ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CLUSTERING TYPES OF CAPTURE FISHERIES PRODUCTS USING THE K-MEANS CLUSTERING ALGORITHM

[1]NURDIN*, [2]BUSTAMI, [3]RINI MEIYANTI, [4]AMALIA FAHADA

[1]Department of Information Technology, Universitas Malikussaleh, Lhokseumawe, Aceh, Indonesia

[2,3,4]Department of Informatics, Universitas Malikussaleh, Lhokseumawe, Aceh, Indonesia.

E-mail: [1]nurdin@unimal.ac.id* (Corresponding Author)

ABSTRACT

North Aceh is one of the districts in Aceh Province which is rich in capture fisheries resources. The North Aceh Regency Maritime and Fisheries Service annually records a large volume of captured fisheries products, reaching tens of thousands of tons with 75 fish divided into 3 types of fish, namely pelagic fish, demersal fish and coral fish spread across 8 sub-districts. is a coastal area. The problem in this research is that there is no system for clustering capture fishery products in North Aceh Regency, so it will be difficult to determine which types of fishery products are classified as low, medium and high catches. This research aims to cluster capture fishery products in North Aceh Regency using the K-Means Clustering algorithm to obtain types of capture fishery products with low catch clusters (C1), medium catch clusters (C2) and high catch clusters (C3). The stages carried out in this research began with the preparation of research instruments and literature review, data collection and analysis and application of the K-Means clustering algorithm. The results obtained from applying the K-Means clustering algorithm are pelagic fish types with a low catch of 86%, a medium catch of 11%, a high catch of 6%. Demersal fish type with low catch 41%, medium catch 53%, high catch 6%. Types of coral fish with low catches of 33%, medium catches of 50% and high catches of 17%. The K-Means clustering algorithm can be used to cluster types of capture fisheries products in North Aceh Regency.

Keywords: *K-Means clustering, Capture fisheries, North Aceh Regency, Clustering algorithm, Data mining*

1. INTRODUCTION

Currently, the development of information technology is growing so rapidly, there are many algorithms, methods and techniques that can be used to solve problems more easily and accurately. One of the uses of information technology is to solve the problem of mapping, clustering fishery products and optimization models for supply chain distribution network planning [1] [2]. North Aceh is one of the districts that has great potential in the marine and fisheries sector, because part of the North Aceh region is a supplier of capture fisheries products [3]. The North Aceh Regency Maritime and Fisheries Service annually records large volumes of captured fisheries products, reaching tens of thousands of tons, which are divided into 3 types of fish, namely pelagic fish, demersal fish and coral fish, spread across 8 sub-districts which are coastal areas in North Aceh Regency. However, until now it is difficult to determine the type of capture fisheries products obtained in any area as a producer of capture fisheries production [4] [5], So need a system for clustering types of captured fishery products using the K-Means clustering algorithm.

Clustering is a data mining method that is unsupervised which is applied without training and does not require targets or outputs. In data mining there are two types of clustering methods used in grouping data, namely hierarchical clustering and non-hierarchical clustering [6]. Clustering is a method for searching and grouping data that has similar characteristics between one data and other data. Clustering is an unsupervised data mining method [7], that is applied without training and does not require targets or outputs. In data mining, there are two types of clustering methods used to group data, namely hierarchical clustering and non-hierarchical clustering [8].

The K-Means method is one of the most widely used clustering methods and has been applied in many fields of science and technology. One of the main problems of the k-means algorithm is that it can produce empty clusters depending on the initial center vector [9]. The K-Means method is a partition algorithm, because K-Means is based on determining the initial number of groups by determining the initial centroid value. The K-Means algorithm uses an iterative process to obtain a cluster database [10]. It takes the desired initial number of clusters as input and produces the final number of clusters as output. The euclidean distance formula is used to find the shortest distance between the center of mass and an object using random data. Data that has the shortest or closest distance to the centroid will form a cluster [11]. The K-Means algorithm is a non-hierarchical data grouping method that attempts to partition existing data into two or more groups, so that data that has the same characteristics is put into the same group and data that has other characteristics is put into another group [12].

The K-Means algorithm has been widely used because of its simple algorithm idea, easy to implement and high efficiency when processing large-scale data [13] [14]. The K-Means clustering algorithm has been widely applied in various case studies, one of which is in this research the application of the K-Means clustering algorithm for clustering types of capture fishery products. The problem in this research is that there is no system for grouping or clustering capture fishery products in North Aceh Regency, so it will be difficult to determine which types of fishery products are classified as low, medium and high catches. This research aims to group or cluster capture fishery products in North Aceh Regency using the K-Means Clustering algorithm to obtain types of capture fishery products with low catch clusters (C1), medium catch clusters (C2) and high catch clusters (C3).

It is important to carry out this research to explore the clustering of types of capture fishery products and which areas with the amount of fishery production include high, medium and low clusters. So that regions with high levels of production can maintain their production levels and for regions with low levels of fisheries production, regional governments can increase productivity by analyzing patterns that can increase the amount of fisheries production.

