

IMPLEMENTASI ALGORITMA K-MEANS CLUSTERING DALAM PENGELOMPOKAN HASIL TANGKAP IKAN DI KEPULAUAN RIAU

Oleh

Nadia Widari Nasution¹, Nurul Ilmi², Joni Eka Candra³

¹ Sistem Informasi, Institut Teknologi Batam

² Pendidikan Teknik Otomotif, Universitas Negeri Makassar

³ Teknik Komputer, Institut Teknologi Batam

E-mail: ¹widari@iteba.ac.id, ²nurul.ilmi@unm.ac.id, ³joni@iteba.ac.id

Article History:

Received: 25-04-2025

Revised: 09-05-2025

Accepted: 28-05-2025

Keywords:

Hasil Tangkap Ikan, K-Means, Clustering, Data Mining, Rapid Miner

Abstract: Provinsi Kepulauan Riau memiliki potensi yang sangat besar dalam pengembangan budidaya, khususnya budidaya laut. Diketahui jumlah produksi perikanan di Kepulauan Riau mengalami peningkatan setiap tahunnya,. Besarnya potensi ikan yang dimiliki oleh Kepulauan Riau berbanding lurus dengan banyaknya jenis ikan yang ada dan dibudidayakan oleh nelayan ataupun pihak ketiga. Namun nelayan kesulitan menentukan jenis ikan berdasarkan lokasi habitat ikan sehingga terkendala dalam menentukan lokasi penangkapan ikan yang dituju. Selain itu proses penangkapan ikan juga dipengaruhi oleh musim dan waktu penangkapannya. Sehingga nelayan perlu mengetahui pengelompokan jenis ikan berdasarkan titik lokasinya agar dapat menyesuaikan dengan waktu dan juga alat tangkap yang digunakan. Clustering adalah suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (similarity) antara satu data dengan yang lain. Salah satu teknik clustering sederhana adalah algoritma K-Means, menggunakan software RapidMiner. Hasil penelitian ini dapat memberikan pengelompokkan jenis ikan berdasarkan titik lokasi yang outputnya sehingga dapat bermanfaat untuk nelayan di Kepulauan Riau.

PENDAHULUAN

Wilayah lautan merupakan kawasan terluas dan memiliki aneka ragam hayati laut serta kekhasannya yang berbeda dengan wilayah daratan di Indonesia. Dengan luasnya lautan, banyaknya pulau-pulau besar dan kecil menjadikan Indonesia sebagai wilayah berpotensi di sektor perikanan. Bidang perikanan menjadi salah satu bagian yang diharapkan mampu menjadi penopang peningkatan kesejahteraan rakyat Indonesia. Sub sektor perikanan dapat berperan dalam pemulihan dan pertumbuhan perekonomian, karena potensi sumberdaya

ikan yang besar dalam jumlah dan keragamannya (Hablum et al, 2019).

Provinsi Kepulauan Riau memiliki potensi yang sangat besar untuk pengembangan budidaya, khususnya budidaya laut. Hal ini didukung dengan luas laut provinsi Kepulauan Riau yang mencapai 24.121.530,0 hektar atau sekitar 95,79% dari total luas wilayahnya, sementara luas daratan hanya 1.059.511,0 hektar atau 4,21% (BPS Kepulauan Riau 2015). Maka dari itu Kepulauan Riau disebut sebagai kepulauan ikan karena ikan menjadi salah satu sumber makanan pokok yang sering dikonsumsi oleh penduduk setempat.

Jumlah produksi perikanan di Kepulauan Riau mengalami peningkatan setiap tahunnya, jumlah produksi perikanan tangkap (ton) pada tahun 2016 sebanyak 303.411,28 dengan nilai tukar nelayan (NTN) sebesar 109,36%, pada tahun 2017 sebanyak 304.038,71 dengan nilai tukar nelayan (NTN) sebesar 110,7%, pada tahun 2018 sebanyak 304.975,02 dengan nilai tukar nelayan (NTN) sebesar 115,26% , pada tahun 2019 sebanyak 309.287,15 dengan nilai tukar nelayan (NTN) sebesar 117,42% dan pada tahun 2020 sebanyak 319.196,89 dengan nilai tukar nelayan (NTN) sebesar 109.70 (Dinas Kelautan dan Perikanan Provinsi Kepulauan Riau, 2021).

Peningkatan daya perikanan Kepulauan Riau adalah akumulasi dari produksi perikanan budidaya oleh kabupaten dan kota yang ada di Kepulauan Riau. Pada Gambar 1 dijabarkan besarnya potensi perikanan dari berbagai kabupaten dan kota yang diklasifikasikan berdasarkan jenis budidaya yang dilakukan.

