

Clustering Data Ekspor Buah-Buahan Berdasarkan Negara Tujuan Menggunakan Algoritma K-Means

Haviz Atma Negara ^{1,*}, Achmad Rizaldi Putra ¹, Ultach Enri ¹

^{1,1} Teknik Informatika; Universitas Singaperbangsa Karawang; Jl. HS. Ronggo Waluyo, Teluk Jambe Timur, Karawang, Jawa Barat, Indonesia – 41361; telp/fax: 0812-95094228; e-mail: haviz.atma17109@student.unsika.ac.id, achmad.rizaldi17033@student.unsika.ac.id, ultach@staff.unsika.ac.id

* Korespondensi: e-mail: haviz.atma17109@student.unsika.ac.id

Diterima: 27 Mei 2021; Review: 16 Juni 2021; Disetujui: 24 Juni 2021

Cara sitasi: Negara HA, Putra AR, Enri U. 2021. Clustering Data Ekspor Buah-Buahan Berdasarkan Negara Tujuan Menggunakan Algoritma K-Means. Bina Insani ICT Journal. Vol. 8 (1): 73-82.

Abstrak: Ekspor merupakan kegiatan ekonomi dalam memasarkan dan jual barang, baik industry, pangan, serta kebutuhan lainnya kepada negara lainnya yang memiliki *kurs* atau nilai mata uang asing yang lebih besar, tujuannya ialah untuk mencari keuntungan yang sebesar-besarnya. Penelitian ini bertujuan untuk menerapkan *data mining* dengan metode *k-means clustering* data ekspor buah-buahan menurut negara tujuannya yang merupakan salah satu komoditas pangan. Penelitian ini menggunakan data pada tahun 2012 sampai 2019 yang diambil melalui situs <https://www.bps.go.id>. Data diolah dengan mengklusterkan data ekspor kedalam 3 kelompok yaitu tinggi, sedang dan rendah. Didapatkan hasil *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 hasil *cluster* negara tujuan ekspor yaitu 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. Informasi pengklusteran data ekspor buah-buahan ini dapat menjadi saran serta masukan bagi pemerintah maupun perusahaan-perusahaan swasta maupun negeri dalam menentukan strategi ekspor buah-buahan dimasa depan.

Kata kunci: buah-buahan, *data mining*, ekspor, *k-means*, *rapid miner*.

Abstract: Export is an economic activity in marketing and selling goods, both industrial, food, and other needs to other countries that have a higher exchange rate or foreign currency value, the aim is to seek the maximum profit. This study discusses the application of data mining using the *k-means clustering* method on fruit export data based on destination countries. This study uses export data from one of Indonesia's food commodities, namely fruits based on the main destination countries in 2012 to 2019 which was taken through the <https://www.bps.go.id> site. The data is processed by clustering the export data into 3 groups, namely high, medium and low. 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 cluster countries with high export levels, namely China & Malaysia, 2 cluster countries with low export levels, namely Vietnam & Thailand, and 6 cluster countries with low export levels, namely Hong Kong, Singapore, Nigeria, India, Japan, United Arab Emirates. This information on clustering fruit export data can be a suggestion and input for the government and private and state companies in determining fruit export strategies in the future.

Keywords: fruits, data mining, exports, *k-means*, *rapid miner*.

1. Pendahuluan

Ekspor merupakan kegiatan ekonomi untuk memasarkan dan jual barang, baik industry, pangan, serta kebutuhan lainnya kepada negara lainnya yang memiliki nilai mata uang lebih besar, tujuannya ialah untuk mencari keuntungan yang sebesar-besarnya. Sebagai negara agraris Indonesia mempunyai sumber daya yang tumpah ruah, oleh karena itu Indonesia harus dapat bersaing dalam bidang ekspor dengan negara kawasan Asia lainnya. Namun faktanya dalam beberapa dekade terakhir Indonesia belum mampu menembus peringkat satu sebagai negara pengekspor yang ada di kawasan Asia. Berdasarkan agility.com/2021Index (2021), Peringkat ekspor negara kawasan Asia-Pasifik: 1) China, 2) India, 3) Indonesia, 5) Malaysia, 8) Vietnam, 11) Thailand, 21) Phillipines.