2. LITERATURE REVIEW

Selvaraj et al, [15] time-series modeling of fish landings in the Colombian Pacific ocean using the ARIMA model. with the results of their research, the ARIMA model is a method that can be used to analyze statistical data. In data-deficient fisheries situations, this method can support the evaluation of fisheries production potential for decision making and management. Data limitations in most recreational fisheries, and the increasing use of catch and release as a fisheries management strategy indicate the need to develop further data integration tools to assess population trends and sustainability of fisheries resources [16]. Management of fishery product resources requires mapping and clustering of fishery products by utilizing the Geographic Information System model [17].

Fitrianah et al, [18] analyzed the determination of potential fishing zones based on mining approach data. The algorithm used in this research is AGRID+. The results of his research explained that the best cluster was formed with daily temporal aggregates, namely the number of cells (m) = 14 with a total of 50 clusters. Utilization of a data mining approach produces 22 groups every day that are identified as potential fishing zones. The use of data mining can provide better input into the data provided, where further research can further explore various useful data mining tools to gain a better understanding of the available data to help improve the welfare of fishermen and aquaculture farmers in Indonesia [19].

Sunarmo et al, [20], implemented machine learning techniques for partition grouping using the K-Means algorithm. This grouping is used to represent the spatial distribution of VMS data in the WPPNRI-711 Fisheries Management Area. Based on the Elbow method, the optimal number of clusters obtained is 7. The results of clustering with the K-Means algorithm show that the distribution of data in each cluster has a value in percentage (90.7%).

Saifullah et al, [21] research aims to detect fish using segmentation, namely segmenting fish images using K-Means clustering. The process carried out is preprocessing to perfect the image. Preprocessing is carried out twice, namely before and after segmentation using K-Means. in preprocessing stage 1 using resize and reshape. The results of preprocessing are segmented using K-Means clustering. The processed object provides a clear picture of the fish object so that K-Means segmentation can help detect fish objects. A new fish image segmentation method combining K-means clustering segmentation algorithm and mathematical morphology has been proposed, which is more accurate and stable than Otsu and other segmentation algorithms [22].

Hablum et al, [23] the K-Means algorithm succeeded in classifying fish catches for the 2015-2017 period using 2 groups, namely group one was categorized as the few catches, and group two was categorized as the most catches. The initial cluster center or centroid is adjusted to the number of variables present. Sugiarto et al, [24] geographic information system for fisheries and livestock areas in Pasuruan Regency. This research shows that the geographic information system of Pasuruan Regency is able to provide information covering fisheries areas, animal husbandry, and the amount of production per year.

Nurdin et al, [25] fisheries yield prediction information system uses a multiple linear regression algorithm. This information system can predict capture fisheries results in Bireun Regency in 2021 of 12,813,870305238 tonnes. The variables used in this research consist of the number of fish caught, the number of motorboats and the number of rainy days. Saifullah et al, [26] aims to detect fish using segmentation, namely segmenting fish images using the K-Means algorithm. Image segmentation is a concept that is often used for object detection.

Anas & Rais [27] mapped risk areas based on the number of natural disasters that have occurred. The clustering method used in this research is the K-Means method. The K-Means method can analyze data well, but is not able to provide detailed information about disaster-prone areas. To overcome this weakness, a Geographic Information System (GIS) was implemented to map the types of disasters.

3. RESEARCH METHODOLOGY

3.1 Description of Problem Formulation

The difficulty of the community and fisheries stakeholders in obtaining information on the types of capture fisheries products and the need for data collection on types of fish with high, medium and low catches at each fishing port in North Aceh Regency. The problem in this study is that there is no clustering system for types of capture fisheries products in North Aceh Regency. Therefore, researchers are interested in creating a model of a clustering system for types of capture fisheries products using the K-Means clustering algorithm. This study is important to determine the clustering of types of capture fisheries products and which areas have high, medium and low fisheries production. So that areas with high production levels can maintain their production levels and for areas with low fisheries production levels, local governments can increase their productivity by analyzing patterns that can increase the amount of fisheries production.

3.2 K-Means Clustering Algorithm

The k-means algorithm is used to determine the number of clusters [28], formed through the use of a specific condition known as criterion, which is involved in the optimal, the splitting method utilizes a condition called as criterion, which is involved in the optimal division of the dataset set by appropriate optimization problems [29]. K-means provides a more comprehensive view of applicant characteristics and needs; using K-means clustering, it is possible to identify the key characteristics of each potential data cluster [30]. Data that has a representative value similarity in one group and data that has a difference in another group so that it allows grouping different data that has a small level of variation. The main principle of this technique is to construct K centroid mass partitions from a set of data, Using the Euclidean Distance formula, calculate the distance between each input data point and each centroid [31] in equation (1).

$$D_{(x,y)} = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} \qquad (1)$$

Where, $D_{(x,y)}$ represents Euclidean Distance, X_1 represents first data training, Y_1 first data testing, X_2 represents second data training, Y_2 represents second data testing, X_n represents n data training, and Y_n represents n data testing.

