

## CLUSTERING HASIL TANGKAP IKAN DI PELABUHAN PERIKANAN NUSANTARA (PPN) TERNATE MENGGUNAKAN ALGORITMA K-MEANS

Rofika Julianti Hablum<sup>1</sup>, Amal Khairan<sup>2</sup>, Rosihan<sup>3</sup>

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Khairun

Jl. Jati Metro, Kota Ternate Selatan

E-mail: rofika.julianti.hablum28@gmail.com<sup>1</sup>, ibntawakkal@gmail.com<sup>2</sup>, rosihan.unkhair@outlook.com<sup>3</sup>

*This research aims to classifying (clustering) the results of catching fish per month for the period to 2015 to 2017 by using the k-means algorithm and determine the most superior fish. Therefore the existence of this research by using the algorithm of K-means clustering can assist to provide the information about superior species of fish or the most abundant species and fish that appear less in the sea waters of Ternate, in order to facilitate the fishermen in preparing for arrest next fish. Data taken by the PPN Ternate is the monthly fishing catch data for the 2015 to 2017 period. Cluster evaluation methods use the DBI method (Davies Bouldin Index) to find out how good or good clusters are used in the research, the authors have calculated several clusters, namely 2 cluster, 3 cluster, 4 cluster, and 5 cluster to determine the comparison, the result is the DBI value in 2 cluster which is closest to value of 0. The DBI value obtained in 2 clusters in 2015 = 0.395, year 2016 = 0.276, and in 2017 = 0.54. The results of the research used 18 fish data to produce 2 clusters, which there were 16 data entered cluster one (C<sub>1</sub>) and 2 data that goes into the cluster two (C<sub>2</sub>).*

**Keywords:** Fish, Cluster, K-means, DBI (Davies Bouldin Index)

**Abstrak** - Penelitian ini bertujuan untuk mengelompokkan (*clustering*) hasil tangkap ikan perbulan periode 2015 sampai 2017 menggunakan algoritma *k-means* dan menentukan ikan yang paling unggul. Maka dari itu dengan adanya penelitian ini dengan menggunakan algoritma *k-means clustering* dapat sedikit membantu dalam memberikan informasi tentang jenis ikan apa yang unggul atau jenis ikan yang paling banyak maupun ikan yang sedikit muncul pada perairan laut Ternate, agar dapat mempermudah para nelayan dalam mempersiapkan penangkapan ikan selanjutnya. Data yang diambil pada PPN Ternate yaitu data hasil tangkap ikan perbulan periode 2015 sampai 2017. Metode evaluasi *cluster* menggunakan metode DBI (*Davies Bouldin Index*) untuk mengetahui seberapa bagus atau baik *cluster* yang dipakai pada penelitian, penulis sudah melakukan perhitungan beberapa *cluster*, yaitu 2 *cluster*, 3 *cluster*, 4 *cluster*, dan 5 *cluster* untuk menentukan perbandingan, hasilnya adalah nilai DBI pada 2 *cluster* yang paling mendekati nilai 0. Nilai DBI yang didapat pada 2 *cluster* tahun 2015 = 0,395, tahun 2016 = 0,276, dan tahun 2017 = 0,54. Hasil penelitian menggunakan 18 data ikan dengan menghasilkan 2 *cluster*, dimana ada 16 data yang masuk pada *cluster* satu (C<sub>1</sub>) dan 2 data yang masuk pada *cluster* dua (C<sub>2</sub>).

**Kata Kunci:** Ikan, Cluster, K-means, DBI (*Davies Bouldin Index*).

## I. PENDAHULUAN

### 1.1. Latar Belakang

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. Perikanan merupakan salah satu bidang 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.

Ternate adalah sebuah kota yang berada di bawah kaki gunung api Gamalama (1.025 m di atas permukaan laut) yang terletak di pulau Ternate. Secara administratif, Ternate berada di provinsi Maluku Utara, Indonesia. Kota Ternate merupakan kota kepulauan yang memiliki luas daratan sebesar 162,03 km<sup>2</sup>, sementara luas lautannya 5.547,55 km<sup>2</sup> [1]. Potensi dari Kota Ternate di sektor perikanan sangatlah besar dilihat berdasarkan pada data statistik Pelabuhan Perikanan Nusantara (PPN) Ternate.

PPN (Pelabuhan Perikanan Nusantara) Ternate adalah salah satu pelabuhan perikanan yang lokasinya cukup strategis berada di dalam kota tepatnya berada di jalan Pasar Inpres, Kelurahan Bastiong Talangame Kecamatan Ternate Selatan Kota Ternate Provinsi Maluku Utara. Akses menuju sarana penunjang seperti bandara dan pelabuhan umum dapat dijangkau dengan mudah. Secara geografis PPN Ternate terletak pada posisi 00 46' 0,36" LU dan 127 22' 41,10" BT (Statistik Perikanan Tangkap PPN Ternate, 2015).