Sebagai negara dengan urutan ketiga negara berkembang di kawasan Asia-Pasifik pada tahun 2021. Indonesia memiliki peningkatan kecepatan ekspor yang tergolong bagus, pencapaian tersebut masih kalah dengan China dan India yang menempati peringkat satu dan dua. Penggolongan tingkat ekspor tersebut berdasarkan beberapa indikator penilaian seperti kekuatan logistik dan fundamental bisnis. Sebagai negara agraris Indonesia mampu bersaing dengan negara berkembang lainnya bahkan negara maju yang ada di kawasan Asia-Pasifik melalui pemanfaatan sumber daya yang ada pada negeri ini, hal tersebut bisa saja terwujud jika petani dapat lebih diberdayakan dan ditingkatkan kesejahteraannya. Harapannya produktifitas petani akan meningkat sehingga sumber daya yang ada pada negeri ini pun akan meningkat pula. Berdasarkan lokadata.id/artikel (2021), Peringkat ekspor berdasarkan komoditas barang Indonesia pada kuartal I menurut golongan barang: 1) Pada tahun 2017 rerata nilai ekspor per bulan mengalami fluktuasi di periode 2017-2020 sebesar US\$13,6 miliar meliputi: a) Bahan Bakar sebesar 22%, b) Minyak nabati dan hewani sebesar 15%, c) Hasil industri lain sebesar 13%, d) Mesin dan pengangkutan sebesar 13%, d) Hasil industri bahan sebesar 12%, e) Bahan mentah sebesar 9%, f) Bahan kimia sebesar 8%, g) Bahan makanan sebesar 7%, h) Barang dan transaksi khusus sebesar 1%, dan i) Minuman dan tembakau sebesar 1%; 2) Pada tahun 2018 rerata nilai ekspor per bulan mengalami kenaikan fluktuasi di periode 2017-2020 sebesar US\$1,1 miliar meliputi: a) Bahan Bakar sebesar 22%, b) Minyak nabati dan hewani sebesar 15%, c) Hasil industri lain sebesar 13%, d) Mesin dan pengangkutan sebesar 13%, d) Hasil industri bahan sebesar 12%, e) Bahan mentah sebesar 9%, f) Bahan kimia sebesar 8%, g) Bahan makanan sebesar 7%, h) Barang dan transaksi khusus sebesar 1%, dan i) Minuman dan tembakau sebesar 1%; 3) Pada tahun 2019 rerata nilai ekspor per bulan mengalami penurunan fluktuasi di periode 2017-2020 sebesar US\$1,2 miliar meliputi: a) Bahan Bakar sebesar 23%, b) Minyak nabati dan hewani sebesar 16%, c) Hasil industri lain sebesar 13%, d) Mesin dan pengangkutan sebesar 12%, d) Hasil industri bahan sebesar 10%, e) Bahan mentah sebesar 8%, f) Bahan kimia sebesar 8%, g) Bahan makanan sebesar 7%, h) Barang dan transaksi khusus sebesar 2%, dan i) Minuman dan tembakau sebesar 1%; 4) Pada tahun 2020 rerata nilai ekspor per bulan mengalami kenaikan fluktuasi di periode 2017-2020 sebesar US\$1,38 miliar meliputi: a) Bahan Bakar sebesar 19%, b) Minyak nabati dan hewani sebesar 17%, c) Hasil industri lain sebesar 14%, d) Mesin dan pengangkutan sebesar 12%, d) Hasil industri bahan sebesar 11%, e) Bahan mentah sebesar 8%, f) Bahan kimia sebesar 8%, g) Bahan makanan sebesar 7%, h) Barang dan transaksi khusus sebesar 4%, dan i) Minuman dan tembakau sebesar 1%.

Komoditas ekspor Indonesia masih didominasi sektor migas, sedangkan bahan mentah dan bahan makanan masih berada pada tingkatan bawah yakni peringkat enam dan delapan, data tersebut berdasarkan hasil audit perperiode data komoditas ekspor yang ada di Indonesia. Dengan kekayaan alam yang berlimpah serta anugerah tanah yang subur harusnya dapat lebih dimanfaatkan lagi, sehingga dapat mendorong pertumbuhan ekspor melalui sektor pertanian tidak hanya pada sektor migas. Berdasarkan persoalan tersebut strategi dalam mendorong pertumbuhan ekspor barang yang ada di Indonesia, strategi yang dapat dilakukan salah satunya adalah menggolongkan atau mengklasterkan negara-negara tujuan ekspor yang mempunyai tingkat ketergantungan pangan yang tinggi sedang dan rendah terhadap barang ekspor Indonesia. Strategi tersebut dapat menggunakan data terdahulu untuk diolah kemudian hasilnya dapat dijadikan bahan evaluasi kedepan dalam menentukan strategi ekspor.

