

CLUSTERING DATA EKSPOR BUAH-BUAHAN BERDASARKAN NEGARA TUJUAN MENGGUNAKAN ALGORITMA K-MEANS

Haviz Atma Negara¹, Achmad Rizaldi Putra², Ultach Enri³

^{1,2,3}Teknik Informatika, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang,
Karawang41361, Indonesia
¹haviz.atma17109@student.unsika.ac.id, ²achmad.rizaldi17033@student.unsika.ac.id,
³ultach@staff.unsika.ac.id

ABSTRAK

Indonesia merupakan salah satu negara pengekspor ke negara-negara maju dan berkembang. Eksportir bertujuan untuk memperoleh keuntungan demi menunjang perekonomian dan kemakmuran masyarakat. Penelitian ini membahas tentang penerapan data mining pada ekspor buah-buahan menurut negara tujuan menggunakan k-means clustering method. Data yang digunakan dalam penelitian ini adalah data Ekspor Buah-buahan Menurut Negara Tujuan Utama dari tahun 2012-2019 yang diambil dari dokumen-dokumen keterangan ekspor impor yang dihasilkan oleh Direktorat Jenderal Bea dan Cukai. Data akan diolah dengan melakukan clustering dalam 3 cluster yaitu cluster tingkat ekspor tinggi, cluster tingkat ekspor sedang dan cluster tingkat ekspor rendah. Centroid data untuk cluster tingkat ekspor tinggi sebesar 2054519.3, centroid data untuk cluster tingkat ekspor sedang sebesar 489020.3, centroid data untuk cluster tingkat ekspor rendah sebesar 20.2. Sehingga diperoleh penilaian berdasarkan indeks ekspor buah-buahan dengan 2 negara *cluster* tingkat ekspor tinggi yakni negara Tiongkok & Malaysia, 2 negara *cluster* tingkat ekspor rendah yakni Vietnam & Thailand, dan 6 negara *cluster* tingkat ekspor rendah yakni Hongkong, Singapura, Nigeria, India, Jepang, Uni Emirat Arab. Sehingga diperoleh informasi hasil pengelompokan data baru yang dapat menjadi masukan bagi pemerintah maupun perusahaan-perusahaan agar memprioritaskan serta dapat lebih meningkatkan kegiatan ekspor buah-buahan berdasarkan klaster yang telah dilaksanakan.

Kata Kunci— *Clustering, K-Means, Rapid Miner, Data Mining, Ekspor, Buah-buahan.*

ABSTRACT

Indonesia is one of the exporting countries to developed and developing countries. Exporters aim to obtain profits for the sake of supporting the economy and the prosperity of society. This study discusses the application of data mining to fruit exports according to destination countries using the k-means clustering method. The data used in this research is the data on the export of fruits by main destination countries from 2012-2019 which are taken from the export-import information documents produced by the Directorate General of Customs and Excise. The data will be processed by clustering in 3 clusters, namely the high export level cluster, the medium export level cluster and the low export level cluster. The data centroid for the high export level cluster is 2054519.3, the data centroid for the medium export level cluster is 489020.3, the data centroid for the low export level cluster is 20.2. So that an assessment is obtained based on the fruit export index with 2 high export cluster countries, namely China & Malaysia, 2 countries with low export level clusters namely Vietnam & Thailand, and 6 countries with low export level clusters namely Hong Kong, Singapore, Nigeria, India, Japan, , United Arab Emirates. In order to obtain information on the results of new data grouping that can be input for the government and companies to prioritize and be able to further increase the export activities of fruits based on the clusters that have been implemented.

Keywords— *Clustering, K-Means, Rapid Miner, Data Mining, exports, fruits.*

1. PENDAHULUAN

Sejak dahulu Indonesia sudah dikenal sebagai salah satu negara agraris, yang artinya sektor pertanian memegang peranan penting dari keseluruhan perekonomian nasional. Hal ini dapat ditunjukkan dari banyaknya penduduk atau tenaga kerja yang hidup atau bekerja dalam sektor pertanian atau dari produk nasional yang berasal dari sektor pertanian.