The process stages in implementing the K-Means Clustering algorithm are as follows [31], determine the value of k as the number of clusters to be formed, initialize k cluster centers in a random way from the dataset, calculating the distance of each input data to each centroid using the Euclidean Distance formula, classify each data based on the closest distance to the centroid, update the centroid value, the new centroid value is obtained from the cluster average, repeat from step 2 to 5, until nothing changes in the members of each cluster

3.3 Data Collection and Variable Type

This research was conducted in North Aceh Regency, Aceh Province, Indonesia, covering the capture fisheries area on the coast of North Aceh. The data used in this research was obtained from the Maritime Affairs and Fisheries Service. The amount of data used is 74 fisheries data which is divided into 3 types of fish, namely: 36 pelagic fish, 32 demersal fish and 6 coral fish. The data variables used in this research are the name of the fish and the number of fish caught in 8 sub-districts in North Aceh Regency, where each sub-district is initialized with the criteria: X1 (Dewantara Subdistrict), X2 (Lapang Subdistrict), X3 (Muara Batu Subdistrict), X4 (Samudra Subdistrict), X5 (Seunuddon Subdistrict), X6 (Syamtalira Bayu Subdistrict), X7 (Tanah Jambo Aye Subdistrict) and X8 (Tanas Pasir Subdistrict).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. RESULT AND DISCUSSION

4.1 Dataset

The dataset used in this study consists of the production results of capture fisheries types in North Aceh Regency in Table 1.

Table 1: Capture Fisheries Production Dataset

No.	Fish Names	X1	X2	X3	X4	X5	X6	X7	X8
1	Ikan Terbang (*Flying Fish*)	0,00	12,73	368,76	26,48	0,00	0,00	0,00	0,00
2	Belanak (*Mullet*)	28,20	20,37	84,62	16,10	62,32	20,48	39,91	13,05
3	Bentong (*Oxeye/Bigeyo Scad*)	5,33	6,08	9,50	8,24	3,04	12,04	2,66	3,17
4	Kerapuh balong (*Honeycom grouper*)	7,79	4,90	15,08	10,05	8,17	2,39	3,35	9,93
5	Bawal putih (*Silver pomfred*)	3,02	5,03	14,57	18,34	11,43	2,01	2,58	5,90
6	Julung-Juling (*Gerfish*)	20,53	16,97	91,10	14,32	79,43	26,35	27,11	22,80
7	Banyar (*Indian Mackerel*)	13,94	23,18	68,92	21,54	35,73	17,01	35,73	15,08
8	Kembung (*Short Body*)	45,45	58,84	102,10	37,08	49,67	36,62	49,67	32,68
9	Layang (*Scad*)	28,46	17,13	39,91	17,61	53,09	11,14	53,09	15,07
10	Tembang (*Frigescole*)	29,40	29,26	154,18	19,64	124,28	60,94	32,56	20,39
11	Siro (*Bali Sarilla*)	12,79	17,23	74,24	20,46	21,28	5,57	9,13	14,32
12	Selar (*Trevalles*)	11,42	5,71	53,58	16,83	19,89	8,61	4,81	12,41
13	Sunglir (*Rainbow*)	23,57	14,32	73,60	22,04	22,42	3,80	11,53	15,33
14	Cakalang (*Skipjack tuna*)	37,03	44,45	406,90	23,52	154,24	21,28	5,95	88,47
15	Tongkol Krai (*Frigate Tuna*)	32,56	27,74	33,05	12,41	27,62	10,84	25,38	33,71
16	Kenyar (*Stripped Bonita*)	8,61	18,49	11,46	11,14	5,95	3,97	11,28	11,68
17	Terubuk (*Hilso Shad*)	39,02	105,28	61,70	15,33	19,64	11,66	10,52	6,46
18	Kapas kapas (*Fels cravelly*)	32,92	15,70	71,63	30,79	13,95	10,68	6,66	22,87
19	Selanget (*Chacusda*)	14,70	8,61	69,05	18,11	19,51	5,33	10,77	12,55
…	………………	……	……	……	……	…..	…..	……	…..
74	Bawal Putih (*Silver pomfred*)	3,02	5,03	14,57	18,34	11,43	2,01	2,58	5,90

4.2 Application of the K-Means Clustering Algorithm

The following are the stages of the clustering process for types of captured fisheries products in North Aceh Regency using the K-Means algorithm:

1. Determine the number of clusters in this clustering of types of capture fishery products, 3 clusters are used, consisting of: Low (C1), medium (C2) and high (C3) catch clusters.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. Choose the initial center of mass randomly researchers chose the following random centroids to be used in the manual clustering calculation process in this study. The following is Table 2. Random Centroid Iteration-1, three types of fisheries were taken as data samples.

Table 2: Initial Centroid Values of Fish Types

Centroid Point	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
C1	23.22	24.63	66.28	19.92	42.82	18.58	21.76	19.23
C2	27.19	61.43	388.99	25.82	113.75	23.69	15.99	50.62
C3	2588.1	12.40	65.45	5.57	20.98	16.47	18.70	13.55

3. After finding the centroid value for the initial calculation that will be used to calculate the distance between the data and the centroid. The equation used in calculating the distance in this study is using the Euclidean Distance formula.