Data mining merupakan proses pengolahan data dengan menggali informasi mentah yang diambil dari suatu data untuk kemudian diolah dan mendapatkan informasi penting serta bermanfaat [1]. Terdapat banyak metode dalam melakukan pengolahan data dengan teknik data *mining* salah satunya adalah *k-mean clustering*. Metode ini bekerja dengan membagi data

dengan spesifik yang sama pada beberapa *cluster*, sedangkan data dengan ciri-ciri yang berbeda masuk kedalam kelompok lainnya [2].

Hajar, dkk (2020) dalam penelitiannya mengelompokkan data ekspor menurut negara tujuannya pada minyak kelapa sawit dengan algoritma *kmeans*. Pada penelitian tersebut menghasilkan *centroid* data untuk *cluster* tingkat ekspor tinggi sebesar 5290,90 (India) dan *centroid* data untuk *cluster* tingkat ekspor rendah sebesar 12,20 (Tiongkok, Singapura, Malaysia, Pakistan, Banglades, Sri Lanka, Mesir, Belanda, dan Jerman) [3]. Windarto, A. P. (2017) menggunakan metode *k-mean clustering* dengan bantuan aplikasi *Rapid Miner* untuk mengelompokkan data ekspor salah satu komoditas pangan yaitu buah-buahan berdasarkan negara tujuan tahun pada 2002-2015 [4].

Dari penelitian sebelumnya diketahui bahwa algoritma *k-means* bisa dipakai untuk mengelompokkan data ekspor berdasarkan negara tujuan. Sehingga pada penelitian ini juga akan menggunakan metode *k-mean clustering*. Perbedaan pada penelitian sebelumnya terletak pada banyak *cluster* yang digunakan. Penelitian sebelumnya mengklasterkan data ekspor kedalam 2 kelompok saja yaitu rendah dan tinggi, sedangkan penelitian yang dilakukan sekarang mengklasterkan data ekspor kedalam 3 kelompok yaitu rendah, sedang dan tinggi. Data yang digunakan pun berbeda yakni data ekspor buah-buahan menurut negara tujuannya pada tahun 2012 sampai 2019, data tersebut dipilih untuk diolah kemudian hasilnya dapat dijadikan bahan evaluasi pemerintah kedepan dalam menentukan strategi ekspor pada komoditas pangan. Perbedaan lainnya yaitu penelitian ini tidak hanya menggunakan aplikasi *Rapid Miner* dalam pengolahan data, tetapi juga menggunakan aplikasi *Microsoft Excel* untuk dilakukan perbandingan terhadap hasil pengolahan data.

Data mining ialah analisis data untuk menemukan pengetahuan atau informasi dari sejumlah data dengan kapasitas besar yang disimpan di dalam repositori. Saat ini, ada berbagai istilah yang digunakan untuk menggambarkan proses ini, termasuk analitik, analitik prediktif, data besar, pembelajaran mesin, dan penemuan pengetahuan dalam database. Tetapi istilah-istilah ini semua memiliki kesamaan tujuan untuk menggali informasi yang dapat ditindaklanjuti dari kumpulan data besar [5].

Pola tersebut tidak bisa didapatkan dengan manual harus melewati proses data mining menggunakan metode dan teknik tertentu. Proses yang dilakukan pada data mining didasarkan dari *artificial intelligent*, *machine learning*, dan *statistic* untuk mengidentifikasi pola yang tersembunyi sehingga mendapatkan informasi yang bermanfaat [6].

Sumber data untuk *mining* dapat berupa *databases*, *data warehouses*, atau *repository* informasi lainnya. Ada beberapa teknik pada data *mining* yaitu *classification*, *prediction*, *clustering*, *association*, *description*, dan *estimation* [7].