Pelabuhan Perikanan Nusantara Ternate dibangun pada tahun 1978 dengan berbagai sarana dasar, sarana fungsional, dan sarana penunjang. PPN Ternate pada awal berdirinya adalah Pelabuhan Perikanan pantai. Pada pertengahan tahun 2001 dinaikkan statusnya menjadi Pelabuhan Perikanan Nusantara. Hal tersebut didasari pada perkembangan operasional PPN Ternate yang terus mengalami peningkatan dari tahun ke tahun [2]. Volume tangkap ikan dari tahun ke tahun pada PPN Ternate mengalami kenaikan dengan menggunakan beberapa alat tangkap seperti *purse Seine*, *Hand Line*, *Pole and Line*, dan *giop*. Adapun jenis ikan hasil tangkap PPN Ternate adalah ikan cakalang, tongkol, ekor kuning, kembung, madidihang, selar, cumi, tembang, julung-julung, sarden, sunglir, kakap, kerapu, lencam, lolosi, kwee, beronang, cendro, biji nangka, dan swanggi.

Berdasarkan uraian diatas, permasalahan yang di temukan yaitu dengan banyak ragamnya jenis ikan hasil tangkap pada PPN Ternate belum diketahui jenis ikan apa yang unggul disetiap tahunnya, disamping itu juga belum terkelompokkan jenis ikan apa yang paling banyak dan sedikit hasil tangkapnya. Maka dari itu dilakukan penelitian untuk *Clustering Hasil Tangkap*

*Ikan Di Pelabuhan Perikanan Nusantara (Ppn) Ternate Menggunakan Algoritma K-Means*", selain itu dilakukan validasi dengan *davies bouldin index* sebagaimana pernah dilakukan oleh [9] pada pemetaan potensi udang.

*Clustering* adalah suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan yang lain. *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.

*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 [3].

### 1.2. Tujuan Penelitian

1. Mengelompokkan (*clustering*) hasil tangkap jenis ikan perbulan periode 2015-2017 menggunakan algoritma *K-Means* dan menentukan ikan yang paling unggul.
2. Membangun aplikasi menggunakan pemrograman PHP untuk implementasi algoritma *K-Means*.

### 1.3. Manfaat Penelitian

1. Sebagai informasi untuk pengelolaan jenis ikan pada perikanan tangkap di PPN Ternate.
2. Sebagai bahan acuan dan petunjuk untuk merencanakan operasi penangkapan ikan bagi sektor perikanan khususnya PPN Ternate.
3. Bahan masukan bagi perguruan tinggi yang berada dikota ternate maupun diluar daerah baik negeri maupun swasta yang memiliki program studi Perikanan mengenai *clustering* hasil tangkap ikan menggunakan algoritma *K-Means*.

## II. TINJAUAN PUSTAKA

### 2.1. Data Mining

*Data mining* adalah suatu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menentukan keteraturan, pola atau hubungan dalam set data berukuran besar. Salah satu tugas utama dalam *data mining* adalah pengelompokkan *clustering* dimana data yang dikelompokkan belum mempunyai contoh kelompok .

*Data mining* adalah suatu proses pencarian korelasi, pola dan tren baru yang berguna dalam media penyimpanan data berukuran besar menggunakan

teknologi pengenalan pola seperti teknik-teknik statistik dan matematis [4].

## 2.2. Normalisasi

Normalisasi adalah proses penskalaan nilai atribut dari data sehingga bisa jatuh pada range tertentu. Pada proses normalisasi dilakukan dengan metode *Min-Max*, yaitu metode normalisasi dengan melakukan transformasi linier terhadap data asli, berikut persamaan pada metode *Min-Max*:

$$\text{nilai normalisasi} = \frac{x_{awal} - x_{min}}{x_{max} - x_{min}} \dots\dots\dots 1$$

Keterangan:

Xawal = nilai awal

Xmax = nilai maksimal atau nilai terbesar

Xmin = nilai minimal atau nilai terkecil [5].

## 2.3. Clustering

Pada dasarnya *clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. *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*. Dalam data mining ada dua jenis metode *clustering* yang digunakan dalam pengelompokkan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering*.

*Hierarchical clustering* adalah suatu metode pengelompokkan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan pada objek lain yang memiliki kedekatan kedua. Demikian seterusnya kemudian *cluster* akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*.

Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical clustering* justru dimulai dengan menentukan terlebih dahulu jumlah *cluster* yang diinginkan (dua *cluster*, tiga *cluster*, atau lain sebagainya). Setelah jumlah *cluster* diketahui, baru proses *cluster* dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *k-means clustering* [3].