Calculate the distance of the first data with the first centroid

$$D_{C1} = \sqrt{\begin{matrix}(0-23.22)^2 + (12.73-24.63)^2 + (368.76-66.28)^2 \\ +(26.48-19.92)^2 + (0-42.82)^2 + (0-18.58)^2 \\ +(0-21.76)^2 + (0-19.23)^2\end{matrix}}$$

$D_{C1} = 309$

Calculate the distance between the first data and the second centroid

$$D_{C2} = \sqrt{\begin{matrix}(0-27.19)^2 + (12.73-61.43)^2 + (368.76-388.99)^2 \\ +(26.48-25.82)^2 + (0-113.75)^2 + (0-23.69)^2 \\ +(0-15.99)^2 + (0-50.62)^2\end{matrix}}$$

$D_{C2} = 141$

Calculate the distance between the first data and the third centroid

$$D_{C3} = \sqrt{\begin{matrix}(0-2588.18)^2 + (12.73-12.40)^2 + (368.76-65.45)^2 \\ +(26.48-5.57)^2 + (0-20.98)^2 + (0-16.47)^2 \\ +(0-18.70)^2 + (0-13.55)^2\end{matrix}}$$

$D_{C3} = 2606$

4.3 Pelagic Fish Types Clustering Calculation

1. Euclidean Distance Calculation

After finding the centroid for the initial calculation that will be used to calculate the distance between the data and the centroid. The following are the results of the distance using Euclidean Distance. The calculation results using Euclidean Distance were carried out on all sample data so that the final iteration clustering results data were obtained (fourth iteration) in Table 3.

Table 3: The results of the last iteration of pelagic fish Types

No.	C1	C2	C3	Distance	Cluster
1	309	141	2606	141	C2
2	34	315	2561	34	C1
3	79	403	2584	79	C1
4	63	390	2578	63	C1
5	2565	2584	0	0	C3
6	53	305	2576	53	C1
7	88	333	2547	88	C1
8	46	305	2569	46	C1
9	19	334	2574	19	C1
10	68	297	2544	68	C1
11	44	361	2560	44	C1
12	44	354	2565	44	C1
13	32	334	2575	32	C1
14	279	99	2568	99	C2
15	40	356	2577	40	C1
16	30	334	2565	30	C1
17	128	243	2563	128	C1
18	35	341	2574	35	C1
19	88	347	2551	88	C1
20	49	321	2573	49	C1
21	84	408	2587	84	C1
22	69	394	2584	69	C1
23	37	333	2541	37	C1
24	43	363	2557	43	C1
25	430	110	2591	110	C2
26	42	368	2563	42	C1
27	42	369	2556	42	C1
28	71	398	2580	71	C1
29	366	62	2579	62	C2
30	41	344	2569	41	C1
31	23	325	2573	23	C1
32	24	321	2566	24	C1
33	118	232	2567	118	C1
34	36	311	2581	36	C1
35	97	279	2544	97	C1
36	86	345	2541	86	C1

2. Final Iteration Results of Pelagic Fish

After testing each type of pelagic fish data, the calculation stops at the fourth iteration. The following are the results of testing the types of pelagic fish in Table 4.

Table 4: Results of the Last Iteration of Pelagic Fish Types

No.	Fish Names	Iteration Fourth	Description
1	Ikan Teri (*Herklotsichthys dispilonotus*)	**C2**	Medium Catch
2	Belanak (*Mullet*)	**C1**	

			Low Catch
3	Bentong (*Oxeye/Bigeyo Scad*)	**C1**	Low Catch
4	Cendro (*Tylosurus crocodilus*	**C1**	Low Catch
5	Daun Bambu (Queen Fish)	**C3**	High Catch
6	Ikan Terbang (*Flying Fish*)	**C1**	Low Catch
7	Japuh (*Rainbow Srdine*)	**C1**	Low Catch
8	Julung-Julung (*Gerfish*)	**C1**	Low Catch
9	Banyar (*Indian Mackerel*)	**C1**	Low Catch
10	Kembung (*Short Body*)	**C1**	Low Catch
11	Layang (*Scad*)	**C1**	Low Catch
12	Lemuru (*Lemuru Fish*)	**C1**	Low Catch
13	Siro (*Bali Sarilla*)	**C1**	Low Catch
14	Selar (*Trevalles*)	**C2**	Medium Catch
15	Selar Hijau (*GreenTrevalles*)	**C1**	Low Catch
16	Sunglir (*Rainbow*)	**C1**	Low Catch
17	Tembang (*Frigescole*)	**C1**	Low Catch
18	Selanget (*Chacusda*)	**C1**	Low Catch
19	Terubuk (*Hilso Shad*)	**C1**	Low Catch
20	Tetengek *(Terpedo scad)*	**C1**	Low Catch
21	Semar *(Semar Fish)*	**C1**	Low Catch
22	Alba Kora *(Albacora)*	**C1**	Low Catch
23	Tuna Mata Besar (*Thunnus obesus*)	**C1**	Low Catch
24	Tuna Sirip Biru Selatan *(Southern Bluefin Tuna)*	**C1**	Low Catch
25	Tongkol Abu-Abu *(Longtell Tuna)*	**C2**	Medium Catch
26	Tongkol Komo (*Euthynnus affinis*)	**C1**	Low Catch
27	Tongkol Krai *(Frigate Tuna)*	**C1**	Low Catch
28	Kenyar *(Stripped Bonita)*	**C1**	Low Catch
29	Cakalang *(Skipjack tuna)*	**C2**	Medium Catch
30	Lemadang *(Commong Dophin fish)*	**C1**	Low Catch
31	Ikan Layaran *(Istiophorus platypterus)*	**C1**	Low Catch
32	Ikan Pedang *(Swo)*	**C1**	Low Catch
33	Tenggiri (*Scomberomorus commerson*)	**C1**	Low Catch
34	Tenggiri Papan *(Scomberomorus guttatus)*	**C1**	Low Catch
35	Cucut Tikus *(Mousetail Fish)*	**C1**	Low Catch
36	Ikan Pelagis Lainnya *(Other Pelagic Fish)*	**C1**	Low Catch

Based on the results of the clustering calculations of pelagic fish types using the K-Means algorithm, this can be seen in the graph below in Figure 1.