Metode ini banyak dipakai adalah mengelompokkan atau melakukan clustering pada suatu data. *Clustering* artinya membagi data yang tidak memiliki label menjadi beberapa kelompok [8]. Metode *k-means clustering* bekerja dengan membagi data dengan spesifikasi yang sama pada satu *cluster*/kelompok, sedangkan data dengan spesifikasi berbeda dimasukkan pada kelompok lain. Sehingga metode ini cocok untuk mengklasterkan tingkat ekspor buah-buahan. *Variabel* yang digunakan dalam penelitian ini yaitu jumlah ekspor berat bersih (*Netto*) dan nilai *Free on Board* (FOB). FOB adalah penyerahan barang yang dilakukan diatas kapal yang selanjutnya akan melakukan pengangkatan barang. Teknik yang diadopsi *data mining*, antara lain: *machine learning*, *pattern recognition*, *vizualization*, *algorithm*, *high-performance computing*, *application*, *information retrieval*, *data warehouse*, *database systems*, dan *statistics* (Jiawei Han dkk., 2021). Data ekspor buah-buahan pada tahun 2012-2019 menurut negara tujuan utama akan diolah dengan menggunakan *clustering* yang dibagi menjadi 3 kelompok yaitu tinggi, sedang dan rendah. Terdapat beberapa tahapan dalam pengelompokan data menggunakan algoritma *K-Means* yaitu: 1) Menentukan banyaknya pembagian kelompok, 2) Secara acak menentukan kelompok yang akan menerima data, 3) Data dari masing-masing kelompok dihitung centroid atau nilai rata-ratanya. *Centroid* dari tiap kelompok yang ada diperoleh dari rata-rata seluruh data disetiap fiturnya. Jika M merupakan jumlah total data, dan i merupakan fitur ke-i pada sebuah kelompok, dan p merupakan dimensi untuk data, maka persamaan menghitung centroid fitur ke-i adalah. Persamaan 1 dilakukan sebanyak p dimensi dari i=1 sampai dengan i=p, menggunakan rumus 1:

$$C_i = \frac{1}{M} \sum_{j=1}^M x_j \dots\dots\dots (1)$$

4) Menentukan data kedalam *centroid* atau rata-rata terdekat. *Euclidean* adalah salah satu cara untuk mengukur jarak data menuju pusat kelompok. Untuk menentukan *Distance space Euclidean* dapat dihitung menggunakan rumus 2:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \dots \dots \dots (2)$$

Metode *K-Means* membandingkan jarak antara data dengan *centroid* pada tiap kelompok yang ada merupakan penentuan kembali data ke dalam masing-masing kelompok. Kemudian data dipindahkan lagi ke dalam kelompok yang memiliki jarak *centroid* terdekat secara tegas. Pengalokasian atau pemindahan data ini menggunakan rumus 3:

$$a_{il} = \begin{cases} 1 \\ 0 \end{cases} \quad d = \min\{D(x_i, c_l)\} \dots \dots \dots (3)$$

Perhitungan menurut jarak serta nilai dari keanggotaan dalam kelompok merupakan fungsi objektif dalam penggunaan algoritma *K-Means*. Perhitungan tersebut menggunakan rumus 4:

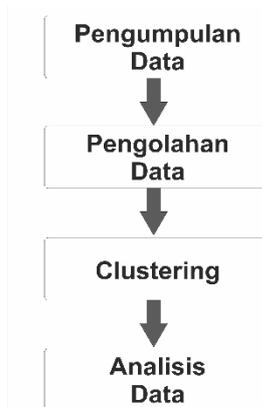
$$J = \sum_{i=1}^n \sum_{l=1}^k a_{il} D(x_i, c_l)^2 \dots \dots \dots (4)$$

n ialah jumlah data, sedangkan k ialah jumlah kelompok, dan a_{il} ialah nilai keanggotaan dari titik data x_i ke kelompok c_l yang diikuti. Untuk a memiliki nilai 0 atau 1. Namun apabila data merupakan anggota dari suatu kelompok maka nilai $a_{il} = 1$. Dan jika tidak maka nilai $a_{il} = 0$. Jika tidak, nilai $a_{il} = 0$.

5) Jika terdapat data yang ternyata berpindah kelompok dan terdapat perubahan dari nilai *centroid* yang telah ditentukan, atau terjadinya perubahan dari nilai fungsi objektif yang digunakan diatas nilai ambang yang ditentukan, maka kembali lagi kepada langkah tiga.

2. Metode Penelitian

Penelitian ini dilakukan dengan metode *k-means clustering* memakai menggunakan aplikasi *Rapid Miner* dan *Microsoft Excel* untuk dilakukan perbandingan terhadap hasil pengolahan data, data kemudian akan diklasterkan kedalam 3 kelompok yaitu rendah, sedang dan tinggi. Beberapa tahap penelitian yang akan dilakukan bisa dilihat melalui Gambar 1 di bawah ini.