(Soekartawi, 1999) Pertanian dalam arti luas terdiri dari lima sektor, yaitu tanaman pangan, perkebunan, peternakan, perikanan dan kehutanan. Kelima sektor pertanian tersebut bila ditangani dengan serius sebenarnya akan mampu memberikan sumbangan yang besar bagi perkembangan perekonomian Indonesia mendatang. Salah satu cara penanganannya yaitu dengan berorientasi pada bisnis pertanian atau agrobisnis.[1]

Seiring berjalannya waktu dirasa sangat disayangkan jika Indonesia tidak memanfaatkan hasil alamnya dengan maksimal, demi menunjang perekonomian dan kemakmuran masyarakat ekspor pangan dilakukan untuk mendapatkan keuntungan lebih besar yang bisa dimanfaatkan untuk menunjang pembangunan ekonomi nasional.

(Agus Perdana Windarto., 2017) pada penelitiannya melakukan penilaian terhadap hasil ekspor buah-buahan berdasarkan negara tujuan dengan menerapkan metode clustering K-Means. Data diolah untuk memperoleh nilai dari produksi ekspor buah-buahan berdasarkan negara tujuan. Data tersebut diolah menggunakan Rapidminer untuk ditentukan nilai centroid dalam 3 cluster yaitu cluster tingkat ekspor tinggi, cluster tingkat ekspor sedang dan cluster tingkat ekspor rendah. Centroid data untuk cluster tingkat ekspor tinggi 904.276,5, Centroid data untuk cluster tingkat ekspor sedang 265.501 dan Centroid data untuk cluster tingkat ekspor rendah 34.280,1. Sehingga diperoleh penilaian berdasarkan indeks ekspor buah-buahan dengan 2 negara cluster tingkat ekspor tinggi yakni India dan Pakistan, 3 negara cluster tingkat ekspor sedang yakni Singapura, Bangladesh dan Negara lainnya dan 6 negara cluster tingkat ekspor rendah yakni Hongkong, Tiongkok, Malaysia, Nepal, Vietnam dan Iran. Hasil yang dari penelitian dapat digunakan untuk mengetahui jumlah ekspor buah-buahan menurut negara tujuan.[2]

Berdasarkan uraian diatas tujuan dari penelitian ini adalah menerapkan algoritma k-means untuk mengcluster ekspor buah-buahan berdasarkan negara tujuan. Nantinya hasil dari penelitian ini dapat menjadi masukan kepada pemerintah, negara yang menjadi prioritas tertinggi pada kegiatan ekspor buah-buahan berdasarkan klaster yang telah dilakukan.[3]

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data Mining merupakan disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data yang besar. Data

mining bekerja dengan metode mengolah data terstruktur dengan menggunakan konsep machine learning. Berikut merupakan pendapat beberapa ahli tentang data mining :

[1] Menurut (Witten *et al.*, 2011) Melakukan ekstraksi untuk mendapatkan informasi penting yang sifatnya implisit dan sebelumnya tidak diketahui, dari suatu data.

[2] Menurut (Darmawan *et al.*, 2018) Data Mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika.

[3] Sedangkan menurut (Han *et al.*, 2011) Ekstraksi pola atau pengetahuan yang menarik (non-trivial, implisit, yang sebelumnya tidak diketahui dan berpotensi berguna dari data dalam jumlah yang besar.

2.2. K-Means Clustering

Salah satu metode yang cukup banyak dipakai adalah clustering. Clustering artinya membagi data yang tidak memiliki label menjadi beberapa kelompok. Variabel yang digunakan dalam penelitian ini : jumlah ekspor berat bersih (ton) dan nilai Free On Board (FOB). FOB adalah penyerahan barang yang dilakukan diatas kapal yang selanjutnya akan melakukan pengangkatan barang. Data ekspor buah-buahan Menurut Negara Tujuan Utama dari tahun 2012-2019 akan diolah dengan menggunakan clustering yang dibagi menjadi 3 cluster yaitu cluster tingkat ekspor tinggi, cluster sedang dan cluster rendah.

Pengelompokan data dengan metode K-Means dilakukan dengan algoritma sebagai berikut :

1. Tentukan jumlah kelompok yang akan dibagi.
2. Alokasikan data ke dalam kelompok secara acak pusat kelompok.
3. Hitung (centroid/rata-rata) dari data yang ada di masing-masing kelompok. Lokasi centroid setiap kelompok diambil dari rata-rata (mean) semua nilai data pada setiap fiturnya. Jika M menyatakan jumlah data dalam sebuah kelompok, i menyatakan fitur ke-i dalam sebuah kelompok, dan p menyatakan dimensi untuk data, maka persamaan menghitung centroid fitur ke-i digunakan persamaan.
4. Persamaan 1 dilakukan sebanyak p dimensi dari i=1 sampai dengan i=p , menggunakan rumus 1 :

$$C_i = \sum_{j=1}^M x_j^i$$

(1)