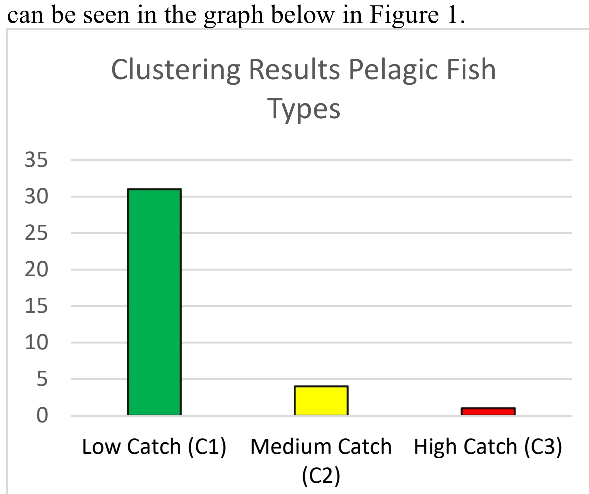

Figure 2: Clustering Results Graph of Pelagic Fish Types

4.4 Demersal Fish Types Clustering Calculation

1. Euclidean Distance Calculation

The following are the distance results using Euclidean Distance, after calculations using Euclidean Distance were carried out on all sample data so that the grouping results data were obtained in the last iteration (fourth iteration) in Table 5.

Table 5: The results of the last iteration of demersal fish Types

No.	C1	C2	C3	Distance	Cluster
1	43	68	572	43	C1
2	20	77	597	20	C1
3	63	57	564	57	C2
4	47	37	552	37	C2
5	87	49	528	49	C2
6	139	90	491	90	C2
7	13	57	584	13	C1
8	48	27	550	27	C2
9	14	62	586	14	C1
10	51	51	559	51	C2
11	109	62	490	62	C2
12	24	54	581	24	C1
13	64	49	548	49	C2
14	129	102	521	102	C2
15	21	79	606	21	C1
16	22	69	589	22	C1
17	41	39	561	39	C2
18	17	53	578	17	C1
19	78	53	543	53	C2
20	46	46	565	46	C2
21	38	45	564	38	C1
22	54	27	552	27	C2
23	55	28	542	28	C2
24	25	55	579	25	C1
25	18	70	596	18	C1
26	44	34	546	34	C2
27	461	403	198	198	C3
28	169	122	438	122	C2
29	18	70	598	18	C1
30	147	120	546	120	C2
31	46	101	618	46	C1
32	746	694	198	198	C3

2. Results of the last iteration of demersal fish types
After testing each type of demersal fish data, the calculation stops at the fourth iteration in Table 6.

Table 6: Results of the Last Iteration of Demersal Fish

No.	Fish Names	Iteration Fourth	Description
1.	Manyung *(Giant Catfish)*	C1	Low Catch
2.	Ikan sebelah	C1	Low Catch
3.	Kuwe *(Jack trevalles)*	C2	Medium Catch
4.	Bawal hitam *(Black pomfred)*	C2	Medium Catch
5.	Bawal putih *(Silver pomfred)*	C2	Medium Catch
6.	Golok-Golok *(Dorab wolf)*	C2	Medium Catch
7.	Beloso *(Greyber lizardfish)*	C1	Low Catch
8.	Gerot-Gerot *(Saddle grune)*	C2	Medium Catch
9.	Ikan Nomei *(bombay duck)*	C1	Low Catch
10.	Kapas-Kapas *(fels cravelly)*	C2	Medium Catch
11.	Peperex *(Leiognathus)*	C2	Medium Catch
12.	Lencan *(Eniperos)*	C1	Low Catch
13.	Kakap Putih *(Barramundi)*	C2	Medium Catch
14.	Kakap Merah *(Red snapper)*	C2	Medium Catch
15.	Jenaha *(Jenaha Fish)*	C1	Low Catch
16.	Kurisi *(Omate hreofin)*	C1	Low Catch
17.	Kuniran *(Sulpur goatfish)*	C2	Medium Catch
18.	Biji Nangka *(yellowstrip goatfish)*	C1	Low Catch
19.	Biji Nangka Karang *(india goatfish)*	C2	Medium Catch
20.	Kurau *(four finger treodfin)*	C2	Medium Catch
21.	Kuro (treodfin)	C1	Low Catch
22.	Swanggi *(purplespot red folgeye)*	C2	Medium Catch
23.	Serin *(red big eye)*	C2	Medium Catch
24.	Gulamah *(croacker)*	C1	Low Catch
25.	Alu-Alu *(creat barracuda)*	C1	Low Catch
26.	Kerong-Kerong *(jerbua terapan)*	C2	Medium Catch
27.	Layur *(hairtail)*	C3	High Catch
28.	Pari Kembang (	C2	Medium Catch
29.	Pari Kalelawar *(dritrays)*	C1	Low Catch

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

30.	Pari burung *(eaglerays)*	C2	Medium Catch
31.	Sembilang	C1	Low Catch
32.	Ikan Damersial Lainnya *(Other Demersal Fish)*	C3	High Catch

Based on the results of the clustering calculations of demersal fish types using the K-Means algorithm, this can be seen in the graph below in Figure 3.