Sumber: Siyanto Y. (2017)

Gambar 1. Tahapan *K-mean Clustering*

Tahapan yang ada punya fungsi masing-masing dan tujuan yang berbeda tentunya, berikut merupakan penjelasan tahapan yang dilakukan metode *k-means clustering* dalam mengelompokkan data ekspor buah-buahan:

Pengumpulan Data

Pengumpulan data dimaksudkan untuk menentukan data yang akan diolah untuk dijadikan sebuah informasi yang bermanfaat, penelitian ini menggunakan data ekspor sektor pangan yaitu buah-buahan berdasarkan negara tujuan utama pada tahun 2012 sampai 2019 yang diambil melalui situs <https://www.bps.go.id>. Negara tujuan utama tersebut terdapat 10 negara didalamnya yaitu, Malaysia, Vietnam, Tiongkok, Thailand, India, Jepang, Hongkong, Uni Emirat Arab, Singapura dan Nigeria. Negara-negara tersebut dipilih karena merupakan langganan yang paling banyak permintaan ekspor bidang pangan terutama terhadap sumber daya pangan berupa buah-buahan yang hidup subur ditanah Indonesia, variabel yang

digunakan adalah jumlah ekspor berat bersih (*netto*) dan nilai *Free on Board* (FOB). Data dianalisis dengan mengklasterkan ekspor buah-buahan menurut negara tujuan utama menjadi tiga *cluster* yakni *cluster* tingkat ekspor tinggi, sedang dan rendah. Jumlah data yang didapatkan pada penelitian ini sebanyak sepuluh data untuk variabel Netto dan 10 data untuk variabel FOB. Sepuluh data yang sudah dikumpulkan terdiri dari tahun 2012, 2013, 2014, 2015, 2016, 2017, 2018, dan 2019. Pada tabel 1 dibawah ini merupakan data yang didapatkan berdasarkan variabel Netto.

Tabel 1. Data Variabel Netto

Negara tujuan	2012	2013	2014	2015	2016	2017	2018	2019	TOTAL
Berat bersih/Netweight: 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	15 389,4	18 926,3	72 197,0	80 622,7	112 885,6	384 566,2	165 429,3	76 551,2	926 567,7
India	49 832,7	20 842,6	11 642,8	24 544,6	10 130,7	13 123,8	8 822,9	22 225,0	161 165,1
Jepang	74,2	71,4	71,2	3 240,7	7 377,4	4 919,7	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	2,7	2,6	20,2

Sumber: Hasil Penelitian (2021)

Sedangkan pada tabel 2 dibawah ini merupakan data yang didapatkan berdasarkan variabel FOB.

Tabel 2. Data Variabel FOB

Negara tujuan	2012	2013	2014	2015	2016	2017	2018	2019	TOTAL
Nilai FOB/FOB value : 000 US \$									
Malaysia	18 107,2	25 295,2	25 983,5	31 855,1	65 807,0	41 239,6	71 619,6	62 980,3	342 887,5
Vietnam	23 594,8	33 593,7	57 926,4	83 165,4	91 234,5	87 133,3	73 012,7	87 420,8	537 081,6
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	353 858,6
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	128,2	133,7	183,5	2 119,3	5 078,6	2 636,0	2 730,2	2 828,4	15 837,9
Hongkong	7 470,3	1 823,1	1 544,6	4 253,0	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	6 881,0	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

Sumber: Hasil Penelitian (2021)

Pengolahan Data

Tahapan pengolahan data dilakukan dengan menggunakan *Microsoft Excel*. Tahapan ini dilakukan menggunakan algoritma *K-Means* untuk mendapatkan data kelompok buah-buahan berdasarkan negara tujuannya. Data yang sudah didapatkan akan diolah untuk dapat diklaster. Dari data yang diperoleh kemudian dijumlahkan pada setiap aspeknya, kemudian pada tahapan ini sudah didapatkan perhitungan dari nilai yang ada nantinya akan digunakan pada proses *clustering*. Untuk menentukan *cluster* tiap-tiap negara berdasarkan nilai ekspornya dilakukan tahapan sebagai berikut: mulai lalu menentukan jumlah kelompok dan memindahkan data ke dalam kluster dan menghilangkan *centroid* dari data pada masing-masing kluster dan dilanjutkan dengan memindahkan masing-masing data ke dalam *centroid* yang mana akan ada

pilih untuk mengubah nilai *centroid*? jika Ya, maka akan masuk pada proses menghilangkan centroid dari data pada masing-masing kluster dan jika Tidak maka akan selesai [8].