5. Alokasikan masing-masing data ke centroid/rata-rata terdekat. Ada beberapa cara yang dapat dilakukan untuk mengukur jarak data ke pusat kelompok, diantaranya adalah Euclidean. Pengukuran jarak pada ruang jarak (distance space) Euclidean dapat dicari menggunakan rumus 2 :

$$d = \sqrt{(x1 - x2)^2 + (y1 - y2)^2}$$

(2)

Pengalokasian kembali data ke dalam masing-masing kelompok dalam metode K-Means didasarkan pada perbandingan jarak antara data dengan centroid setiap kelompok yang ada. Data dialokasikan ulang ke kelompok yang memiliki centroid jarak terdekat secara tegas. Pengalokasian data ini menggunakan rumus 3 :

$$ail = \{0 \quad d = \min\{D(xi, ci)\}$$

(3)

Fungsi objektif yang digunakan pada metode K-Means dihitung berdasarkan jarak dan nilai keanggotaan dalam kelompok, menggunakan rumus 4 :

$$J = \sum_{i=1}^n \sum_{i=1}^k a_{ic} D(xi, ci)^2$$

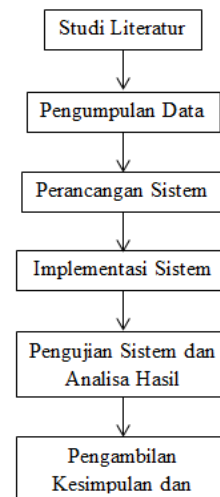
(4)

n adalah jumlah data, k adalah jumlah kelompok, a_{i1} adalah nilai keanggotaan titik data x_i ke kelompok c_1 yang diikuti. a mempunyai nilai 0 atau 1. Apabila data merupakan anggota suatu kelompok, nilai $a_{i1} = 1$. Jika tidak, nilai $a_{i1} = 0$. Jika tidak, nilai $a_{i1} = 0$.

6. Kembali ke langkah 3, jika masih ada data yang berpindah kelompok, ada perubahan nilai centroid yang telah ditentukan, atau apabila perubahan nilai pada fungsi objektif yang digunakan masih di atas nilai ambang yang ditentukan.

3. METODE YANG DIUSULKAN

Penelitian ini menggunakan dataset ekspor buah-buahan berdasarkan negara tujuan. Dataset tersebut kemudian dikelompokkan menggunakan algoritma K-Means.



Gambar 3.1. Alur Penelitian

Pengelompokkan data ekspor buah-buahan dengan menggunakan algoritma K-Means ini dilakukan dengan beberapa tahapan, yaitu :

3.1 Tahap Pengumpulan data

Dalam penerapan data mining produksi ekspor buah-buahan menurut negara tujuan, diperlukan data terkait tentang hal itu. Sumber data penelitian diperoleh dari data yang dikumpulkan berdasarkan dokumen-dokumen keterangan ekspor impor yang dihasilkan oleh Direktorat Jenderal Bea dan Cukai melalui situs <https://www.bps.go.id>. Selain itu sejak tahun 2015 data ekspor juga berasal dari PT. Pos Indonesia, catatan instansi lain di perbatasan, dan hasil survei perdagangan lintas batas laut. Data yang digunakan dalam penelitian ini adalah data Ekspor Buah-buahan Menurut Negara Tujuan Utama dari tahun 2012-2019 yang terdiri dari 10 negara yakni Malaysia, Vietnam, Tiongkok, Thailand, India, Jepang, Hongkong, Uni Emirat Arab, Singapura dan Nigeria. Variabel yang digunakan adalah jumlah ekspor berat bersih (netto) dan nilai Free On Board (FOB). Data akan diolah dengan melakukan clustering ekspor buah-buahan berdasarkan negara tujuan utama dalam 3 cluster yakni cluster tingkat ekspor tinggi, cluster tingkat ekspor sedang dan cluster tingkat ekspor rendah.

3.2 Tahap Pengolahan Data Menggunakan Microsoft Excel

Tahapan ini merupakan tahapan pengolahan data menggunakan algoritma K-Means untuk mendapatkan kelompok data ekspor buah-buahan berdasarkan negara tujuan. Data yang telah diperoleh akan diolah terlebih dahulu untuk dapat dikluster. Dalam tahap sebelumnya, data setiap negara tujuan ekspor buah-buahan akan dijumlah setiap aspeknya sehingga pada tahapan ini sudah diperoleh perhitungan nilai yang akan diproses pada tahap clustering.