## 2.4. Algoritma K-means

*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 [3].

Algoritma *k-means* juga termasuk metode *Partitioning Method*, yaitu membangun berbagai partisi dan kemudian mengevaluasi partisi tersebut

dengan beberapa kriteria [4].

Berikut rumus pengukuran jarak dalam algoritma *k-means*:

$$d_{(x,y)} = \sqrt{(x_i - y_i)^2 + (x_i - y_i)^2} \dots\dots\dots 2$$

Keterangan:

d = titik dokumen

x = data *record*

y = data *centroid*

Jarak yang terpendek antara *centroid* dengan dokumen menentukan posisi *cluster* suatu dokumen. Misalnya dokumen A mempunyai jarak yang paling pendek ke *centroid* 1 dibanding ke yang lain, maka dokumen A masuk ke group 1. Hitung kembali posisi ke *centroid* baru untuk tiap-tiap *centroid* ( $C_{i-j}$ ) dengan mengambil rata-rata dokumen yang masuk pada *cluster* awal ( $G_{i-j}$ ). Iterasi dilakukan terus hingga posisi group tidak berubah (Dwi, 2012). Berikut rumus dari penentuan *centroid*:

$$C(i) = \frac{x_1 + x_2 + x_3 + \dots + x_n}{\sum x} \dots\dots\dots 3$$

Keterangan:

$x_1$  = nilai data *record* ke-1

$x_2$  = nilai data *record* ke-2

$\sum x$  = jumlah data *record*

Proses algoritma *K-means* sebagai berikut:

- a. Pilih secara acak objek sebanyak k, objek-objek tersebut akan dipresentasikan sebagai *mean* pada *cluster*.
- b. Untuk setiap objek dimasukan kedalam *cluster* yang tingkat kemiripan objek terhadap *cluster* tersebut tinggi. Tingkat kemiripan ditentukan dengan jarak objek terhadap *mean* atau *centroid cluster* tersebut.
- c. Hitung nilai *centroid* yang baru pada masing-masing *cluster*.
- d. Proses tersebut diulang hingga anggota pada kumpulan *cluster* tersebut tidak berubah.

## 2.5. Davies Bouldin Index

*Davies Bouldin Index (DBI)* adalah cara validasi *cluster* yang dibuat oleh D.L. Davies. DBI merupakan salah satu metode evaluasi *cluster* pada suatu metode pengelompokan yang didasarkan pada nilai kohesi dan separasi. Dalam suatu pengelompokan, kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap *centroid* dari *cluster* yang diikuti. Sedangkan separasi didasarkan pada jarak antar *centroid* dari *cluster*nya.

*Sum of square within (SSW) cluster* merupakan persamaan yang digunakan untuk mengetahui matrik kohesi dalam sebuah *cluster* ke-i yang dirumuskan sebagai berikut:

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_i) \dots\dots\dots 4$$

Keterangan:

$m_i$  = jumlah data dalam *cluster* ke-i

$c_i$  = *centroid cluster* ke-i

$d()$  = jarak setiap data ke *centroid* yang dihitung menggunakan jarak *euclidean*

*Sum of square between cluster* (SSB) merupakan persamaan yang digunakan untuk mengetahui separasi antar *cluster*

yang dihitung menggunakan persamaan:

$$SSB_{i,j} = d(c_i, c_j) \dots\dots\dots 5$$

Setelah nilai kohesi dan separasi diperoleh, kemudian dilakukan pengukuran rasio ( $R_{ij}$ ) untuk mengetahui nilai perbandingan antara *cluster* ke-I dan *cluster* ke-j. *Cluster* yang baik adalah *cluster* yang memiliki nilai kohesi sekecil mungkin dan separasi yang sebesar mungkin. Nilai rasio dihitung menggunakan persamaan berikut:

$$R_{ij} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \dots\dots\dots 6$$

Nilai rasio yang diperoleh tersebut digunakan untuk mencari nilai *davies bouldin index* (DBI) dari persamaan berikut:

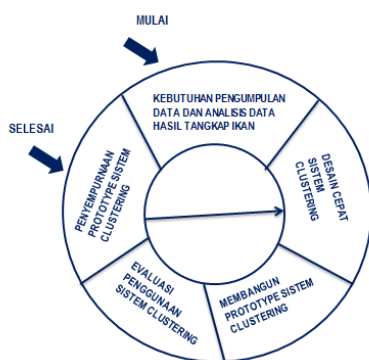
$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \dots\dots\dots 7$$

Dari persamaan tersebut, k merupakan *cluster* yang digunakan. Semakin kecil nilai DBI yang diperoleh (non-negatif  $\geq 0$ ), maka semakin baik *cluster* yang diperoleh dari pengelompokan *K-Means* yang digunakan [6].