Kabupaten/ Kota	Luas Daratan (ha)	Panjang Garis Pantai (km)	Budidaya di Laut (ha)		Budidaya di darat (ha)	
			Pesisir	Laut Lepas	Air Payau	Air Tawar
Batam	211.772	172	10.710	50.422	288	150
Bintan	194.613	435	6.684	37.517	250	389
Karimun	287.320	642	10.211	47.232	190	65
Natuna	205.845	460	6.021	20.393	507	22
Anambas	59.014	132	1.993	19.005	0	50
Lingga	23.920	473	19.054	226.538	828	83
Tanjungpinang	77.027	53	0	0	0	60
Jumlah	1.059.511	2.368	54.672	401.108	2.063	819
Luas Laut	24.121.530			455.780		

Gambar 1. Potensi Perikanan Kabupaten/Kota di Kepulauan Riau

Sumber: Dinas Kelautan dan Perikanan, 2021

Besarnya potensi ikan yang dimiliki oleh Kepulauan Riau berbanding lurus dengan banyaknya jenis ikan yang ada dan dibudidayakan oleh nelayan ataupun pihak ketiga. Namun nelayan kesulitan menentukan jenis ikan berdasarkan lokasi habitat ikan sehingga terkendala dalam menentukan lokasi penangkapan ikan yang dituju. Selain itu proses penangkapan ikan juga dipengaruhi oleh musim dan waktu penangkapannya. Sehingga nelayan perlu mengetahui pengelompokan jenis ikan berdasarkan titik lokasinya agar dapat menyesuaikan dengan waktu dan juga alat tangkap yang digunakan.

Clustering adalah suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan yang lain (Andini et al, 2022). *Clustering* merupakan salah satu metode data mining yang bersifat tanpa arahan (*unsupervised*), maksudnya metode ini diterapkan tanpa adanya latihan (*training*) dan tanpa ada guru (*teacher*) serta tidak memerlukan target output.

Teknik *clustering* banyak digunakan seiring meningkatnya jumlah data yang tersedia. Salah satu teknik *clustering* sederhana adalah K-Means. Algoritma K-Means mengelompokkan data berdasarkan jarak terdekat antara data dengan pusat kluster yang sudah ditentukan (Febriani et al, 2021). Teknik K-Means merupakan salah satu metode data *clustering* non-hirarki yang mengelompokkan data dalam bentuk satu atau lebih *cluster*/kelompok. Data-data yang memiliki karakteristik yang berbeda dikelompokkan dengan cluster yang lain sehingga data yang ada dalam satu kelompok memiliki tingkat variasi yang kecil (Hablum et al, 2019).

Pada penelitian ini akan dilakukan pengelompokkan atau klasterisasi jenis ikan yang ada di Kepulauan Riau agar memudahkan nelayan dalam menentukan lokasi ikan dan mengetahui jenis ikan yang unggul.

LANDASAN TEORI

Data Mining

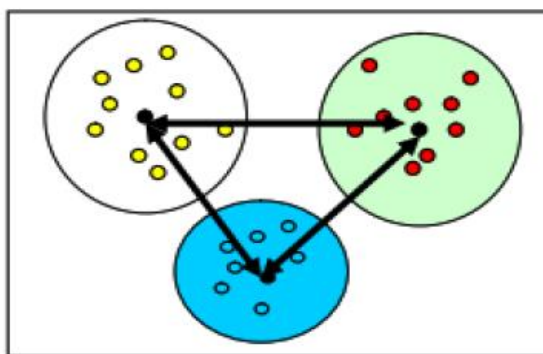
Data Mining adalah proses mengeksplorasi dan menganalisis sejumlah besar data mentah untuk menemukan pola dan pengetahuan tersembunyi yang bermanfaat. Data mining melibatkan integrasi berbagai teknik dari bidang statistik, *machine learning*, database, dan visualisasi untuk melakukan ekstraksi informasi dan pola dari data.

Ada beberapa tahapan penting dalam proses data mining, diantaranya adalah (Manikandan & Selvakumar, 2022):

- a. Pembersihan Data
- b. Integrasi Data
- c. Seleksi Data
- d. Transformasi Data
- e. Pemilihan Teknik Data Mining
- f. Evaluasi Pola dan Pengetahuan
- g. Interpretasi dan Presentasi Pengetahuan

K-Means Clustering

K-Means salah satu metode pengelompokan data non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok (Amalina, 2022). Metode ini mempartisi data untuk meminimalisir varians kedalam kelompok sehingga data berkarakteristik sama di masukkan kedalam kelompok yang lain.