Clustering

Setelah data diolah menggunakan *Microsoft Excel* maka selanjutnya dilakukan juga pengolahan data dengan *RapidMiner Studio* untuk dilakukan perbandingan. *Cluster* adalah klasifikasi tanpa adanya pengawasan yang membagi seluruh data pada beberapa kelas. Hal ini dikerjakan berdasarkan penggunaan persamaan serta juga langkah-langkah *Euclidean Distance*. Metode analisis digunakan untuk membagi serangkaian data menjadi beberapa grup berdasarkan kesamaan yang telah ditentukan sebelumnya. Dibutuhkan sebuah *flowchart* pada data yang telah tersedia untuk menentukan cluster. *Flowchart* dapat berguna untuk menentukan alur perhitungan dalam menemukan hasil dari penerapan *cluster* pada data yang akan diproses.

Analisis Data

Tahapan ini dilakukan analisis dengan bantuan *tools RapidMiner* untuk menganalisis data ekspor buah-buahan berdasarkan negara tujuannya. Dengan bantuan aplikasi *RapidMiner* beberapa proses pengolahan data bisa dilakukan seperti *data mining*, *machine learning*, *predictive analytics* dan *text mining* [8]. Data yang diperoleh dihitung dengan menghitung bobot dari tiap indeks. Setelah proses sebelumnya ditentukan kedalam tiga *cluster*, maka selanjutnya analisis dari hasilnya.

3. Hasil dan Pembahasan

Pengumpulan Data

Data ekspor buah-buahan pada tahun 2012-2019 yang sudah didapatkan akan dihitung terlebih dahulu untuk melakukan *clustering*. Hasil penjumlahan dari dua kriteria penilaian yaitu berat bersih (*netto*) dan nilai (*FOB*) seperti yang ditunjukkan pada gambar 2.

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	15 389.4	18 926.3	72 197.0	80 622.7	112 885.6	384 566.2	165 429.3	76 551.2	926 567.7
India	49 832.7	20 842.6	11 642.8	24 544.6	10 130.7	13 123.8	8 822.9	22 225.0	161 165.1
Jepang	74.2	71.4	71.2	3 240.7	7 377.4	4 919.7	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	2.7	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	25 983.5	31 855.1	65 807.0	41 239.6	71 619.6	62 980.3	342 887.5
Vietnam	23 594.8	33 593.7	57 926.4	83 165.4	91 234.5	87 133.3	73 012.7	87 420.8	537 081.6
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	353 858.6
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	128.2	133.7	183.5	2 119.3	5 078.6	2 636.0	2 730.2	2 828.4	15 837.9
Hongkong	7 470.3	1 823.1	1 544.6	4 253.0	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	6 881.0	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	173 746.4	238 276.8	274 591.1	346 854.0	269 852.1	290 356.5	1835 644.6

Sumber: Hasil Penelitian (2021)

Gambar 2. Data ekspor buah-buahan tahun 2012-2019

Pengolahan Data

Selanjutnya, langkah pertama dalam proses pengolahan data yaitu mengakumulasikan nilai yang diambil dari data ekspor buah-buahan berdasarkan negara tujuan: 1) Hongkong dengan netto 44650,9 ton dan nilai FOB 43464,7 ribu US\$, 2) Tiongkok dengan netto 1282626,8 ton dan nilai FOB 353858,6 ribu US\$, 3) Singapura dengan netto 53076 ton dan nilai FOB 39897,2 ribu US\$, 4) Malaysia dengan netto 2054519,3 ton dan nilai FOB 342887,5 ribu US\$, 5) Nigeria dengan netto 20,2 ton dan nilai FOB 4650,6 ribu US\$, 6) Vietnam dengan netto 371138,4 ton dan nilai FOB 531081,6 ribu US\$, 7) India dengan netto 16165,1 ton dan nilai FOB 224260,9 ribu US\$, 8) Thailand dengan netto 926567,7 ton dan nilai FOB 241145,4 ribu

US\$, 9) Jepang dengan netto 25203,6 ton dan nilai FOB 15837,9 ribu US\$, dan 10) Uni Emirat Arab dengan netto 46924,6 ton dan nilai FOB 32560,1 ribu US\$.

Input Data

Dataset sebelumnya yang sudah diakumulasikan akan dimasukkan kedalam *tools* RapidMiner, sehingga akan diklasterkan menjadi dua.