### III. Metode Penelitian

#### 3.1. Metode Pengembangan Perangkat Lunak

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



Gambar 1 Tahapan *Prototype Clustering* Hasil Tangkap Ikan

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

1. Kebutuhan Pengumpulan dan Analisis Data Hasil Tangkap Ikan  
Mengumpulkan kebutuhan sistem yaitu dengan mengambil data penelitian di PPN Ternate kemudian hasilnya di analisa untuk menentukan kebutuhan sistem.
2. Desain yang Cepat Sistem *Clustering*  
Setelah menentukan kebutuhan sistem maka dilakukan desain secara cepat dengan pemodelan DFD dan ERD dari sistem *clustering* tangkap ikan.
3. Membangun *Protototype* Sistem *Clustering*  
Dari desain yang di pergunkan dibuat program yang merepresentasikan sistem untuk *meng-clustering* hasil produksi ikan dengan menerapkan algoritma *K-Means*
4. 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*.
5. Penyempurnaan *Protototype* Sistem *Clustering*  
Jika hasil dari evaluasi pengguna sistem *meng-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 sistem *clustering* produksi ikan.

#### 3.2. Pengumpulan Data

##### 1. Observasi

Peneliti melakukan observasi langsung ke Kantor Dinas PPN Ternate untuk mengambil data yang ada. Dengan demikian, peneliti dapat mengetahui data yang terkait dengan judul ini.

##### 2. Studi Literatur

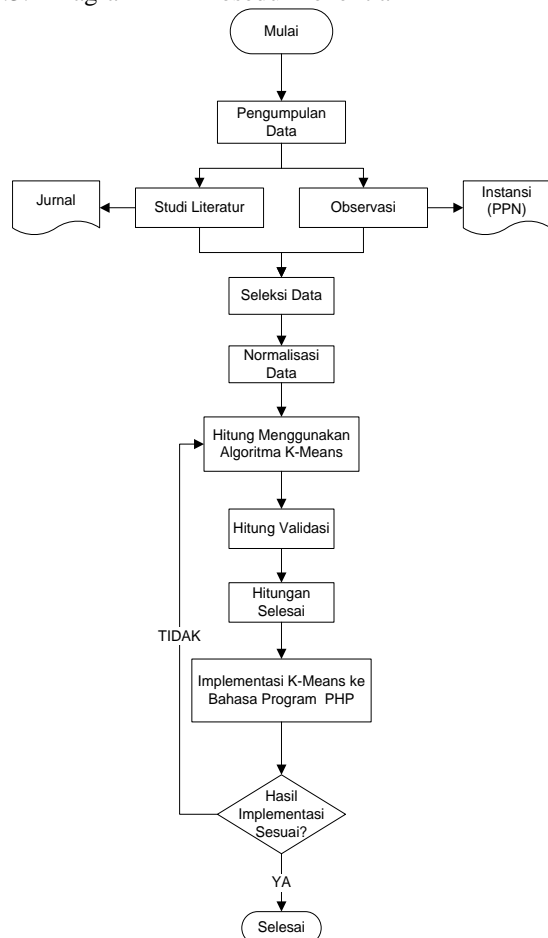
Pada studi literatur dimana peneliti mengambil dari jurnal yang terkait dengan judul penelitian. Seperti halnya jurnal tentang *clustering data mining* yang menggunakan metode seperti *K-Means*.

Dalam penelitian ini, terdapat dua jenis data, yaitu:

- a. Data primer  
Data primer dalam penelitian ini adalah data yang diambil dari Kantor Dinas PPN Ternate. Data yang dipergunakan oleh peneliti adalah data akomodasi hasil tangkap ikan perbulannya selama periode 2015 sampai 2017.
- b. Data sekunder  
Data sekunder yang dipergunakan dalam penelitian ini adalah teori-teori, penelitian

terdahulu, maupun keterangan-keterangan dari ahli yang berkompeten yang bersumber dari internet yang berisikan tentang berita terkini mengenai jenis-jenis ikan di Indonesia.

### 3.3. Diagram Alir Prosedur Penelitian



Gambar 2 Diagram Alir Penelitian

Langkah – langkah diagram alir penelitian:

1. Pengumpulan Data  
Peneliti mengambil data penelitian dengan observasi yaitu dengan mengambil data langsung ke instansi terkait dengan judul penelitian yakni PPN Ternate, sedangkan studi literatur yaitu peneliti mengambil data pada jurnal yang terkait dengan algoritma *K-Means*.
2. Seleksi Data  
Untuk memilih himpunan data (dataset) yang akan digunakan pada penulisan ini, yaitu data hasil tangkap ikan perbulan selama periode 2015-2017.
3. Normalisasi Data  
Sebelum data dihitung menggunakan algoritma *k-means*, semua data hasil tangkap ikan tahun 2015 sampai tahun 2017 dinormalisasi terlebih dahulu karena adanya perbedaan besaran angka yang cukup jauh pada masing-masing variabel data.
4. Hitung *Clustering* Menggunakan Algoritma *K-Means*.