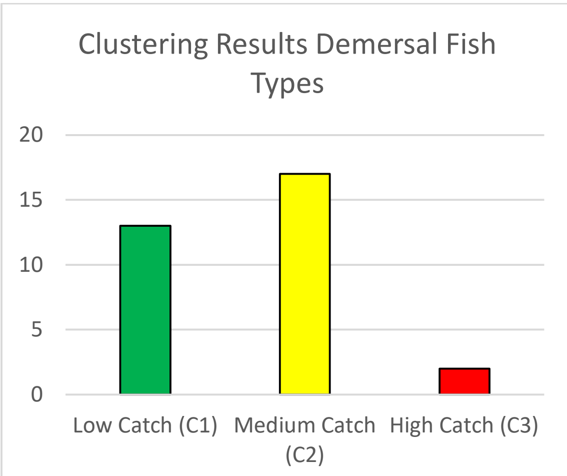

Figure 3: Clustering Results Graph of Demersal Fish Types

4.5 Coral Fish Types Clustering Calculation

1. Euclidean Distance Calculation

The following are the distance results using Euclidean Distance, after calculations using Euclidean Distance on all sample data, the grouping results data were obtained in the last iteration (second iteration) in Table 7.

Table 7: The results of the last iteration of Coral fish Types

No.	C1	C2	C3	Distance	Cluster
1	45	18	42	18	**C2**
2	59	34	0	0	**C3**
3	34	12	33	12	**C2**
4	5	34	60	5	**C1**
5	5	37	59	5	**C1**
6	33	13	34	13	**C2**

2. The results of the last iteration of coral fish types

After testing the data for each type of coral fish, the calculation stops at the second iteration. The following are the clustering results in Table 8.

Table 8: Results of the Last Iteration of Coral Fish

No.	Fish Names	Iteration Second	Description
1.	Ekor Kuning *(yellow tail fusiller)*	**C2**	Medium Catch
2.	Kerapu Karang *(Gluelinedseabass)*	**C3**	High Catch
3.	Kerapu Bebek *(Humipbuckhin)*	**C2**	Medium Catch
4.	Kerapu Balong *(honeycom grouper)*	**C1**	Low Catch
5.	Kerapu Lumpur (*Plectropomus leopardus*)	**C1**	Low Catch
6.	Ikan Karang Lainnya *(Other Coral Fish)*	**C2**	Medium Catch

Based on the results of the calculation of the grouping of coral fish types using the K-Means algorithm, it can be seen in the following graph in Figure 4.

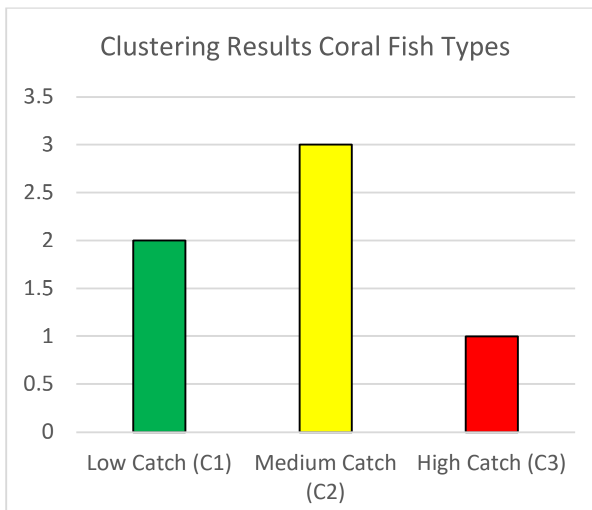

Figure 4: Clustering Results Graph of Coral Fish Types

4.6 Final Clustering Results

The following are the final results of clustering of types of capture fishery products in North Aceh Regency using the K-Means clustering algorithm in graphic form in Figure 5.

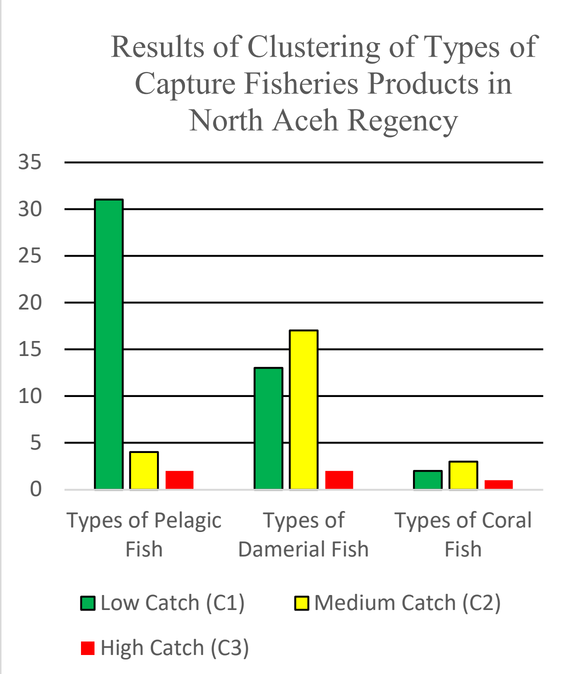

Figure 5: Clustering Results Graph of Fishery Product Types in North Aceh Regency