3.3 Tahap Implementasi K-Means Cluster Dengan Rapid Miner

Sebagai bahan perbandingan pengolahan data yang sudah dilakukan menggunakan Microsoft Excel juga dilakukan pengolahan data kembali menggunakan aplikasi Rapminer Studio. Clustering merupakan klasifikasi tanpa pengawasan dan merupakan proses partisi sekumpulan objek data dari satu set menjadi beberapa kelas. Hal ini dapat dilakukan dengan menerapkan berbagai persamaan dan langkah-langkah mengenai jarak algoritma, yaitu dengan Euclidean Distance. Analisis kluster ialah metode yang dipakai untuk membagi rangkaian data menjadi beberapa grup berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya. Dalam menentukan cluster berdasarkan data yang telah tersedia, dibutuhkan sebuah flowchart untuk memudahkan dalam menentukan alur perhitungan sebagai alur untuk menemukan hasil dari penerapan cluster terhadap data yang akan diproses.

3.4 Tahap Analisis Data

Pada tahapan ini dilakukan analisis data ekspor buah buahan menurut negara tujuan aplikasi tool. Rapid Minner. RapidMiner adalah sebuah lingkungan machine learning data mining, text mining dan predictive analytics [8]. Data yang diperoleh diolah dengan menggunakan perhitungan bobot dari tiap indeks. Pada tahapan sebelumnya, telah ditentukan akan dicluster ke dalam 3 cluster yakni cluster tingkat ekspor tinggi, cluster tingkat ekspor sedang dan cluster tingkat ekspor rendah. Pada tahapan inilah akan dianalisis hasilnya.

4. HASIL PENELITIAN

Dalam melakukan clustering, data yang diperoleh akan dihitung terlebih dahulu berdasarkan jumlah ekspor buah buahan pada tahun 2012-2019 berdasarkan negara tujuan. Hasil penjumlahan berdasarkan 2 kriteria penilaian yakni berat bersih (netto) dan nilai FOB seperti yang ditunjukkan pada gambar 4.1

Negara Tujuan	2012	2013	2014	2015	2016	2017	2018	2019	TOTAL
Berat bersih/Net weight : Ton									
Malaysia	112 023,0	190 794,9	210 025,4	267 848,8	352 106,3	283 836,2	320 337,9	317 546,8	2054 519,3
Vietnam	20 284,8	28 224,8	41 654,4	61 466,5	58 223,8	52 210,7	47 619,3	61 454,1	371 138,4
Tiongkok	32 634,8	23 308,2	158 424,4	206 541,1	204 561,7	268 223,9	188 629,4	200 303,3	1282 626,8
Thailand	15 289,4	18 926,3	72 197,0	80 622,7	112 885,6	384 586,2	165 429,3	76 551,2	926 567,7
India	49 632,7	10 842,6	11 647,8	24 544,6	10 130,7	13 123,8	8 821,9	12 215,0	161 165,1
Jepang	74,2	71,4	71,2	3 240,7	7 377,4	4 915,7	4 675,1	4 773,9	25 203,6
Hongkong	10 128,8	2 102,1	1 871,9	7 251,3	5 926,8	1 596,0	3 620,9	12 153,1	44 650,9
Uni Emirat arab	2 272,2	1 977,4	4 238,0	5 514,4	4 133,0	4 044,7	10 377,6	14 377,3	46 934,6
Singapura	9 456,1	8 757,7	6 820,8	6 004,0	4 481,4	5 449,7	6 193,5	5 912,8	53 076,0
Nigeria	0,4	0,6	1,1	2,5	8,7	1,6	3,7	2,6	20,2
Jumlah	252 096,4	295 006,0	506 947,0	663 036,6	759 835,4	1017 972,5	755 708,6	715 300,1	4965 902,6
Nilai FOB/FOB value : 000 US\$									
Malaysia	18 107,2	25 295,2	25 983,9	33 885,1	89 807,0	41 239,6	71 619,8	62 980,3	342 887,5
Vietnam	23 594,8	33 593,7	57 926,4	83 185,4	91 234,5	87 133,3	73 022,7	87 420,8	527 081,6
Tiongkok	16 031,3	8 007,5	49 400,8	51 210,2	51 939,5	71 745,9	50 815,2	54 708,2	359 858,6
Thailand	3 372,9	5 691,9	15 351,6	20 513,7	31 142,5	107 104,7	42 325,7	15 642,4	241 145,4
India	55 746,3	24 617,5	14 654,7	36 142,2	16 851,7	28 155,9	14 645,1	33 447,5	224 260,9
Jepang	126,2	135,7	183,5	2 119,3	5 076,6	2 636,0	2 730,2	2 828,4	15 837,9
Hongkong	7 470,3	1 823,1	1 544,6	4 252,0	4 617,2	1 343,9	3 393,0	19 019,6	43 464,7
Uni Emirat arab	1 998,6	1 695,9	3 202,5	4 005,1	3 383,5	3 121,1	6 248,8	8 904,6	32 560,1
Singapura	7 465,9	6 681,0	5 179,2	4 168,0	2 718,9	3 946,4	4 589,1	4 948,7	39 897,2
Nigeria	119,0	193,6	319,6	844,8	1 817,7	427,2	472,7	456,0	4 650,6
Jumlah	134 034,6	107 933,1	173 744,4	238 276,8	274 591,1	346 854,0	269 852,1	290 356,3	1835 644,6

