pkID	Int	1
2	Tanggal	Date	11/10/1978
3	Nama Kapal	Varchar	Samudera 12
4	Kode Lokasi	Int	50
5	Kode Area	Int	1120
6	Kode Area 5 deg	Int	25
7	Kode Area 1 deg	Int	1010
8	Latitude	Decimal	-11.43333
9	Longitude	Decimal	117.9166667
10	Basket	Int	283
11	Pancing	Int	6
12	Jumlah Pancing	Int	1698
13	Longline	Int	50940
14	Yellowfin Tuna	Int	1
15	Bigeye Tuna	Int	10
16	Albacore	Int	0
17	S. Bluefin Tuna	Int	0
18	Black Marlin	Int	0
19	Blue Marlin	Int	1
20	Strip Marlin	Int	2
21	Swordfish	Int	1
22	Temperature	Decimal	29.6

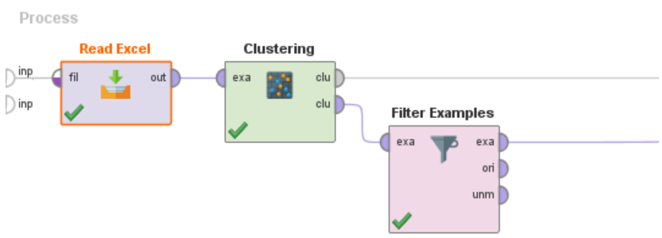

Fig. 2. Models of Rapidminer.

B. Pre-Processing Data

The data that has been obtained has to go through the Pre-processing process first to get the best value when the program execution process takes place. Data cleaning and Reduction will be carried out in this phase to improve some data that is damaged. The data received is not good because there is a lot of data that is empty and attributes that are not needed. Therefore, the results of the Pre-processing of this data will be shown in Table II.

The author has to preprocessing the data manually in the dataset. There are 35,000 data that must be checked one by one. This preprocessing only uses 2 methods:

a) Cleaning: Cleaning data is used to clean up missing values so the data can be processed properly because if the data has a missing value, the data cannot be processed [19]. Missing values can be deleted or replaced with a value of 0. In this tuna catch data, almost 1,000 datas are lost and this takes a long time to complete, considering the preprocessing process is done manually.

b) Reduction: In this step the author chooses, and focuses attention on simplification, abstraction, and transformation of the rough data obtained [20]. This technique is used to simplify some attributes because there are several attributes that are not used in this study. Like the attribute type of fish, because the author only uses 3 types of fish, namely: Yellowfin, Bigeye Tuna, and Albacore, while other attributes must be removed in order to facilitate data to be processed. In addition to the type of fish, the author also requires the attributes of longitude and latitude as variables in determining the tuna catchment area.

TABLE II: DATA AFTER PREPROCESSING

Attribute	Description	Data Type	Data
1	pkID	Int	1
2	Tanggal	Date	11/10/1978
3	Latitude	Decimal	-11.43333
4	Longitude	Decimal	117.9166667
5	Yellowfin Tuna	Int	1
6	Bigeye Tuna	Int	10
7	Albacore	Int	0

Table II. is the result of data that is in accordance with the study needs. The pkID attribute is the ID of the data row. Whereas the attribute date is the determinant of the direction of the west monsoon and east monsoon, while only the last 3 years have been taken for this study.

C. *Using Model (DBSCAN)*

Data that has been checked and preprocessed, will enter the execution stage to find out the results of the DBSCAN algorithm. As shown in Fig. 2. this modeling uses RapidMiner tools to help process the data that has been collected. In this section, the author will provide an explanation of the modeling used in RapidMiner with the DBSCAN Algorithm.

a) Read Excel Operator: This operator is used to import dataset files from Excel to RapidMiner to be processed properly using RapidMiner and in this operator we can also specify which attributes we will use later to execute.

b) Clustering (DBSCAN) Operator: This operator performs clustering with DBSCAN. DBSCAN (for density-based spatial clustering of applications with noise) is a density-based clustering algorithm because it finds a number of clusters starting from the estimated density distribution of corresponding nodes. In this operation, the author will be asked to provide the MinPts value. and Eps value. Then the author will determine the value of Eps: 1.5 and MinPts: 5. From the value of Eps and MinPts this will form a cluster of tuna fish.

c) Fillter Example Operator: This Operator selects which Examples of an ExampleSet are kept and which Examples are removed. In this Fillter Example operator, the author will use it to clean up the noises in the cluster. By removing cluster_0 and cluster_1 so as not to cover other clusters that have been formed.