5. CONCLUSIONS

The K-Means Clustering algorithm calculation process to obtain low, medium and high catch cluster results uses 74 fisheries data which is divided into 3 types of fisheries spread across 8 sub-districts which are coastal areas in North Aceh Regency. The results obtained from applying the K-Means clustering algorithm are pelagic fish types with low catches of 86%, medium catches of 11%, high catches of 6%. Demersal fish type with low catch 41%, medium catch 53%, high catch 6%. A type of coral fish with a low catch of 33%, a medium catch of 50%, and a high catch of 17%. The K-Means clustering algorithm can be used to group types of capture fishery products in North Aceh Regency. The K-Means Clustering algorithm has weaknesses caused by determining the starting point of the centroid. The cluster results formed from the K-Means Clustering algorithm are very dependent on the specified initial cluster starting point value, this makes it very difficult to obtain unique initial centroid results.

Further research is suggested to use other clustering algorithms, such as the K-Medoids algorithm as a comparison to find out which algorithm is better in clustering. The results of this study can be developed using a geographic information system approach.

ACKNOWLEDGMENTS

We acknowledge the support grant and research funding received from The LPPM Universitas Malikussaleh in 2024.

REFERENCES:

[1] N. Nurdin, M. Zarlis, Tulus, and S. Efendi, “Optimization and Computing Model of Fish Resource Supply Chain Distribution Network,” *J. Phys. Conf. Ser.*, vol. 1898, no. 1, 2021, doi: 10.1088/1742-6596/1898/1/012022.

[2] N. Nurdin, Bustami, and Maryana, “Robust optimization approach for agricultural commodity supply chain planningg,” *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 2, pp. 304–315, 2021

[3] F. Fajriana, “Analisis Algoritma K-Medoids pada Sistem Klasterisasi Produksi Perikanan Tangkap Kabupaten Aceh Utara,” *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol.*7, no.*2, pp. 263-269, 2021.

[4] N. Nurdin, M. Zarlis, Tulus and S. Efendi, “Data Driven Optimization Approach to Fish Resources Supply Chain Planning in Aceh Province”, *IOP Conf. Series: Journal of Physics: Conf. Series* 1255 (2019) 012081, 2019.

[5] N. Nurdin, M. Zarlis, Tulus and S. Efendi, “Mixed Integer Linear Programming Model for Integrated Fish Supply Chain Planning” *Journal of Theoretical and Applied Information Technology*, Vol.98, No.12, pp. 2017-2028, 2020.

[6] T. Velmurugan and T. Santhanam, “Computational complexity between K-means and K-medoids clustering algorithms for normal and uniform distributions of data points” *Journal of Computer Science*, vol. 6, no. 3, pp. 363-368, 2010. https://doi.org/10.3844/jcssp.2010.363

[7] J. Zhou, Q. Wang, C.C. Hung and F. Yang, “Credibilistic clustering algorithms via alternating cluster estimation” *J. Intell. Manuf*, vol. 28, pp. 727–738, 2017.

[8] T. Velmurugan and T. Santhanam, “Computational complexity between K-means and K-medoids clustering algorithms for normal and uniform distributions of data points” *Journal of Computer Science*, vol. 6, no. 3, pp. 363-368, 2010. https://doi.org/10.3844/jcssp.2010.363

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[9] A. Malki, M. Rizk, M. El-Shorbagy and A. Mousa, "Hybrid Genetic Algorithm with K-Means for Clustering" Problems" *Open Journal of Optimization*, vol. 5, no. 2, pp. 71-83, 2016. https://doi.org/10.4236/OJOP.2016.52009

[10] P. Arora and S. Varshney, "Analysis of K-means and K-medoids algorithm for big data", *Procedia Computer Science*, 78: 507-512.

[11] Mardalius, "Implementasi Algoritma K-Means Clustering Untuk Menentukan Kelas Kelompok Bimbingan Belajar Tambahan (Studi Kasus : Siswa Sma Negeri 1 Ranah Pesisir)", *Proceding SEMILOKA R. 2017 "Teknologi Mobile,"* vol. 1, p. 105, 2017, doi: 10.31219/osf.io/6mec3

[12] J. Xiao, J. Lu, and X. Li. "Davies Bouldin Index based hierarchical initialization K-means", Intelligent Data Analysis, vol. 21, no. 6, pp. 1327-1338, 2017. https://doi.org/10.3233/IDA-163129

[13] H. Ismkhan, "I-k-means-plus: An iterative clustering algorithm based on an enhanced version of the k-means" *Pattern Recognition*, vol. 79, pp. 402–413, 2018. https://doi.org/10.1016/j.patcog.2018.02.01

[14] K. P. Sinaga, I. Hussain and M. S. Yang, "Entropy K-Means Clustering with Feature Reduction Under Unknown Number of Clusters" *IEEE Access*, vol. 9, pp. 67736–67751, 2021.