Gambar 4.1 Data ekspor buah-buahan tahun 2012-2019

Selanjutnya, tahap pertama dalam proses pengolahan data yaitu mengakumulasi nilai yang

diambil dari data ekspor buah-buahan berdasarkan negara tujuan yang dapat dilihat pada tabel 4.1.

Tabel 4.1 Data Akumulasi Ekspor Buah-buahan

NO	Negara Tujuan	Netto (TON)	Nilai FOB (Ribu US\$)
1	Hongkong	44650,9	43464,7
2	Tiongkok	1282626,8	353858,6
3	Singapura	53076	39897,2
4	Malaysia	2054519,3	342887,5
5	Nigeria	20,2	4650,6
6	Vietnam	371138,4	531081,6
7	India	16165,1	224260,9
8	Thailand	926567,7	241145,4
9	Jepang	25203,6	15837,9
10	Uni Emirat Arab	46924,6	32560,1

4.1 Input Data

Data yang digunakan pada penelitian ini adalah data ekspor buah-buahan berdasarkan negara tujuan. Data ini diperoleh dari data yang dikumpulkan berdasarkan dokumen-dokumen keterangan ekspor impor yang dihasilkan oleh Direktorat Jenderal Bea dan Cukai melalui situs <https://www.bps.go.id>. Data yang telah diakumulasi akan dimasukkan ke dalam tool Rapid Minner. Sehingga dapat dicluster menjadi dua. Setelah dimasukkan ke dalam Rapid Minner.

4.2 Centroid Data

Dalam penerapan algoritma *k-means* dihasilkan nilai titik tengah atau centroid dari data yang didapat dengan ketentuan bahwa *clusterisasi* yang diinginkan adalah 3, penentuan *cluster* dibagi atas tiga bagian yakni *cluster* tingkat ekspor tinggi (C1), *cluster* tingkat ekspor sedang (C2), dan *cluster* tingkat ekspor rendah (C3). Maka nilai titik tengah atau centroid juga terdapat 3 titik. Penentuan titik *cluster* ini dilakukan dengan mengambil nilai terbesar (*maximum*) untuk *cluster* tingkat ekspor tinggi (C1), nilai rata-rata (*average*) untuk *cluster* tingkat ekspor sedang (C2), dan nilai terkecil (*minimum*) untuk *cluster* tingkat ekspor rendah (C3). Nilai titik tersebut dapat diketahui pada tabel 4.2

• Iterasi Pertama

Tabel 4.2 Centroid Data Iterasi 1

Atribut	Cluster 1 (Tinggi)		Cluster 2 (Sedang)		Cluster 3 (Rendah)	
Nilai	Netto	FOB	Netto	FOB	Netto	FOB
	2054519 .3	5370 81.6	482090. 3	1835 64.5	20.2 0.6	465 0.6

4.3 Clustering Data

Dengan menggunakan centroid tersebut maka dapat dilakukan proses *clustering* data yang telah didapat menjadi 3 *cluster*. Proses *cluster* dengan mengambil jarak terdekat dari setiap data yang diolah. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 1 dan *clustering* data dapat digambarkan pada gambar 4.2.

- *Menghitung jarak setiap data ke pusat cluster*
Jarak setiap data ke pusat *cluster* ditentukan dengan rumus *Euclidian Distance*.