Centroid Data

Penggunaan algoritma *k-means* menghasilkan *centroid* dengan jumlah *cluster* 3 sesuai dengan kebutuhan yaitu dari *cluster* tingkat ekspor tinggi (C1), *cluster* tingkat ekspor sedang (C2), dan *cluster* tingkat ekspor rendah (C3). Tiga kelompok tersebut bisa dianggap sebagai tiga titik. Menentukan titik *cluster* pada penelitian ini dilakukan dengan memilih nilai yang paling besar (*maximum*) yang termasuk ke dalam *cluster* tingkat ekspor tinggi (C1), nilai rata-rata (*average*) yang termasuk ke dalam *cluster* tingkat ekspor sedang (C2), dan nilai yang paling kecil (*minimum*) yang termasuk ke dalam *cluster* tingkat ekspor rendah (C3). Pada tabel 4 dibawah ini merupakan nilai titik tengah yang diperoleh pada penelitian ini. Iterasi Pertama pada centroid dijelaskan pada tabel 3.

Tabel 3. Centroid Data Iterasi 1

Atribut	Cluster 1 (Tinggi)		Cluster 2 (Sedang)		Cluster 3 (Rendah)	
	Netto	FOB	Netto	FOB	Netto	FOB
Nilai	2054519.3	537081.6	482090.3	183564.5	20.2	4650.6

Sumber: Hasil Penelitian (2021)

Cluster

Dengan adanya centroid maka selanjutnya proses *clustering* dapat dilakukan dengan penggunaan data yang telah didapat menjadi 3 *cluster* atau kelompok. Proses dimulai dengan mengambil jarak terdekat berdasarkan tiap data yang diolah. Proses pencarian jarak terpendek, pengelompokan data pada iterasi 1 dan *clustering* data dapat digambarkan pada gambar 3. 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				

Sumber: Hasil Penelitian (2021)

Gambar 3. Perhitungan Jarak Pusat Cluster Iterasi 1

Pengelompokan data kedalam *cluster* Data yang mempunyai jarak terdekat akan masuk keadalam *cluster* tersebut. Pengelompokan objek dengan data iterasi 1 pada negara tujuan: 1) Hongkong dengan netto 44650,9 ton dan nilai FOB 43464,7 ribu US\$ serta kelompok 3, 2) Tiongkok dengan netto 1282626,8 ton dan nilai FOB 353858,6 ribu US\$ serta kelompok 1, 3) Singapura dengan netto 53076 ton dan nilai FOB 39897,2 ribu US\$ serta kelompok 3, 4) Malaysia dengan netto 2054519,3 ton dan nilai FOB 342887,5 ribu US\$ serta kelompok 1, 5) Nigeria dengan netto 20,2 ton dan nilai FOB 4650,6 ribu US\$ serta kelompok 3, 6) Vietnam dengan netto 371138,4 ton dan nilai FOB 537081,6 ribu US\$ serta kelompok 2, 7) India dengan netto 16165,1 ton dan nilai FOB 2242690,9 ribu US\$ serta kelompok 3, 8) Thailand dengan

netto 926567,7 ton dan nilai FOB 242245,4 ribu US\$ serta kelompok 2, 9) Jepang dengan netto 25203,6 ton dan nilai FOB 15837,9 ribu US\$ serta kelompok 3, dan 10) Uni Emirat Arab dengan netto 46934,6 ton dan nilai FOB 32560,1 ribu US\$ serta kelompok 3.

Pada iterasi kedua, setelah proses iterasi 1 selesai perlu dilakukan kembali proses iterasi sampai menemukan hasil yang sama dengan proses iterasi 1. Kemudian dihitung kembali untuk mencari *centroid* untuk iterasi 2 yang dapat dilihat pada tabel 4.

Tabel 4. *Centroid* Data Iterasi 2

Atribut	Cluster 1 (Tinggi)		Cluster 2 (Sedang)		Cluster 3 (Rendah)	
	Netto	FOB	Netto	FOB	Netto	FOB
Nilai	1668573	348373,1	648853,1	389113,5	31008,4	60111,9

Sumber: Hasil Penelitian (2021)

Jika nilai *centroid* atau titik tengah sudah ditemukan, proses untuk menentuka jarak terdekat dilakukan kembali. Proses pencarian serta penentuan jarak terpendek, pengelompokan data pada iterasi 2 dapat digambarkan pada gambar 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

Sumber: Hasil Penelitian (2021)

Gambar 4. Perhitungan Jarak Pusat *Cluster* Iterasi 2

Pengelompokan data kedalam *cluster* Data yang mempunyai jarak terdekat akan masuk keadalam *cluster* tersebut dijelaskan pada gambar 5.

