CLUSTERING PRODUK EKSPOR INDONESIA BERDASARKAN TINGKAT PERMINTAAN MENGGUNAKAN METODE K-MEANS TAHUN 2020-2022

Mujibulloh [1], Martanto [2], Umi Hayati [3]
[1] Teknik Informatika, STMIK IKMI Cirebon
[2] Manajemen Informatika, STMIK IKMI Cirebon
[3] Teknik Informatika, STMIK IKMI Cirebon
Jalan Perjuangan No. 10 B Cirebon, Indonesia
mujibulloh0828@gmail.com

ABSTRAK

Indonesia saat ini merupakan salah satu eksportir ke berbagai negara maju dan berkembang. Meski demikian, terdapat permasalahan dalam mengidentifikasi pola dan prospek ekspor yang dapat membantu perusahaan dalam industri ini untuk membuat keputusan yang lebih baik. Penelitian ini bertujuan untuk menganalisis produk ekspor utama Indonesia dan memberikan prospek ekspor bagi perusahaan yang terlibat dalam industri ini dengan menggunakan teknik analisis K-Means Clustering. Peneliti Menggunakan data dari Badan Pusat Statistik Indonesia, penelitian ini mengambil sampel periode 2020-2022 yang kemudian dianalisis untuk mengidentifikasi cluster berdasarkan volume ekspor. Hasil penelitian menunjukkan bahwa metode K-Means efektif dalam mengklasifikasikan produk ekspor ke dalam tiga cluster: ekspor rendah (C0) dengan 99 data, ekspor dan impor sedang (C1) dengan 3 data, dan ekspor tinggi (C2) dengan 1 data. Dengan nilai DBI terkecil sebesar 0.3665 untuk cluster tiga, Hasil ini memberikan kontribusi dalam pemahaman lebih lanjut tentang pola ekspor dan impor dan dapat digunakan sebagai dasar untuk pengambilan keputusan yang lebih baik di dalam industri ini.

Kata kunci: *Data mining, K-Means clustering, ekspor, KDD.*

1. PENDAHULUAN

Ekspor memegang peran penting dalam mendorong pertumbuhan dan pembangunan ekonomi suatu negara, termasuk Indonesia. Kegiatan ekspor dapat memberikan dampak positif terhadap perekonomian dengan mendorong produsen dalam negeri untuk memanfaatkan sumber daya produksi mereka dengan lebih baik dan menjadi lebih kompetitif di pasar global. Faktor-faktor yang mempengaruhi kegiatan ekspor bisa berasal dari dalam atau luar negeri. Dalam era digital, proses ekspor dan impor menjadi lebih mudah dengan adanya akses informasi mengenai buyer dan supplier. Namun, pemanfaatan data historis melalui metode data mining dalam konteks ekspor barang masih belum banyak dilakukan secara komprehensif.

Kegiatan ekspor sangat penting bagi suatu negara. Jika nilai impor suatu negara lebih besar dari nilai ekspornya, ini bisa berdampak negatif pada perekonomian negara tersebut. Pada era digital dan perkembangan teknologi informasi, penggunaan sistem informasi yang terintegrasi dan pemanfaatan data mining bisa memberikan keuntungan besar bagi perusahaan yang terlibat dalam ekspor barang. Data mining adalah metode yang bisa menggali informasi berharga dari data historis, mengidentifikasi pola dan hubungan yang tidak terlihat secara langsung, dan memberikan wawasan yang mendalam untuk pengambilan keputusan yang lebih baik. Namun, dalam konteks ekspor barang, penerapan metode data mining masih belum banyak dilakukan secara komprehensif. Banyak perusahaan ekspor yang masih mengandalkan pengalaman dan pengetahuan intuitif dalam mengelola proses ekspor, tanpa memanfaatkan potensi data yang dimiliki.