Tahapan proses dimana data yang sudah dinormalisasi akan dihitung dan *cluster* dengan menggunakan cara kerja algoritma *K-Means*. Sebelum data dihitung peneliti menentukan kriteria sesuai dengan tahap *K-Means*. Karena yang akan dihitung adalah hasil tangkap ikan perbulan, maka dapat ditentukan kriteria yaitu: Januari, Februari, Maret, April, Mei, Juni, Juli, Agustus, September, Oktober, November, dan Desember.

5. Validasi Data  
Setelah proses *clustering* menggunakan algoritma *k-means* selesai, maka data *centroid* akhir pada tahap *k-means* akan di hitung menggunakan metode DBI (*Davies Bouldien Index*) untuk menentukan *cluster* terbaik yang digunakan pada penelitian ini.
6. Implementasi *K-Means* Menggunakan Bahasa Program PHP  
Pada tahap dimana data yang sudah dihitung menggunakan tahap *K-Means* akan diuji dalam program menggunakan bahasa pemrograman PHP.
7. Hasil Pengujian Sesuai?  
Pada tahap ini, jika hasil pengujian program tidak sesuai, maka diperbaiki bahasa program PHPnya, kemudian diuji lagi, jika belum sesuai maka kembali ke hitungan menggunakan tahap *K-Means* lalu diuji lagi dengan program, setelah hasil pengujian program sesuai, maka dilanjutkan ke kesimpulan.

### 3.4. Analisis Sistem Algoritma *K-means*

Didalam pembuatan laporan ini menggunakan algoritma *k-means* untuk menyelesaikan masalah penentuan *cluster* pada masing-masing jenis ikan hasil tangkap pada PPN Ternate dan penentuan keunggulan ikan pertahun. Sebelum tahap menghitung algoritma *k-means* menggunakan rumus pengukuran jarak (untuk mencari nilai pusat klaster atau *centroid* yang cocok), ditentukan kriteria terlebih dahulu, ada 12 kriteria yang ditentukan yaitu, bulan Januari, Februari, Maret, April, Mei, Juni, Juli, Agustus, September, Oktober, November, dan Desember. setelah hasil nilai pengukuran jarak didapat, maka akan ditentukan letak masing-masing klaster dalam bentuk tabel. Nilai pusat *cluster* akan didapat pada beberapa iterasi hingga nilai *cluster* tidak berpindah pada iterasi terakhir dan sudah terbagi menjadi dua *cluster*.

Sistem yang akan dibangun pada penelitian ini secara umum digunakan untuk mengimplementasikan hasil hitungan algoritma *k-means*, selain itu sistem ini juga dapat memberikan informasi yang dibutuhkan oleh pengguna, misalnya akan ditampilkan hasil *cluster* yang sudah dibentuk dan ikan yang paling unggul pada jenis-jenis ikan hasil tangkap pada instansi terkait, sehingga informasi tersebut dapat dimanfaatkan untuk memberitahukan misalnya kepada para nelayan, ikan apa yang paling unggul atau jenis ikan yang paling banyak, sehingga para nelayan sudah bisa mempersiapkan jenis alat tangkap ikan

yang sesuai (misalnya jaring lingkaran/ purse seine digunakan untuk menangkap ikan tongkol, layang, cakalang, kembung) dalam melakukan atau merencanakan penangkapan ikan selanjutnya yang dianggap paling unggul atau paling banyak terlihat diperaian Ternate.

### 3.5. Metode Pengujian Sistem

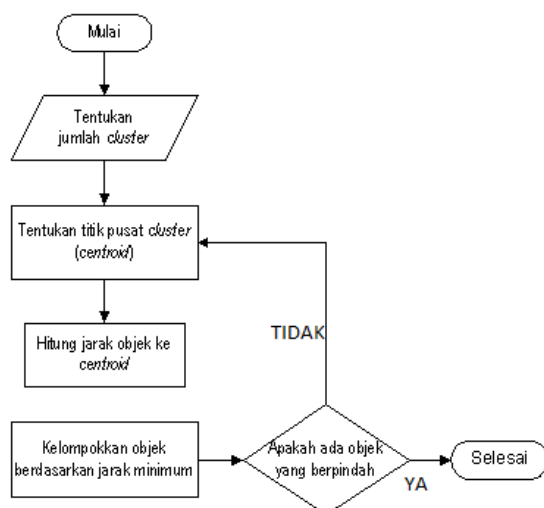
Pengujian perangkat lunak digunakan untuk mengetahui apakah perangkat lunak dapat berjalan dengan baik atau tidak, dalam penelitian ini pengujian perangkat lunak yang digunakan adalah *Black box*. *Black box* testing digunakan untuk mengetahui apakah program yang dibuat sudah sesuai dengan kebutuhan fungsionalnya atau tidak. Skenario pengujian dapat dilihat pada tabel 1.