NO	Negara Tujuan	Netto (TON)	Nilai FOB (Ribu US\$)	C1	C2	C3	Jarak Terpendek
1	Hongkong	44650.9	43464.7	2069596	459326.8	59147.6	59147.6
2	Tiongkok	1282626.8	353858.6	793340.2	818449.1	1329295.3	793340.2
3	Singapura	53076	39897.2	2062272	452430.7	63696.5	63696.5
4	Malaysia	2054519.3	342887.5	194194.1	1580480	2082155.3	194194.1
5	Nigeria	20.2	4650.6	2122369	514200.1	0.0	0.0
6	Vietnam	371138.4	537081.6	1683381	370519.5	649008.1	370519.5
7	India	16165.1	224260.9	2062218	467699.1	220203.0	220203.0
8	Thailand	926567.7	241145.4	1166127	448191.7	956253.1	448191.7
9	Jepang	25203.6	15837.9	2095189	486700.7	27556.5	27556.5
10	Uni Emirat Arab	46934.6	32560.1	2070009	460611.3	54588.5	54588.5
	cluster 1 (tinggi)	2054519.3	537081.6				
	cluster 2 (sedang)	482090.3	183564.5				
	cluster 3 (rendah)	20.2	4650.6				

Gambar 4.2 Perhitungan Jarak Pusat Cluster Iterasi 1

- *Pengelompokan data kedalam cluster*
Data yang mempunyai jarak terdekat akan masuk kedalam *cluster* tersebut.

Pengelompokan Objek				
NO	Negara Tujuan	Netto (TON)	Nilai FOB (Ribu US\$)	Kelompok
1	Hongkong	44650.9	43464.7	3
2	Tiongkok	1282626.8	353858.6	1
3	Singapura	53076	39897.2	3
4	Malaysia	2054519.3	342887.5	1
5	Nigeria	20.2	4650.6	3
6	Vietnam	371138.4	537081.6	2
7	India	16165.1	224260.9	3
8	Thailand	926567.7	241145.4	2
9	Jepang	25203.6	15837.9	3
10	Uni Emirat Arab	46934.6	32560.1	3

Gambar 4.3 Pengelompokan Data Iterasi 1

- Iterasi Kedua

Setelah proses iterasi 1 selesai perlu dilakukan kembali proses iterasi sampai menemukan hasil yang sama dengan proses iterasi 1. Pada tahap selanjutnya kembali dilakukan proses mencari nilai titik tengah atau centroid untuk iterasi 2 yang dapat dilihat pada tabel 4.3

Tabel 4.3 Centroid Data Iterasi 2

Atribut	Cluster 1		Cluster 2		Cluster 3	
	(Tinggi)		(Sedang)		(Rendah)	
Nilai	Netto	FOB	Netto	FOB	Netto	FOB

	1668573	3483	648853,	3891	31008,4	601
		73,1	1	13,5		11,9

Setelah mendapatkan nilai titik tengah atau centroid, proses sama dilakukan dengan mencari jarak terdekat. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 2 dapat digambarkan pada gambar 4.4.

- *Menghitung jarak setiap data ke pusat cluster*
Jarak setiap data ke pusat *cluster* ditentukan dengan rumus *Euclidian Distance*.

penentuan cluster baru							
clust netto (TON)	Nilai FOB (US\$)						
I	1668573	348373.1					
II	648853.1	389113.5					
III	31008.4	60111.9					
NO	Negara Tujuan	Netto (TON)	Nilai FOB (Ribu US\$)	C1	C2	C3	Jarak Terpendek
1	Hongkong	44650.9	43464.7	1652299	696084.3	21523.2	21523.2
2	Tiongkok	1282626.8	353858.6	385985.2	634753.5	1285626.6	385985.2
3	Singapura	53076	39897.2	1644685	690581.2	29926.8	29926.8
4	Malaysia	2054519.3	342887.5	385985.3	1406426	2043173.6	385985.3
5	Nigeria	20.2	4650.6	1703588	754185.6	63531.3	63531.3
6	Vietnam	371138.4	537081.6	1311086	314674.5	585822.9	314674.5
7	India	16165.1	224260.9	1657062	653812.3	164818.7	164818.7
8	Thailand	926567.7	241145.4	749713	314674.4	913673.7	314674.4
9	Jepang	25203.6	15837.9	1676676	726824.2	44652.9	44652.9
10	Uni Emirat Arab	46934.6	32560.1	1652104	699597.2	31823.7	31823.7

Gambar 4.4 Perhitungan Jarak Pusat Cluster Iterasi 2

- *Pengelompokan data kedalam cluster*
Data yang mempunyai jarak terdekat akan masuk kedalam *cluster* tersebut.