D. *Cluster Analysis*

This will be the final step before reaching the results, at this stage the author will give the results of clustering of several types of fish (Yellowfin, Albacore, and Bigeye Tuna) which will be associated with the direction of the west monsoon from October - April and east monsoon from April - October 1989-1991.

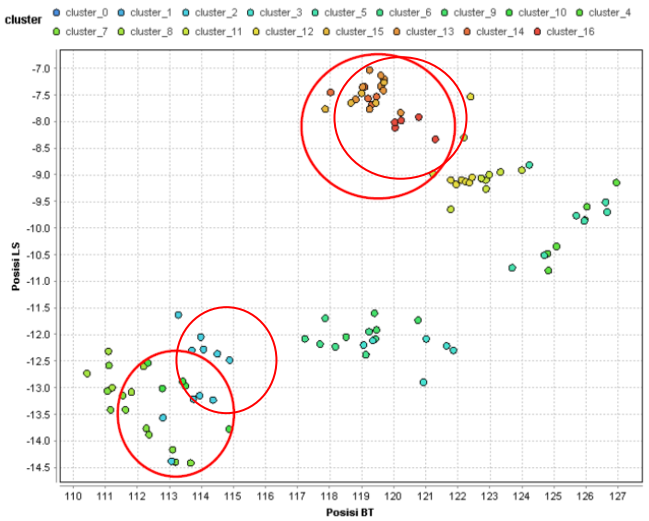

Fig. 3. West Monsoon 1989.

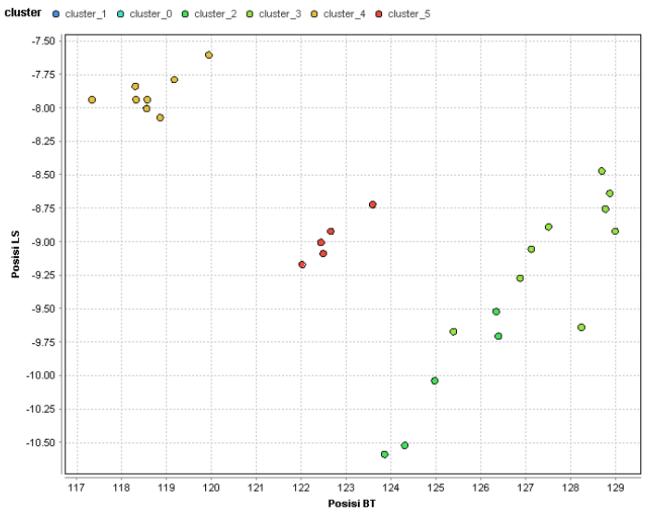

Fig. 4. West Monsoon 1990.

Fig. 3 and Fig. 6 are the result of clusters on the 1989 West Monsoon and East Monsoon from 1989 to 1990. In the Fig. 3, it can be seen that the movement of the fish in formed clusters is very different, but there is 1 cluster point that looks motionless or only moves slightly and it doesn't change too much. In 1989, the West Monsoon cluster formation was dominated by Bigeye tuna. Only 2 clusters are dominated by Yellowfin tuna and 1 cluster dominated by Albacore tuna while the others are Bigeye tuna. In 1989 the East Moonson cluster was also dominated by Bigeye tuna, but the difference was that there were 3 clusters dominated by Yellowfin tuna and 1 cluster by Albacore tuna. This indicates that 2 seasons this year the movement of fish is not so significant and only changed a little.