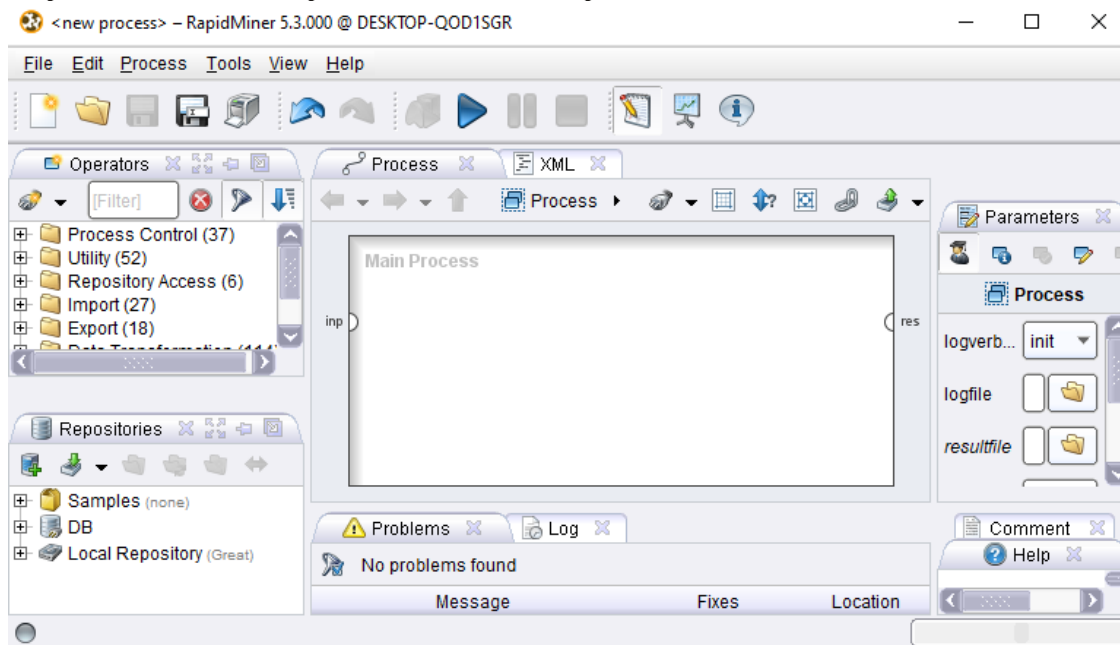


Gambar 2. K-Means Clustering

RapidMiner

RapidMiner adalah platform perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Ini menyediakan beragam alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. RapidMiner dirancang untuk mudah digunakan dan memungkinkan pengguna untuk dengan mudah membangun dan menguji berbagai model, bahkan tanpa pengalaman pemrograman. RapidMiner menawarkan antarmuka drag-and-drop yang memungkinkan pengguna untuk membangun alur kerja untuk memproses dan menganalisis data. Ini mendukung beragam sumber data, termasuk file datar, basis data, dan platform big data seperti Hadoop dan Spark.

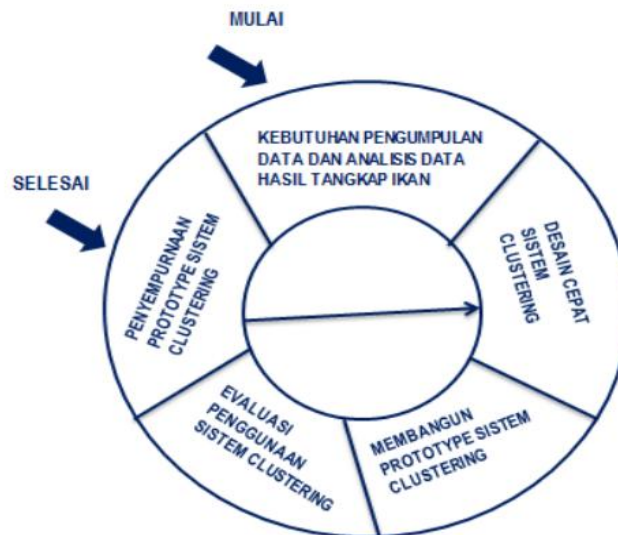
Perangkat lunak ini juga mencakup beragam operator yang sudah dibangun, yang merupakan blok bangunan dari alur kerja, yang mencakup semua tahap proses data mining, seperti pembersihan data, pemilihan fitur, dan pemodelan.