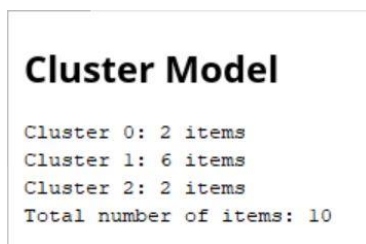
Pengelompokan Objek Baru				
NO	Negara Tujuan	Netto (TON)	Nilai FOB (Ribu US\$)	Kelompok
1	Hongkong	44650.9	43464.7	3
2	Tiongkok	1282626.8	353858.6	1
3	Singapura	53076	39897.2	3
4	Malaysia	2054519.3	342887.5	1
5	Nigeria	20.2	4650.6	3
6	Vietnam	371138.4	537081.6	2
7	India	16165.1	224260.9	3
8	Thailand	926567.7	241145.4	2
9	Jepang	25203.6	15837.9	3
10	Uni Emirat Arab	46934.6	32560.1	3

Gambar 4.5 Pengelompokan Data Iterasi 2

Berdasarkan pengamatan yang telah dilakukan karena hasil dari iterasi 1 dan iterasi 2 sama seperti yang ditunjukkan oleh gambar 4.3 dan gambar 4.5, maka tidak perlu dilakukan proses iterasi kembali.

4.4 Analisa Data

Pengolahan dataset ekspor buah-buahan berdasarkan negara tujuan telah didapatkan, kemudian akan dilakukan pengolahan dataset kembali menggunakan Rapidminer sebagai bahan perbandingan. Hasil yang didapatkan berdasarkan Rapidminer sebagai berikut :



Gambar 4.6 Hasil Cluster Algoritma K-Means

Gambar diatas menunjukkan hasil dari 10 data ekspor buah-buahan berdasarkan negara tujuan dengan menggunakan algoritma K-Means yang terbagi menjadi 3 cluster yaitu cluster 0 sebanyak 2 negara, cluster 1 sebanyak 6 negara, cluster 2 sebanyak 2 negara.

Clustering data ekspor buah-buahan berdasarkan negara tujuan dengan algoritma K-Means yang menggunakan Microsoft Excel maupun rapidminer menghasilkan pengelompokkan yang sama pada setiap kelompoknya. Berdasarkan hasil pengolahan data didapatkan *Cluster* tingkat ekspor tinggi (C1) yang mempunyai ekspor buah-buahan tinggi yakni Negara Tiongkok & Malaysia. *Cluster* tingkat ekspor sedang (C2) yang mempunyai ekspor buah-buahan sedang yakni Negara Vietnam & Thailand. *Cluster* tingkat ekspor rendah (C3) yang mempunyai ekspor buah-buahan rendah yakni Negara Hongkong, Singapura, Nigeria, India, Jepang, Uni Emirat Arab.

Paramater untuk mengukur kinerja algoritma K-Means dilakukan dengan menghitung Davies Bouldin Index dan Avg. Within centroid distance. Davies Bouldin Index merupakan algoritma yang menghasilkan *cluster* dengan jarak *interr-cluster* rendah (kesamaan *inter-cluster tinggi*) dan jarak antar *cluster* tinggi (kemiripan antar *cluster* rendah) akan memiliki Davies Bouldin Index yang rendah, algoritma *cluster* yang menghasilkan kumpulan *cluster* dengan indeks Davies Bouldin terkecil dianggap sebagai algoritma terbaik berdasarkan kriteria ini. Sedangkan Avg. Within centroid distance merupakan rata-rata dalam jarak *cluster* dihitung dengan rata-rata jarak antara centroid dan semua contoh *cluster*.

PerformanceVector

```

PerformanceVector:
Avg. within centroid distance: 26578804986.110
Avg. within centroid distance_cluster_0: 49509992721.117
Avg. within centroid distance_cluster_1: 2963910878.467
Avg. within centroid distance_cluster_2: 74492299574.035
Davies Bouldin: 0.318
    
```

Gambar 4.7 Hasil Kinerja

Gambar diatas menunjukkan hasil kinerja dari pengelompokkan data ekspor buah-buahan berdasarkan negara tujuan dengan nilai Davies Bouldin Index

sebesar 0.318 dan Avg. Within centroid distance sebesar 26578804986.110.