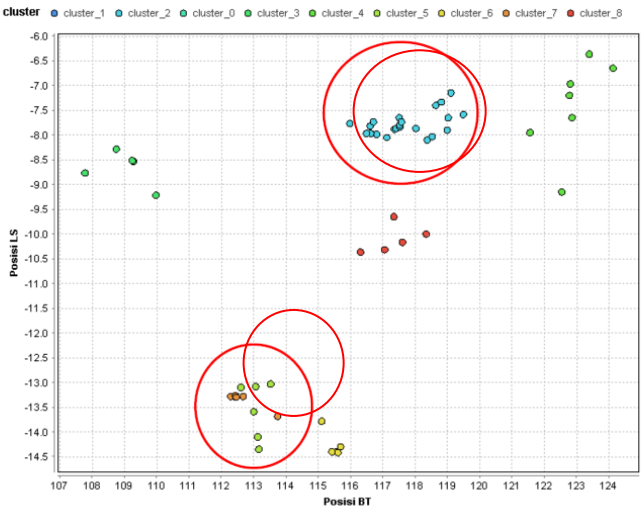

Fig. 5. West Monsoon 1991.

Fig. 6. East Monsoon 1989.

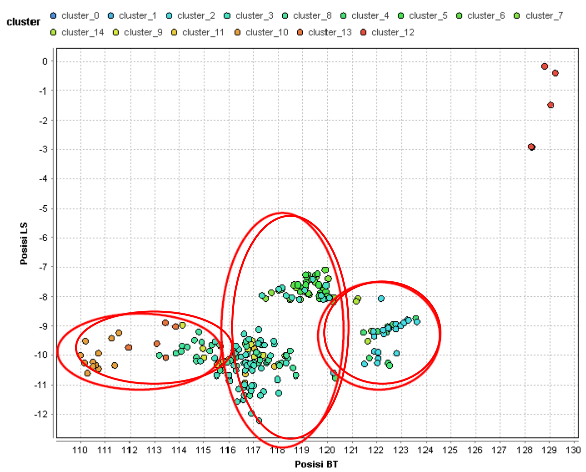

Fig. 7. East Monsoon 1990.

Fig. 4 and Fig. 7 have very large differences, seen in clusters that have been formed. Data on Fig. 4 has very little catch data and has an impact on poor formation of clusters Fig. 4, while in Fig. 7, clusters that are formed are very dense and optimal because the data of fish catches in 1990 on East Moonson have quite a lot of data. Of the 7 clusters formed in 1990 West Monsoon, there were only 2 clusters dominated by Bigeye tuna and 5 clusters dominated by Yellowfin tuna. In 1990 East Moonson only had 2 clusters dominated by Bigeye tuna and 1 cluster dominated by Albacore tuna, while the remaining clusters were dominated by Yellowfin tuna.

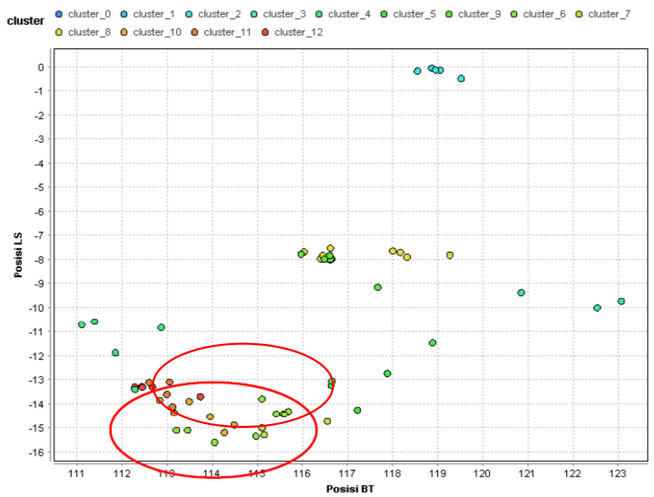

Fig. 8. East Monsoon 1991

This indicates that the movement of the season in that year does not have much impact on the cluster results but the catch data produced is very different considering that the West Monsoon data is not as large as the East Monsoon.

The two seasons in 1991 had a very large difference, seen in Fig. 5 and Fig. 8. The movement of tuna this year is very different each season, it can be seen that the clusters formed from West Monsoon and East Monsoon are very different. In West Monsoon there are about 4 clusters which are dominated by Bigeye tuna and 4 clusters which are also dominated by Yellowfin tuna, this indicates that in West Monsoon in 1991 it had a diverse cluster of tuna fish In East Monsoon there are about 3 clusters which are dominated by Albacore tuna and 2 clusters which are dominated by Bigeye tuna while the others are dominated by Yellowfin tuna.

As we can see in Fig. 3 until Fig. 5 of West Monsoon and Fig. 6 until Fig. 8 of East Monsoon, in the red circle that the author has made, there are some fish points that have not changed even though the year and the season has changed. Almost every chart has a red circle which indicates that the fish point has not changed.