Gambar 3. RapidMiner

METODE PENELITIAN

Dalam penelitian ini, metode pengembangan perangkat lunak yang digunakan yaitu metode *prototype*, sebagai alat bantu dalam memberikan gambaran sistem seperti materi dan menu yang perlu dimasukkan dalam *prototyping* yang akan dikembangkan. Berikut tahapan *prototype* yang dibuat:



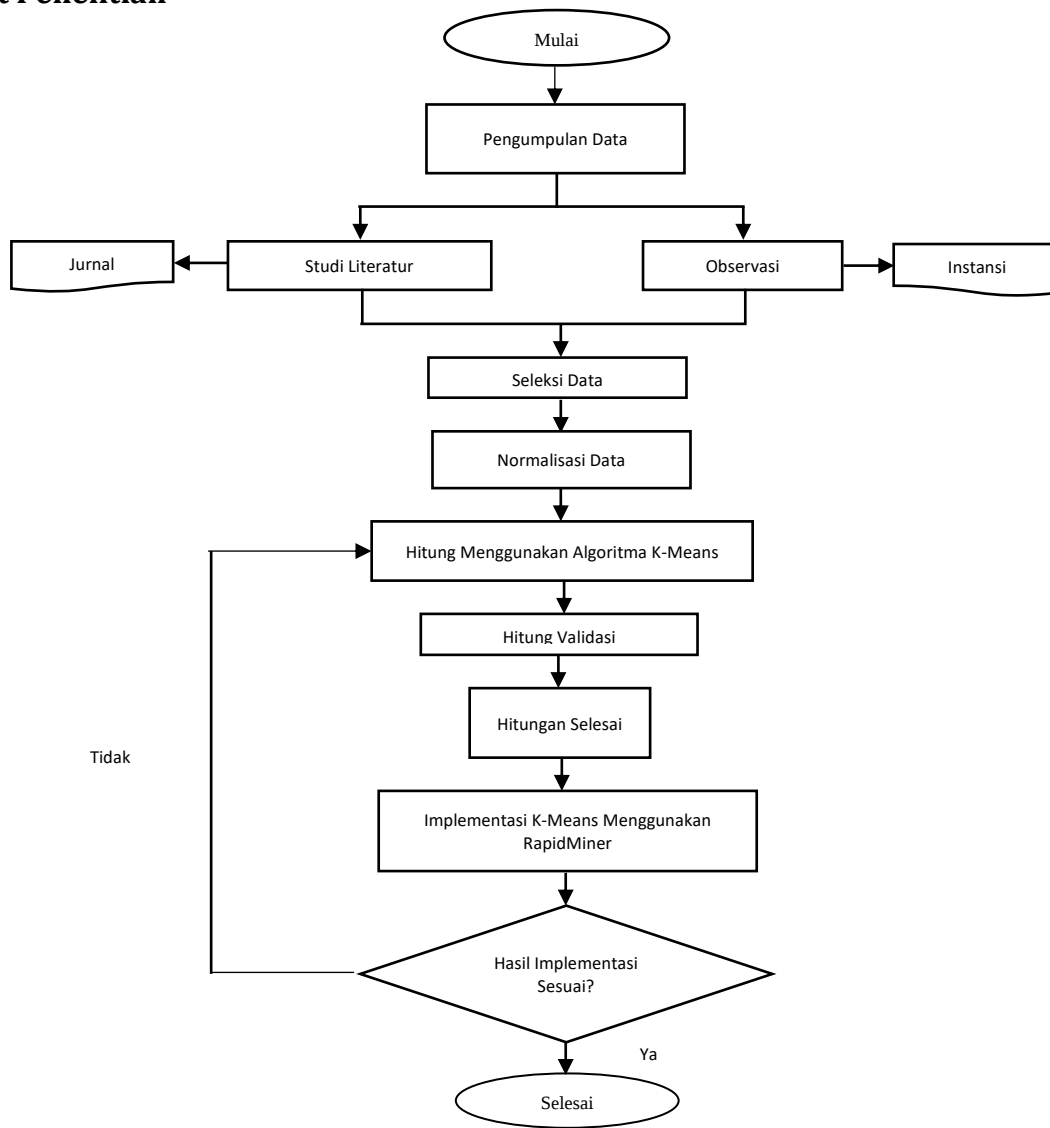
Gambar 5. Tahapan *Prototype Clustering* Hasil Tangkap Ikan

Adapun tahapan *prototype clustering* hasil tangkap ikan sebagai berikut:

- Kebutuhan Pengumpulan dan Analisis Data Hasil Tangkap Ikan**
Mengumpulkan kebutuhan sistem yaitu dengan mengambil data penelitian di Dinas Kelautan dan Perikanan Provinsi Kepulauan Riau kemudian hasilnya di analisa untuk kebutuhan sistem.
- Desain yang Cepat Sistem *Clustering***
Setelan menentukan kebutuhan sistem maka dilakukan desain secara cepat dengan pemodelan DFD dan ERD dari sistem *clustering* tangkap ikan.
- Membangun *Prototype* Sistem *Clustering***
Dari desain yang yang dipergunakan dibuat program yang merepresentasikan sistem untuk *mengclustering* hasil produksi ikan dengan menerapkan algoritma K-Means.
- Evaluasi Pengguna Sistem *Clustering***
Pengguna diminta untuk melakukan uji coba sistem *clustering* agar diketahui kelebihan dan kekurangan, pada tahap ini segala informasi dari pengguna sistem *clustering* dicatat untuk diberikan ke pembuat sistem *clustering*.
- Penyempurnaan *Prototype* Sistem *Clustering***
Jika hasil dari evaluasi pengguna sistem menggunakan *clustering* hasil produksi ikan tidak puas maka dilakukan penyempurnaan sistem sesuai dengan permintaan pengguna, kemudian kembali melakukan desain secara cepat (Langkah 2), dan lanjut ke Langkah berikutnya. Jika masih ditemukan kekurangan pada tahap evaluasi pengguna. Maka dilakukan perbaikan lagi dari Langkah 2. Proses akan berhenti jika Pengguna sudah tidak menemukan kekurangan pada system *clustering* produksi ikan.

Penerapan proses algoritma K-Means menggunakan pemrograman Python yang berguna dalam membantu melakukan kegiatan seperti data mining, klustering, klasifikasi, dan regresi.