[15] J. Selvaraj, V. ARUNACHALAM, K. Coronado-Franco, L. Romero-Orjuela and Y. Ramírez-Yara, "Time-series modeling of fishery landings in the Colombian Pacific Ocean using an ARIMA model" *Regional Studies in Marine Science*, 39, 101477, 2020. https://doi.org/10.1016/J.RSMA.2020.101477

[16] R. Santos, J. Rehage, A. Adams, B. Black, J. Osborne and E. Kroloff, "Quantitative assessment of a data-limited recreational bonefish fishery using a time-series of fishing guides reports" *PLoS ONE*, 12, 2017. https://doi.org/10.1371/journal.pone.0184776

[17] Nurdin, Taufiq, Fajriana, and M. Z. Ulhaq, "Mapping System Model and Clustering of Fishery Products using K- Means Algorithm with Web GIS Approach," International Journal of Intelligent Systems and Applications in Engineering vol. 11, no. 3, pp. 738–749, 2023

[18] 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. https://doi.org/10.7763/ijiet.2016.v6.682

[19] W. Anselmus Teniwut, F. Pentury and Y. Anastasia Ngamel, "Forecasting the welfare of fishermen and aquaculture farmers in Indonesia: Data Mining Approach" *J. Phys. Conf. Ser*, vol. 1175, no. 1, 2019. https://doi.org/10.1088/17426596/1175/1/012066.

[20] Sunarmo, A. Affandi and S. Sumpeno, "Clustering Spatial Temporal Distribution of Fishing Vessel Based lOn VMS Data Using K-Means. 2020 3rd Int. Conf. Inf. Commun. Technol. ICOIACT 2020, pp. 1–6. https://doi.org/10.1109/ICOIACT50329.2020.9331994

[21] S. Saifullah, A. P. Suryotomo and B. Yuwono, "Fish detection using morphological approach based-on k-means segmentation" *Compiler*, vol. 10, no. 1, 2021. https://doi.org/10.28989/compiler.v10i1.946

[22] H. Yao, Q. Duan, D. Li and J. Wang, "An improved KK-means clustering algorithm for fish image segmentation" *Math. Comput. Model,* vol. 58, pp. 790-798, 2012. https://doi.org/10.1016/j.mcm.2012.12.025

[23] R. J. Hablum, A. Khairan and Rosihan. "Clustering Hasil Tangkap Ikan di Pelabuhan Perikanan Nusantara (PPN) Ternate Menggunakan Algoritma K-Means" *JIKO (Jurnal Informatika dan Komputer) Ternate*. vol. 2, no. 1, pp. 26-33, 2019. http://dx.doi.org/10.33387/jiko.v2i1.1053

[24] A. S. Sgiarto, Z.Z. Hani, and A. P. Yosep. "Sistem Informasi Geografis aerah Perikanan Dan Peternakan Kabupaten Pasuruan." *JATI (Jurnal Mahasiswa Teknik Informatika)*. Vol. 8., no. 1, 2020. https://doi.org/10.36040/jati.v4i1.232.

[25] N. Nurdin, F. Fajriana, M. Maryana, and A. Zanati. "Information System for Predicting Fisheries Outcomes Using Regression Algorithm Multiple Linear." *JITE (Journal of Informatics and Telecommunication Engineering*), vol. 5, no. 2, pp. 247-258, 2022.

[26] S. Saifullah, P. S. Andiko, and Y. Bambang. "Fish Detection Using Morphological Aproach Based-On K-Means Segmentation."*Image and Video Processing (Eess.IV)* 1: 1–9. 2018.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[27] S. Annas and Z. Rais. "K-Means and GIS for Mapping Natural Disaster Prone Areas in Indonesia." *Pervasive Health: Pervasive Computing Technologies for Healthcare*, 1: 340–462, 2020.

[28] J. H. Escorcia Guzman, R. A. Zuluaga-Ortiz, D. A. Barrios-Miranda, and E. J. Delahoz-Dominguez, "Information and Communication Technologies (ICT) in the processes of distribution and use of knowledge in Higher Education Institutions (HEIs)," *Procedia Comput. Sci.*, vol. 198, no. 2021, pp. 644–649, 2021, doi: 10.1016/j.procs.2021.12.300.

[29] I. D. Borlea, R. E. Precup, and A. B. Borlea, "Improvement of K-means Cluster Quality by Post Processing Resulted Clusters," *Procedia Comput. Sci.*, vol. 199, pp. 63–70, 2021, doi: 10.1016/j.procs.2022.01.009.

[30] D. Abdullah, S. Susilo, A. S. Ahmar, R. Rusli, and R. Hidayat, "The application of K-means clustering for province clustering in Indonesia of the risk of the COVID-19 pandemic based on COVID-19 data," *Qual. Quant.*, vol. 56, no. 3, pp. 1283–1291, 2022, doi: 10.1007/s11135-021-01176-w.

[31] Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustring dalam Penetuan Siswa Kelas Unggulan," *J. Tekno Kompak*, vol. 15, no. 2, p. 25, 2021, doi: 10.33365/jtk.v15i2.1162