Tabel 4.4. Perbandingan Nilai Davies Bouldin Index (DBI)

K	2	3	4	5
DBI	0.219	0.318	0.113	0.034

Tabel diatas menunjukkan perbandingan nilai Davies Bouldin Index berdasarkan jumlah *cluster* yang ditentukan. Dapat dilihat pada nilai Davies Bouldin Index dengan jumlah cluster 3, nilai yang dihasilkan bukan merupakan yang terbaik. Jumlah *cluster* terbaik berada pada jumlah K sebanyak 5 dengan nilai Davies Bouldin Index sebesar 0.034.

5. KESIMPULAN

Berdasarkan hasil penelitian dan pengamatan yang telah dilakukan Metode Algoritma K-Means Clustering dirasa dapat diterapkan dalam melakukan penilaian terhadap hasil ekspor buah-buahan berdasarkan negara tujuan. Dengan menginput data yang digunakan dalam penelitian ini yaitu data Ekspor Buah-buahan Menurut Negara Tujuan Utama dari tahun 2012-2019 yang diambil dari dokumen-dokumen keterangan ekspor impor yang dihasilkan oleh Direktorat Jenderal Bea dan Cukai.

Penerapan Algoritma K-Means menghasilkan nilai titik tengah atau centroid dari data yang didapat dengan ketentuan bahwa *clusterisasi* yang diinginkan adalah 3, penentuan *cluster* dibagi atas tiga bagian yakni *cluster* tingkat ekspor tinggi (C1), *cluster* tingkat ekspor sedang (C2), dan *cluster* tingkat ekspor rendah (C3). Berdasarkan hasil *clusterisasi* yang telah dilakukan didapatkan *Cluster* tingkat ekspor tinggi (C1) yakni Tiongkok & Malaysia, *cluster* tingkat ekspor sedang (C2) yakni Vietnam & Thailand, *cluster* tingkat ekspor rendah (C3) yakni Hongkong, Singapura, Nigeria, India, Jepang, Uni Emirat Arab.

UCAPAN TERIMA KASIH

Terima kasih kepada program studi Teknik Informatika, Universitas Singaperbangsa Karawang atas dukungan dalam terlaksanakannya penelitian ini.

Daftar Pustaka

- [1] Y Windarto, A. P. (2017). Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method. *Techno. Com*, 16(4), 348-357.
- [2] Anjelita, M., Windarto, A. P., Wanto, A., & Saifullah, S. (2019, August). Analisis Metode K-Means pada Kasus Ekspor Barang Perhiasan dan Barang Berharga Berdasarkan Negara Tujuan.

In Seminar Nasional Sains dan Teknologi Informasi (SENSASI) (Vol. 2, No. 1).

- [3] Hajar, S., Novany, A. A., Windarto, A. P., Wanto, A., & Irawan, E. (2020, February). Penerapan K-Means Clustering Pada Ekspor Minyak Kelapa Sawit Menurut Negara Tujuan. In *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)* (Vol. 1, No. 1, pp. 314-318).
- [4] Siyamto, Y. (2017). Pemanfaatan Data Mining Dengan Metode Clustering Untuk Evaluasi Biaya Dokumen Ekspor di PT Winstar Batam. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 1(2).
- [5] Richel, Y. (2020). *PENERAPAN DATA MINING MENGGUNAKAN METODE K-MEANS CLUSTERING PADA DATA EKSPOR MINYAK PALA* (Doctoral dissertation, Universitas Andalas).
- [6] Lubis, R. H. (2018). ANALISIS KINERJA EKSPOR-IMPOR BUAH-BUAHAN INDONESIA PADA PERDAGANGAN INTERNASIONAL. *Al-Masharif: Jurnal Ilmu Ekonomi dan Keislaman*, 6(1), 103-116.
- [7] North, M. (2012). *Data mining for the masses* (Vol. 615684378). Athens: Global Text Project.
- [8] Hofmann, M., & Klinkenberg, R. (Eds.). (2016). *RapidMiner: Data mining use cases and business analytics applications*. CRC Press.
- [9] Darmawan, A., Kustian, N., Rahayu, W., & Tabebuya. (2018). IMPLEMENTASI DATA MINING MENGGUNAKAN MODEL SVM, 2(3), 299-307.