Tabel 1 Skenario Pengujian

No	Kelas Uji	Butir Uji	Jenis Pengujian
1	Login	Verifikasi username	Blackbox
		Verifikasi password	
2	Input Centroid Awal	Input data centroid awal	Blackbox
3	Input Dataset	Input data tahun selanjutnya	Blackbox

## IV. HASIL DAN PEMBAHASAN

### 4.1. Flowchart Algoritma K-means



Gambar 4 Flowchart Algoritma 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 ditentukan peneliti sebanyak 12 yaitu: Januari, Februari, Maret, April, Mei, Juni, Juli, Agustus, September, Oktober, November, dan 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.

### 4.2. Hasil Hitung K-means dengan 2 Cluster Periode 2015 sampai 2017

Tabel 2 Posisi *Cluster* Iterasi Terakhir Tahun 2015

No	Jenis Ikan	$C_1$	$C_2$	Posisi <i>Cluster</i>
1	Layang	1,790	0,813	$C_2$
2	Tembang	0,142	2,669	$C_1$
3	Cakalang	3,323	0,813	$C_2$
4	Tongkol krai	1,052	1,962	$C_1$
5	Kembung	0,269	2,392	$C_1$
6	Madidihang	0,235	2,440	$C_1$
7	Julung	0,144	2,672	$C_1$
8	Lencam	0,09	2,524	$C_1$
9	Teri	0,132	2,665	$C_1$
10	Kakap	0,082	2,522	$C_1$
11	Kerapu	0,063	2,533	$C_1$
12	Baronang	0,08	2,604	$C_1$
13	Kwee	0,067	2,526	$C_1$
14	Tenggiri	0,153	2,682	$C_1$
15	Biji nangka	0,133	2,664	$C_1$
16	Selar	0,137	2,667	$C_1$
17	Swanggi	0,111	2,646	$C_1$
18	Lolosi biru	0,153	2,682	$C_1$

Deskripsi:

1. *Cluster* pertama memiliki pusat (0,064; 0,078; 0,042; 0,034; 0,031; 0,041; 0,033; 0,03; 0,019; 0,022; 0,042; 0,056) yang dapat diartikan sebagai sedikit. Ada 16 ikan yang termasuk kelompok ini, yaitu ikan Tembang, Tonkogkol krai, Kembung, Madidihang, Julung, Lencam, Teri, Kakap, Kerapu, Baronang, Kwee, Tenggiri, Biji nangka, Selar, Swanggi, Lolosi biru.
2. *Cluster* kedua memiliki pusat (0,774; 0,729; 0,836; 0,807; 0,74; 0,751; 0,81; 0,852; 0,813; 0,642; 0,714; 0,803) yang dapat diartikan sebagai banyak. Ada 2 ikan yang termasuk dalam kelompok ini yaitu ikan Layang dan Cakalang.

Tabel 3 Posisi *Cluster* Iterasi Terakhir Tahun 2016

No	Jenis Ikan	C <sub>1</sub>	C <sub>2</sub>	Posisi <i>Cluster</i>
1	Layang	2,632	0,639	C <sub>2</sub>
2	Tembang	0,122	2,948	C <sub>1</sub>
3	Cakalang	3,151	0,639	C <sub>2</sub>
4	Tongkol krai	0,129	2,956	C <sub>1</sub>
5	Kembung	0,345	2,555	C <sub>1</sub>
6	Madidihang	0,434	2,470	C <sub>1</sub>
7	Julung	0,107	2,934	C <sub>1</sub>
8	Lencam	0,126	2,719	C <sub>1</sub>
9	Teri	0,128	2,955	C <sub>1</sub>
10	Kakap	0,113	2,735	C <sub>1</sub>
11	Kerapu	0,085	2,761	C <sub>1</sub>
12	Baronang	0,081	2,902	C <sub>1</sub>
13	Kwee	0,082	2,763	C <sub>1</sub>
14	Tenggiri	0,129	2,955	C <sub>1</sub>
15	Biji nangka	0,087	2,913	C <sub>1</sub>
16	Selar	0,105	2,927	C <sub>1</sub>
17	Swanggi	0,076	2,892	C <sub>1</sub>
18	Lolosi biru	0,126	2,953	C <sub>1</sub>

Deskripsi:

1. *Cluster* pertama memiliki pusat (0,042; 0,052; 0,042; 0,045; 0,061; 0,017; 0,015; 0,039; 0,02; 0,026; 0,028; 0,029) yang dapat diartikan sebagai sedikit. Ada 16 ikan yang termasuk kelompok ini, yaitu ikan Tembang, Tongkol krai, Kembung, Madidihang, Julung, Lencam, Teri, Kakap, Kerapu, Baronang, Kwee, Tenggiri, Biji nangka, Selar, Swanggi, dan Lolosi biru.
2. *Cluster* kedua memiliki pusat (0,794; 0,964; 0,977; 0,965; 0,929; 0,744; 0,705; 0,975; 0,722; 0,839; 0,765; 0,787) yang dapat diartikan sebagai banyak. Ada 2 ikan yang termasuk dalam kelompok ini yaitu ikan Layang dan Cakalang.

Tabel 4 Posisi *Cluster* Iterasi Terakhir Tahun 2017

No	Jenis Ikan	C <sub>1</sub>	C <sub>2</sub>	Posisi <i>Cluster</i>
1	Layang	1,627	1,065	C <sub>2</sub>
2	Tembang	0,159	2,478	C <sub>1</sub>
3	Cakalang	3,23	1,065	C <sub>2</sub>
4	Tongkol krai	0,572	1,979	C <sub>1</sub>
5	Kembung	0,161	2,296	C <sub>1</sub>
6	Madidihang	0,655	1,732	C <sub>1</sub>
7	Julung	0,155	2,471	C <sub>1</sub>
8	Lencam	0,068	2,265	C <sub>1</sub>
9	Teri	0,179	2,496	C <sub>1</sub>
10	Kakap	0,077	2,256	C <sub>1</sub>
11	Kerapu	0,054	2,277	C <sub>1</sub>
12	Baronang	0,17	2,484	C <sub>1</sub>
13	Kwee	0,059	2,275	C <sub>1</sub>
14	Tenggiri	0,163	2,480	C <sub>1</sub>
15	Biji nangka	0,156	2,473	C <sub>1</sub>
16	Selar	0,161	2,476	C <sub>1</sub>
17	Swanggi	0,16	2,473	C <sub>1</sub>
18	Lolosi biru	0,07	2,351	C <sub>1</sub>

Deskripsi:

1. *Cluster* pertama memiliki pusat (0,04; 0,031; 0,037; 0,053; 0,059; 0,045; 0,036; 0,031; 0,046; 0,063; 0,084; 0,068) yang dapat diartikan sebagai sedikit. Ada 16 ikan yang termasuk kelompok ini, yaitu ikan Tembang, Tongkol krai, Kembung, Madidihang, Julung, Lencam, Teri, Kakap, Kerapu, Baronang, Kwee, Tenggiri, Biji nangka, Selar, Swanggi, dan Lolosi biru.
2. *Cluster* kedua memiliki pusat (0,716; 0,596; 0,918; 0,664; 0,754; 0,632; 0,66; 0,633; 0,591; 0,872; 0,869; 0,65) yang dapat diartikan sebagai banyak. Ada 2 ikan yang termasuk dalam kelompok ini yaitu ikan Layang dan Cakalang.

## V. PENUTUP

### 5.1. Kesimpulan

Berdasarkan hasil penelitian dalam *clustering* hasil tangkap ikan menggunakan *k-means*, dapat disimpulkan bahwa:

1. Algoritma *K-Means* telah digunakan untuk mengelompokkan hasil tangkap ikan periode 2015-2017 dengan menggunakan 2 *cluster*, yaitu *cluster* satu dikategorikan sebagai hasil tangkap sedikit, dan *cluster* dua dikategorikan sebagai hasil tangkap paling banyak. Dengan memperoleh hasil akhir yaitu jenis ikan dapat dikelompokkan kedalam masing-masing *cluster*. Ada 16 Jenis ikan pada *cluster* satu (C<sub>2</sub>) adalah ikan tembang, tongkol krai, kembung, madidihang, julung, lencam, teri, kakap, kerapu, baronang, kwee, tenggiri, biji nangka, selar, swanggi dan lolosi biru. Dan ada 2 jenis ikan pada *cluster* dua yaitu ikan layang dan cakalang. Maka dari itu dapat dilihat bahwa jenis ikan yang unggul adalah pada *cluster* dua (C<sub>2</sub>).
2. *K-means* mempunyai kelemahan yang diakibatkan oleh penentuan pusat awal *cluster* atau *centroid*. Hasil *cluster* yang terbentuk dari algoritma *k-means* sangatlah bergantung pada inisiasi nilai pusat awal *cluster* yang ditentukan, hal ini menyebabkan sangat sulit memperoleh hasil *centroid* awal yang unik, sehingga peneliti menggunakan metode validasi DBI untuk mengevaluasi hasil *clustering* yang diperoleh.
3. Peneliti melakukan perhitungan beberapa *cluster*, yaitu 2 *cluster*, 3 *cluster*, 4 *cluster*, dan 5 *cluster* untuk menentukan perbandingan *cluster* terbaik dengan menggunakan metode validasi DBI (*Davies Bouldin Index*), hasil nilai yang diperoleh:
  - a. 2 *Cluster* menghasilkan nilai DBI pada tahun 2015 = 0,395, tahun 2016 = 0,276 dan tahun 2017 = 0,54.
  - b. 3 *Cluster* menghasilkan nilai DBI pada tahun 2015 = 1,126, tahun 2016 = 0,541, dan tahun 2017 = 0,692.
  - c. 4 *Cluster* menghasilkan nilai DBI pada tahun 2015 = 2,052, tahun 2016 = 0,704, dan tahun 2017 = 0,568.