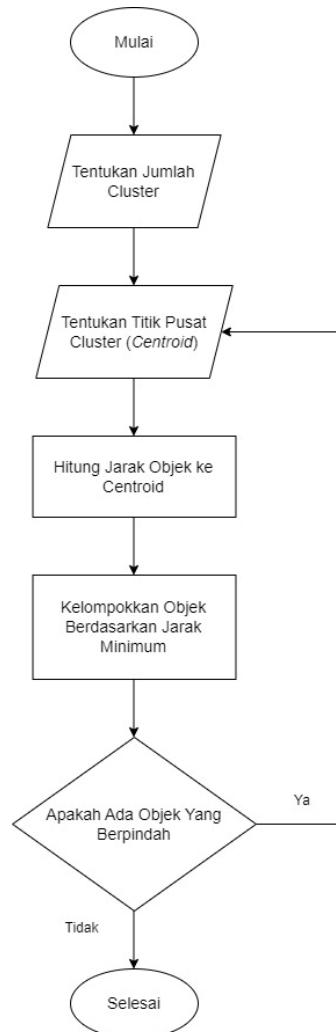
Flowchart Penelitian



Gambar 6. Flowchart Penelitian

HASIL DAN PEMBAHASAN

Flowchart Penelitian



Gambar 7. Flowchart K-Means

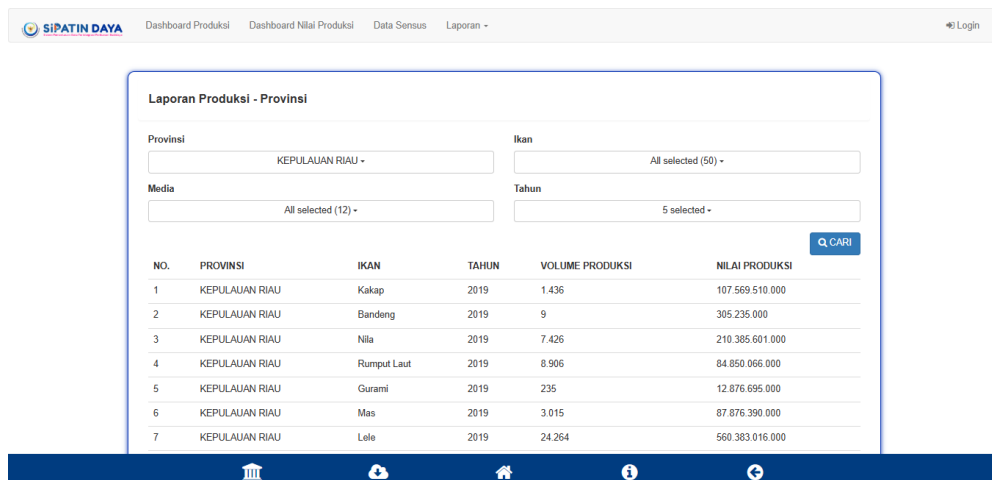
Algoritma K-Means

1. Menentukan jumlah *cluster*, yaitu ada dua *cluster* pada penelitian dengan *cluster* pertama (C_1) dikategorikan sebagai sedikit dan *cluster* kedua (C_2) dikategorikan sebagai banyak.
2. Menentukan pusat cluster atau centroid awal disesuaikan dengan jumlah variabel yang ada. Jumlah variabel yang di tentukan peneliti sebanyak 12, yaitu: Januari, Februari, Maret, April, Mei, Juni, Juli, Agustus, September, Oktober, November, Desember. Maka nilai centroid awal pada masing-masing *cluster* harus 12 data.
3. Menghitung jarak objek ke centroid dengan menggunakan rumus pengukuran jarak.
4. Mengelompokkan objek kedalam *cluster* berdasarkan jarak minimum dalam bentuk tabel, jika objek mengalami perpindahan pada tahap ini maka kembali ke proses ke-2, dan apabila objek tidak mengalami perpindahan maka perhitungan selesai, maka lanjut ke Kesimpulan.

Secara garis besar, jenis sumber daya ikan yang terdapat di perairan laut Kepri adalah: kelompok sumber daya ikan pelagis (tongkol, tenggiri, kembung, layang, teri, dan sebagainya), kelompok sumber daya ikan demersal (kakap merah, kurisi, beloso, bawal, dan sebagainya), kelompok sumber daya ikan karang (kerapu, baronang, napoleon, dan sebagainya), kelompok sumber daya moluska (cumi-cumi, sotong, dan sebagainya), dan kelompok sumber daya krustase (kepiting, rajungan), dan kelompok sumber daya udang.

Data Penelitian

Data produksi ikan di Provinsi Kepulauan Riau untuk tahun 2016 – 2020 dapat dilihat di [SIPATIN DAYA](#) sebagai berikut:



NO.	PROVINSI	IKAN	TAHUN	VOLUME PRODUKSI	NILAI PRODUKSI
1	KEPULAUAN RIAU	Kakap	2019	1.436	107.569.510.000
2	KEPULAUAN RIAU	Bandeng	2019	9	305.235.000
3	KEPULAUAN RIAU	Nila	2019	7.426	210.385.601.000
4	KEPULAUAN RIAU	Rumput Laut	2019	8.906	84.850.066.000
5	KEPULAUAN RIAU	Gurami	2019	235	12.876.695.000
6	KEPULAUAN RIAU	Mas	2019	3.015	87.876.390.000
7	KEPULAUAN RIAU	Lele	2019	24.264	560.383.016.000

Gambar 8. Website SIPATIN DAYA

Data penelitian disajikan dalam bentuk tabel sebagai berikut:

Tabel 1. Data Produksi Ikan Provinsi Kepulauan Riau Tahun 2019

No.	Ikan	Volume Produksi	Nilai Produksi
1	Kakap	1.436	107.569.510.000
2	Bandeng	9	305.235.000
3	Nila	7.426	210.385.601.000
4	Gurami	235	12.876.695.000
5	Mas	3.015	87.876.390.000
6	Lele	24.264	560.383.016.000
7	Patin	6.333	184.141.025.000
8	Kaci-kaci	1.076	77.629.353.000
9	Kerapu	6.238	705.480.546.908
10	Bawal	555	40.068.167.000
11	Gabus	160	9.428.885.000
12	Mujair	304	7.205.501.000
13	Baronang	2	137.400.000

Pengolahan Data

A. Pembersihan Data (*Cleaning Data*)

Proses *cleaning data* merupakan proses pembersihan data yang bertujuan untuk menghilangkan data yang tidak memiliki nilai (null), data yang salah input, data yang tidak relevan, duplikat data dan data yang tidak konsisten karena keberadaannya bisa