Only in Fig. 4 which do not have a fish point remain due to the number of clusters or the amount of fish catch that is not so much in that season and year. This indicates that there are several habitats of tuna fish that do have a habitat at a fixed longitude and latitude.

IV. Result and Discussion

This is the final stage, where at this stage the author will be talking about the best Eps and MinPts value to get the best cluster and the reason why the author uses the Eps = 1.5 and Minpts = 5 values. The Eps and Minpts values that I have specified are the most optimal values, because those values will form several clusters with several solid points as shown in Fig. 3 until Fig. 8. The author will show some values of Eps and Minpts apart from those specified.

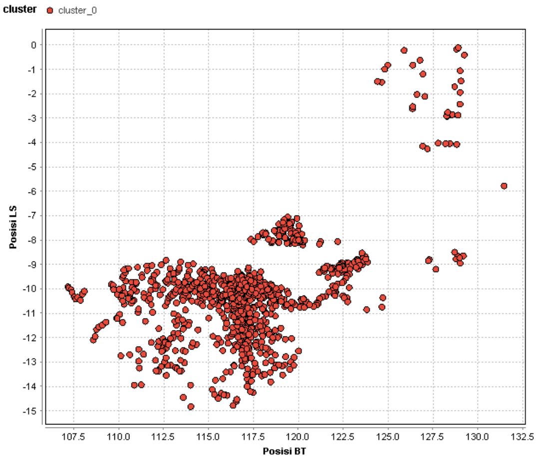

Fig. 9. Eps = 0.5 & MinPts = 5.

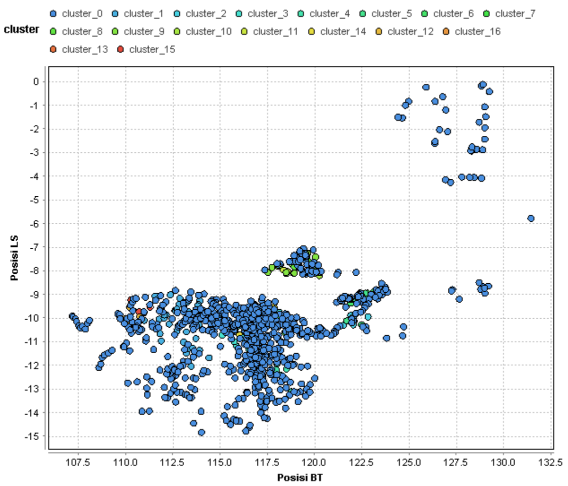

Fig. 10. Eps = 2.0 & MinPts = 5.

Fig. 9 is the result of DBSCAN's spatial cluster analysis using several Eps and MinPts values. The comparison above does only change the Eps value, because the thing that really influences the formation of clusters is the value of the Eps itself. Because the Eps value is used to form a cluster from the specified radius point [21]. In Fig. 10. It is indeed seen forming several clusters, but still too much noise and not so form a dense cluster so that it is not optimal enough if the Eps value used is 2.0. Therefore, the value of Eps = 1.5 is a very optimal value used to form a dense cluster without causing much noise.

TABLE III: ADJUSTMEN OF EPS & MINPTS

Eps	MinPts	Cluster
0.25	5	1
0.25	10	1
0.25	15	1
0.25	20	1
0.25	25	1
0.5	5	1
0.5	10	1
0.5	15	1
0.5	20	1
0.5	25	1
1	5	6
1	10	1
1	15	1
1	20	1
1	25	1
1.25	5	21
1.25	10	4
1.25	15	1
1.25	20	1
1.25	25	1
1.5	5	17
1.5	10	5
1.5	15	3
1.5	20	2
1.5	25	1

Table III describes the experiments of several MinPts and Eps, because of the many experiments only 1 matching value was found to be the basis of the formation of the cluster. There are around 20 trials that are carried out according to Table III. The cluster results are generated from several MinPts and Eps values. Seen some experiments failed because it did not produce a good cluster. But in experiments at the point MinPts: 1.5 and Eps: 5 can produce formed clusters with dense points, while other values are not very able to form clusters with dense points because the values given are not optimal.