- d. 5 Cluster menghasilkan nilai DBI pada tahun 2015 = 0,734, tahun 2016 = 0,715, dan tahun 2017 = 0,755.  
Jika nilai DBI mendekati nilai 0 maka cluster tersebut dianggap baik, pada option a sampai d terlihat bahwa nilai DBI pada 2 cluster tahun 2015 sampai 2017 adalah yang paling mendekati nilai 0, maka 2 cluster yang dipakai pada penelitian ini dapat dinyatakan baik dibandingkan dengan beberapa cluster lainnya.
4. Perhitungan algoritma telah diimplementasikan pada aplikasi yang dibangun menggunakan bahasa program PHP hasilnya sama dengan perhitungan yang dilakukan secara manual.
- 5.2. Saran
1. Penelitian ini menggunakan algoritma *K-means* dan telah diketahui apa kelemahannya, diharapkan pada penelitian selanjutnya agar menggunakan algoritma lainnya, kemudian bandingkan hasilnya.
  2. Menggunakan metode validasi untuk evaluasi cluster lebih dari 1, agar mendapatkan hasil cluster yang lebih baik.
  3. Sistem yang dibangun masih terdapat kekurangan dan keterbatasan, yaitu sistemnya masih terlihat sederhana dan hanya berfokus pada metode yang digunakan, bukan di desain *interface*-nya, jadi diharapkan pada peneliti selanjutnya agar bisa dikembangkan tampilan *interface*-nya lebih mudah digunakan.
  4. Sistem yang bangun menggunakan bahasa program PHP, maka dari itu diharapkan pada peneliti selanjutnya dapat menggunakan bahasa program lainnya, misalnya bahasa program *Java*, *Phyton*.

- Kendal). Jurusan Teknik Informatika FIK UDINUS.
- [7] Akhmad Sholikhin dan Berliana Kusuma, 2013. Pembangunan Sistem Informasi Investasi Sekolah Pada Dinas Pendidikan Kabupaten Rembang Berbasis Web. Indonesia *Jurnal on Networking and Security (IJNS)*.
- [8] Vindhy Agus Setiawan, 2011. *Analisis dan Perancangan Sistem Informasi Simpan Pinjam Pada LKM Gerembeng Bali*. Sekolah Tinggi Manajemen Informatika dan Komputer Yogyakarta.
- [9] Tempola, F., dan Assagaf, A, F. 2018. Clustering of Potency of Shrimp in Indonesia with K-Means Algorithm and Validation of Davies-Bouldin Index. *International Conference on Science and Technology (ICST 2018)*.730-733. (<https://www.atlantis-pub.com/proceedings/icst-18/55910933> : diakses 27 November 2018)

#### DAFTAR PUSTAKA

- [1] Edi Mirmanto, 2010. *Komposisi Flora dan Struktur Hutan Alami di Pulau Ternate*. Bidang Botani Pusat Penelitian Biologi –LIPI.
- [2] Statistik Perikanan Tangkap PPN Ternate, 2015. Kementrian Kelautan dan Perikanan. Pelabuhan Perikanan Nusantara Ternate.
- [3] Johan Oscar Ong, 2013. *Implementasi Algoritma K-Means Clustering untuk Menentukan Strategi Marketing President University*. President University.
- [4] Dwi Novianti Nango, 2012. *Penerapan Algoritma K-Means Untuk Clustering Data Anggaran pendapatan Belanja Daerah di Kabupaten*. Universitas Negeri Gorontalo.
- [5] Eni Irviani, 2018. *Algoritma K-Means Clustering untuk Menentukan Nilai Gizi Balita*. Universitas Bima Sarana Informatika.
- [6] Alith Fajar Muhammad, 2018. *Klasterisasi Proses Seleksi Pemain Menggunakan Algoritma K-Means (Study Kasus: Tim Hockey Kabupaten*