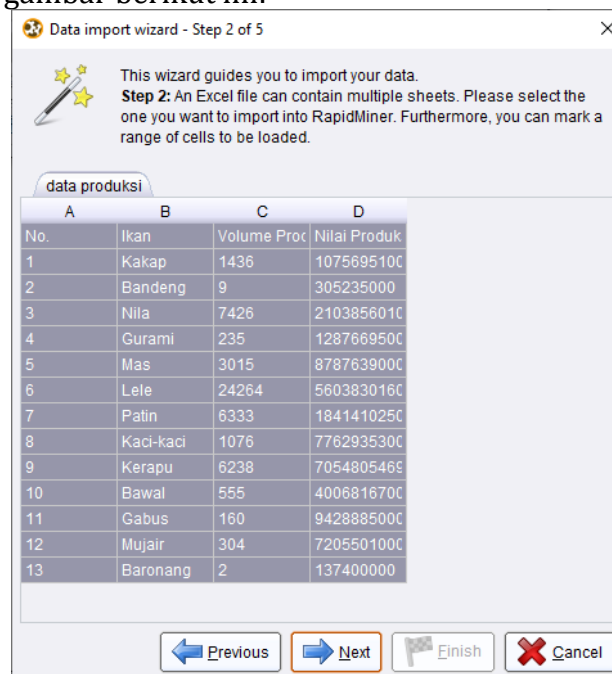
mengurangi mutu atau akurasi dari hasil data mining nantinya. Pembersihan data juga akan mempengaruhi performansi dari system data mining karena data yang akan ditangani akan berkurang jumlah dan kompleksitasnya.

B. Seleksi Data (*Data Selection*)

Untuk penelitian ini data selection dilakukan penghapusan pada duplikasi data, namun pada data penelitian tidak ditemukan data yang duplikat. Sehingga dapat dilakukan tahapan selanjutnya yaitu Transformasi Data.

C. Transformasi Data (*Data Transformation*)

Data transformation digunakan untuk mengubah data dalam bentuk yang sesuai dalam proses data mining. Beberapa 19 teknik untuk *data transformation* adalah *normalization*, pemilihan *attribute*, dan *dicretization*. Adapun *data transformation* pada penelitian ini dapat dilihat pada gambar berikut ini:

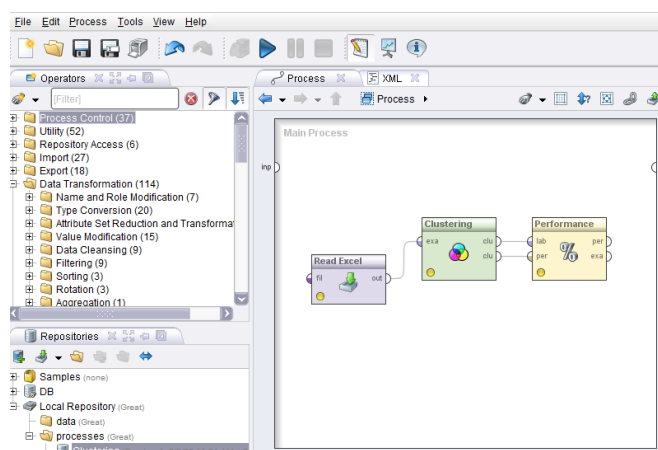


Gambar 7. *Data Transformation*

D. Implementasi

Data mining merupakan proses mencari pola atau informasi menarik dalam data yang terpilih dengan menggunakan metode algoritma K-Means. Berdasarkan gambar model algoritma k-means menjelaskan bahwa operator yang digunakan yaitu retrieve dipergunakan untuk memanggil data set pada penelitian ini. Kemudian clustering k-means untuk memodelkan data set yang telah ada, serta cluster distance performance digunakan untuk menguji hasil pengelompokan, penelitian ini menetapkan 2 kelompok yaitu unggul dan tidak unggul.

Data diproses dengan menggunakan metode k-means dengan parameter $k=2$, measure type=mixed measure. Hal ini dilakukan karena data produksi ikan berupa angka dan menggunakan *Euclidean Distance* untuk standar jarak antar titik klaster.



Gambar 8. Model Algoritma K-Means

Selanjutnya data di proses dengan RapidMiner dan menghasilkan 2 model kluster yang berisi cluster 0 = 6 items, cluster 1 = 7 items. Cluster dimulai dari 0 karena pada bahasa pemrograman angka 0 adalah angka pertama dari urutan penomoran.