Pengelompokan Objek Baru					Iterasi 2
NO	Negara Tujuan	Netto (TON)	Nilai FOB (Ribu US\$)	Kelompok	
1	Hongkong	44650.9	43464.7	3	Cluster 1 sebanyak : 2 Data
2	Tiongkok	1282626.8	353858.6	1	Cluster 2 sebanyak : 2 Data
3	Singapura	53076	39897.2	3	Cluster 3 sebanyak : 6 Data
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	

Sumber: Hasil Penelitian (2021)

Gambar 5. Pengelompokan Data Iterasi 2

Berdasarkan pengamatan yang telah dilakukan karena hasil dari iterasi 1 dan iterasi 2 sama, maka tidak perlu dilakukan proses iterasi kembali.

Analisis data

Pengolahan dataset ekspor buah-buahan berdasarkan negara tujuan telah didapatkan, sebagai perbandingan hasil maka akan dilakukan pengolahan menggunakan *Rapidminer*, adapun hasil dari cluster algoritma K-Means, yaitu:

Cluster Model

Cluster 0: 2 items

Cluster 1: 8 items

Tota number of items: 10

Data tersebut menunjukkan hasil dari 10 data ekspor buah-buahan berdasarkan negara tujuan dengan menerapkan 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 menurut negara tujuannya menggunakan algoritma *K-Means* dengan *tools Microsoft Excel* maupun *rapidminer* menghasilkan pengelompokan 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.

Parameter yang digunakan untuk mengevaluasi performa dari algoritma K-Means yaitu nilai *Davies Bouldin Index* dan *Avg. Within centroid distance*. Berdasarkan *Davies Bouldin Index cluster* dengan jarak *inter-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* adalah rata-rata dari jarak *cluster* dihitung dengan rata-rata jarak antara centroid dan semua contoh *cluster*. Hasil kinerja ditunjukkan dengan:

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

Data diatas menunjukkan hasil kinerja dari pengelompokan data ekspor buah-buahan berdasarkan negara tujuan dengan nilai *Davies Bouldin Index* sebesar 0.318 dan *Avg. Within centroid distance* sebesar 26578804986.110. Tabel 5 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.

Tabel 5. Perbandingan Nilai *Davies Bouldin Index (DBI)*

K	2	3	4	5
DBI	0.219	0.318	0.113	0.034

Sumber: Hasil Penelitian (2021)

4. Kesimpulan

Melalui hasil penelitian serta pengamatan yang telah dilakukan Algoritma *K-Means Clustering* dapat dipakai untuk menentukan hasil ekspor buah-buahan menurut negara tujuannya. Pada penelitian ini diperoleh nilai *cluster* dengan tingkat ekspor tinggi mendapatkan nilai *centroid* sebesar 2054519.3, *cluster* dengan tingkat ekspor sedang mendapatkan nilai *centroid* sebesar 489020.3, dan untuk *cluster* dengan tingkat ekspor rendah mendapatkan nilai *centroid* sebesar 20.2. Dengan menggunakan *tools Microsoft Excel* atau *Rapid Miner* hasil *cluster* yang sama. Berdasarkan hasil pengolahan data didapatkan 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.

Referensi

- [1] A. M. Siregar dan A. Puspabhuana, *Data Mining: Pengolahan Data Menjadi Informasi dengan RapidMiner*, CV Kekata Group, 2017.
- [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] Windarto, A. P. (2017). Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan KMeans Clustering Method. *Techno. Com*, vol. 16, no. 4, pp. 348-357.
- [5] Larose, D. T., & Larose, C. D. (2014). *Discovering Knowledge in Data an Introduction to Data Mining*. New Jersey: John Wiley & Sons, Inc.
- [6] Kamila, I., Khairunnisa, U., & Mustakim. (2019). Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau. *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 5, no. 1, pp. 119-125.
- [7] Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques Third Edition*. Waltham: Elsevier.
- [8] Siyamto, Y. (2017). Pemanfaatan Data Mining Dengan Metode Clustering Untuk Evaluasi Biaya Dokumen Ekspor di PT Winstar Batam. *Jurnal Media Informatika Budidarma*, vol. 1, no. 2.
- [9] Richel, Y. (2020). *Penerapan Data Mining Menggunakan Metode K-Means Clustering Pada Data Ekspor Minyak Pala* (Doctoral dissertation, Universitas Andalas).
- [10] Lubis, R. H. (2018). Analisis Kinerja Ekspor-Impor Buah-Buahan Indonesia Pada Perdagangan Internasional. *Al-Masharif: Jurnal Ilmu Ekonomi dan Keislaman*, vol. 6, no. 1, pp. 103-116.