TABLE IV: RESULT CLUSTER OF WEST MONSOON

Year	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7	cluster_8	cluster_9
1988-1989	Albacore	Bigeye Tuna	Yellowfin	Bigeye Tuna	0	Yellowfin	Bigeye Tuna	Bigeye Tuna	Bigeye Tuna	Bigeye Tuna
1989-1990	Bigeye Tuna	Bigeye Tuna	Yellowfin	Yellowfin	Yellowfin	Yellowfin	-	-	-	-
1990-1991	Bigeye Tuna	Yellowfin	Yellowfin	Yellowfin	Bigeye Tuna	Albacore	Bigeye Tuna	Bigeye Tuna	Yellowfin	-

TABLE V: RESULT CLUSTER OF EAST MONSOON

Year	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7	cluster_8	cluster_9
1989-1989	Yellowfin	Bigeye Tuna	Yellowfin	0	Yellowfin	Bigeye Tuna	Bigeye Tuna	Albacore	Bigeye Tuna	Bigeye Tuna
1990-1990	Albacore	Bigeye Tuna	Yellowfin	Bigeye Tuna	Yellowfin	Yellowfin	Yellowfin	0	Yellowfin	Yellowfin
1991-1991	Albacore	Yellowfin	0	Yellowfin	Bigeye Tuna	Albacore	Bigeye Tuna	Albacore	Yellowfin	Yellowfin

After several attempts to create a group of spatial areas of tuna fishing by determining several MinPts and Eps values, the optimal value in forming a spatial cluster of tuna fishing catches was obtained. As seen in Table IV and Table V, tuna will continue to move along with the progress of the season and year. We can see that in Table IV. West Monsoon, spatial clusters are almost dominated by Bigeye Tuna, while in Table V. East Monsoon, spatial groups are almost dominated by Yellowfin tuna. Thus in West Monsoon will be dominated by Bigeye tuna in these waters and East Monsoon which dominates this waters is Yellowfin tuna. That is why tuna is one type of fish that moves according to the season and direction of the wind.

V. Conclusion

Using the DBSCAN algorithm, spatial clustering related to fishing grounds can provide excellent results in creating spatial grouping in tuna fishing areas. This method must adjust the Eps and MinPts values to provide optimal results in cluster formation based on the pattern of the fish location itself. The optimal Eps and MinPts is 1.5 and 5 as in Fig. 3 until Fig. 8. The experiment was carried out several times to obtain the optimal Eps and MinPts values, as in Table III.

DBSCAN forms a spatial cluster using optimal Eps and MinPts values. The point of cluster density also affects the results of spatial formation, as in Fig. 9 and Fig. 10 which do not form clusters because the value of Eps and MinPts is not optimal. If the MinPts value given is getting bigger, the cluster formed will be even smaller while if the Eps value gets bigger, the cluster formed will be less.

As we can see in Table IV and Table V, that in the west monsoon season it is dominated by Bigeye tuna, while in the east monsoon season it is dominated by Yellowfin tuna. This data is obtained by analyzing the results of the cluster that has been formed.

References

[1] R. Noviyanti, "Kondisi perikanan tangkap Di wilayah pengelolaan perikanan (WPP) Indonesia," *Univ. Terbuka, Jakarta*, no. Gambar 1, 2011.

[2] Y. Yonvitner, "Produktivitas nelayan, kapal dan alat tangkap di wilayah pengelolaan Perikanan Indonesia," vol. 9, no. 2, pp. 254–266, 2007.

[3] K. Salton, "How DBSCAN works and why should we use it?" *Towards Data Science*, 2017.

[4] A. Yuwono, Y. Oslan, S. Kom, and D. D. Dwijono, "Implementasi metode density based spatial clustering of applications with noise untuk mencari arah penyebaran wabah demam berdarah," *J. EKSIS*, vol. 2, no. 1, pp. 11–21, 2009.

[5] M. Sadikin and F. Alfiandi, "Comparative study of classification method on customer candidate data to predict its potential risk," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 6, pp. 4763–4771, 2018.

[6] M. H. Dunham, *Data Mining: Introductiory and Advanced Topics*, 2002.

[7] D. N. Utama and F. Nhita, "Implementasi density based clustering menggunakan Graphics Processing Unit (GPU)," *e-Proceeding of Engineering*, vol. 2, no. 3, pp. 7905–7912, 2015.