Tabel 2. Hasil dari proses RapidMiner

Cluster Model	Jumlah
Cluster 0	6 items
Cluster 1	7 items
Total Cluster	13 items

Pembahasan

Berdasarkan hasil dari nilai performance diatas yang mengkategorikan sebagai Cluster 0 dikategorikan sebagai “Tidak Unggul” sedangkan Cluster 1 sebagai “Unggul”. Dapat dilihat sebagai berikut:

Cluster 0 dengan Kategori Tidak Unggul

Tabel 3. Data Cluster 0 dengan Kategori Tidak Unggul

No.	Atribut	Jenis Ikan	Kategori
1	Cluster 0	Baronang	Tidak Unggul
2	Cluster 0	Bandeng	Tidak Unggul
3	Cluster 0	Gabus	Tidak Unggul
4	Cluster 0	Gurami	Tidak Unggul
5	Cluster 0	Mujair	Tidak Unggul
6	Cluster 0	Bawal	Tidak Unggul

Cluster 1 dengan Kategori Unggul

Tabel 4. Data Cluster 1 dengan Kategori Unggul

No.	Atribut	Jenis Ikan	Kategori
1	Cluster 1	Lele	Unggul
2	Cluster 1	Nila	Unggul
3	Cluster 1	Patin	Unggul
4	Cluster 1	Kerapu	Unggul
5	Cluster 1	Mas	Unggul
6	Cluster 1	Kakap	Unggul
7	Cluster 1	Kaci-kaci	Unggul

Hasil penelitian ini mendapatkan informasi atau pola dari penerapan algoritma k-means dengan data produksi terdapat sebanyak 7 items yang unggul dan terdapat 6 items yang tidak unggul.

KESIMPULAN

Dari penelitian yang sudah dilakukan maka dapat disimpulkan bahwa transaksi yang termasuk cluster 1 adalah kategori unggul dengan jumlah 7 data. Selain itu, pada cluster 0 dengan jumlah 6 data. Dengan dibuatnya penelitian ini maka diharapkan dapat menjadi gambaran untuk Dinas Kelautan dan Perikanan Provinsi Kepulauan Riau dalam menentukan jenis ikan yang paling unggul.

DAFTAR PUSTAKA

- [1] Amalina T, Bima D, Pramana A, Sari BN. Metode K-Means Clustering Dalam Pengelompokan Penjualan Produk Frozen Food. J Ilm Wahana Pendidik [Internet]. 2022; 8(15): 574–83. Available <https://doi.org/10.5281/zenodo.7052276>
- [2] Andini F, Zilfitri D, Filki Y, Ridho M. Algoritma K-Means Clustering dalam Optimalisasi Komposisi Pakan Ternak Ayam Petelur. J Sistim Inf dan Teknol. 2022;5:44–8.
- [3] Badan Pusat Statistik Kepulauan Riau. Potret Potensi Kelautan dan Perikanan Provinsi Kepulauan Riau 2015. 2015. 67 p.
- [4] Dinas Kelautan dan Perikanan Provinsi Kepulauan Riau. Rencana Strategi Dinas Kelautan dan Perikanan Provinsi Kepulauan Riau Tahun 2021-2026. 2021;
- [5] Febriani Y, Sari YP, Octaria D. Metode K-Means Cluster Untuk Mengelompokkan Kota/Kabupaten di Sumatera Selatan Berdasarkan Produksi Ikan Air Tawar. Sainmatika J Ilm Mat dan Ilmu Pengetah Alam. 2021;18(2):175.
- [6] Hablum R, Khairan A, Rosihan R. Clustering Hasil Tangkap Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K Means. JIKO (Jurnal Inform dan Komputer). 2019;2(1):26–33.
- [7] Manikandan L. C., Selvakumar R. K. A Review on Data Mining Concepts and Tools. Int J Sci Res Comput Sci Eng Inf Technol. 2022;3307:600–5.
- [8] Sulistiyawan E, Hapsery A, Arifahanum LJA. Perbandingan Metode Optimasi Untuk Pengelompokan Provinsi Berdasarkan Sektor Perikanan Di Indonesia (Studi Kasus Dinas Kelautan dan Perikanan Indonesia). J Gaussian. 2021;10(1):76–84.
- [9] Tavakolifaradonbe J, Karunarathne S, Vågsæther K. Evaluating the acceptability and accuracy of Phasepy as a Python framework to calculate the interfacial properties and phase equilibrium. Proc 63rd Int Conf Scand Simul Soc SIMS 2022, Trondheim, Norway, Sept 20-21, 2022. 2022;192:198–203.
- [10] Vera JF, Angulo JM. An MDS-based unifying approach to time series K means clustering: application in the dynamic time warping framework. Stoch Environ Res Risk Assess. 2023;37(12):4555–66. 10.
- [11] Widhiyo Sudiyono. The Application of Artificial Intelligence in Djia Stocks 21 To Improve The Investment Profitability Using Python. Int J Econ Bus Account Res. 2022;6(2):1–8. 11.

HALAMAN INI SENGAJA DIKOSONGKAN