[8] T. Alfina and B. Santosa, "Analisa perbandingan metode hierarchical clustering, K-Means dan gabugan keduanya dalam membentuk cluster data (Studi kasus : Problem kerja praktek jurusan teknik industri ITS)," *Anal. PerbandinganMetode Hierarchical Clust. K-means dan Gabungan Keduanya dalam Clust. Data*, vol. 1, no. 1, pp. 1–5, 2012.

[9] D. Fitrianah, "Spatial and temporal clustering analysis on chlorophyll - a data distribution," *Int. J. Eng. Technol.*, vol. 7, no. 2, pp. 261–265, 2018.

[10] B. Ruswanto, "Analisis spasial sebaran kasus tuberkulosis paru ditinjau dari faktor lingkungan dalam dan luar rumah di kabupaten pekalongan," Thesis, Univerditas Diponegoro,2010.

[11] R. T. Ng and J. Han, "Efficient and effective clustering methods for spatial data mining," in *Proc. 20th Int. Conf. Very Large Data Bases*, 1994, pp. 144–155.

[12] A. S. Devi, I. K. G. D. Putra, and I. M. Sukarsa, "Implementasi metode clustering dbscan pada proses pengambilan keputusan," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 6, no. 3, p. 185, 2015.

[13] M. Ester, H. Kriegel, X. Xu, and D. Miinchen, "A density-based algorithm for discovering clusters in large spatial databases with noise," *KDD-96 Proceedings*, 1996.

[14] H.-P. P. Kriegel and Martin, "Density-based clustering of uncertain data hans-peter," *Metall. Trans. A*, vol. 20, no. 11, pp. 672–677, 2005.

[15] T. Boonchoo, X. Ao, and Q. He. (2018). An efficient density-based clustering algorithm for higher-dimensional data. [Online]. Available: https://www.groundai.com/project/an-efficient-density-based-clustering-algorithm-for-higher-dimensional-data/1

[16] A. R. Ajiboye, A. G. Akintola, and A. O. Ameen, "Anomaly detection in dataset for improved model accuracy using DBSCAN clustering algorithm," *African J. Comput. ICT*, vol. 8, no. 1, pp. 39–46, 2015.

[17] D. Fitrianah, H. Fahmi, A. N. Hidayanto, and A. M. Arymurthy, "A data mining based approach for determining the potential fishing zones," *Int. J. Inf. Educ. Technol.*, vol. 6, no. 3, pp. 187–191, 2016.

[18] D. Fitrianah, A. N. Hidayanto, J. L. Gaol, H. Fahmi, and A. M. Arymurthy, "A spatio-temporal data-mining approach for identification of potential fishing zones based on oceanographic characteristics in the Eastern Indian Ocean," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 9, no. 8, pp. 3720–3728, 2016.

[19] D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. P. Cardoso, "Preprocessing techniques for context recognition from accelerometer data," *Pers. Ubiquitous Comput.*, vol. 14, no. 7, pp. 645–662, 2010.

[20] E. Namey, G. Guest, L. Thairu, and L. Johnson, "Data reduction techniques for large qualitative data sets," *Handb. team-based Qual. Res.*, pp. 137–163, 2007.

[21] D. Birant and A. Kut, "ST-DBSCAN: An algorithm for clustering spatial-temporal data," *Data Knowl. Eng.*, vol. 60, no. 1, pp. 208–221, 2007.

Muhammad Ramadhani was born in Palembang, Indonesia, in 1998. He got the bachelor's degree in informatics engineering from the Universitas Mercu Buana, Jakarta, Indonesia from 2015 to 2019.

He is currently a student in the Faculty of Computer Science. His research interests include data mining, machine learning, and deep learning.

Devi Fitrianah was born in Jakarta, Indonesia, in 1978. She received the bachelor's degree in computer science from Bina Nusantara University, West Jakarta, Indonesia, in 2000, the master's degree in information technology and Ph.D. degree in computer science from the Universitas Indonesia, Depok, Indonesia, in 2008 and 2015, respectively.

She is currently a faculty member with the Department of Computer Science, Universitas Mercu Buana, where she is the head of research center. Her research interests include image procesing, data mining, applied remote sensing, and geographic information system.