Untuk mengatasi permasalahan ini, penelitian ini bertujuan untuk menganalisis barang ekspor Indonesia yang banyak diminati oleh pasar internasional dengan menggunakan metode data mining. Metode ini menekankan pada data histori ekspor Indonesia untuk mengkelompokkan nilai barang yang paling banyak diminati melalui variabel berat dan harga barang ekspor. Teknik analisis yang digunakan dalam penelitian ini adalah data mining k-means clustering. Tujuan dari penelitian ini adalah untuk menentukan produk ekspor yang paling diminati di pasar internasional dan yang kurang bersaing dalam pasar internasional. Hasil dari penelitian ini dapat memberikan wawasan berharga yang dapat digunakan untuk efisiensi dan efektivitas ekspor pada eksportir pemula untuk memilih komoditas yang sudah mampu bersaing di pasar internasional.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian sebelumnya dilakukan oleh Windarto, (2017), membahas mengenai penerapan data mining pada ekspor buah buahan menggunakan clustering dari berbagai buah buahan yang di ekspor oleh Indonesia, metode dalam penelitian ini menggunakan K-Means Clustering agar menemukan buah buahan yang diminati pasar internasional. Tujuan dari penelitian ini adalah menerapkan k-means dalam clu ster ekspor buah buahan berdasarkan negara tujuan. Hal ini dapat menjadi masukan kepada pemerintah dan organisasi yang berkontribusi kepada ekspor buah buahan untuk merencanakan strategi ekspor yang lebih luas [2].

Penelitian selanjutnya dilakukan oleh Zatira et al., (2021), membahas mengenai bagaimana pengaruh perdagangan internasional mempengaruhi kenaikan prekonomian Indonesia. Penelitian ini bertujuan untuk menguji pengarus perdagangan internasional yang di ukur dengan nilai ekspor dan impor terhadap pertumbuhan ekonomi Indonesia yang diukur dengan Gross Domestic Produk (GDP). Penelitian ini menggunakan model regresi kuantitatif untuk menganalisis data. Pengaruh perdagangan internasional terhadap pertumbuhan ekonomi Indonesia selama tahun penelitian adalah subjek penelitian ini [3].

Penelitian selanjutnya dilakukan oleh .Pooja et al., (2022) membahas penerapan data mining clustering pada nilai ekspor kertas berdasarkan Pelabuhan asal Indonesia menggunakan algoritma K-Means. Studi ini bertujuan untuk mencapai informasi kluster Pelabuhan Pelabuhan yang berada di kategori menengah. Hasil dari penelitian ini bisa menjadi informasi penting terhadap pemerintahan dan perusahaan di bidang logistic mengenai pemetaan pelabuan asal dengan valuasi rendah hingga tinggi untuk kegiatan ekspor kertas di Indonesia [5].

2.2. Data Mining

Data mining adalah proses yang melibatkan pengekstrakan informasi berharga dan tersembunyi dari dataset besar. Proses ini menggunakan berbagai teknik dan algoritma analisis data untuk mengidentifikasi pola dan tren yang dapat membantu dalam pengambilan keputusan. Data mining mencakup beberapa tahap, termasuk pra-pemrosesan data, eksplorasi data, pemodelan, evaluasi, dan interpretasi hasil. Tujuan utama dari data mining adalah untuk menghasilkan informasi baru dari tumpukan data yang besar. Data mining juga dikenal sebagai Knowledge Discovery in Database (KDD), yang merupakan proses otomatis mencari data dalam ruang memori yang sangat besar untuk menemukan pola dengan menggunakan teknik seperti klasifikasi, asosiasi, atau clustering. Beberapa teknik yang digunakan dalam data mining antara lain algoritma naïve bayes, decision tree, dan lainnya [6][5]. Data mining dapat digunakan dalam berbagai bidang, oleh karena itu pada penelitian ini digunakan untuk menentukan produk ekspor yang paling diminati di pasar internasional dan yang kurang bersaing dalam pasar internasional menggunakan algoritma yang ada pada data mining.

2.3. Clustering

Clustering adalah teknik pengumpulan data yang digunakan untuk menganalisis dan memecahkan masalah dalam pengelompokan data menjadi partisi yang lebih efisien dari dataset ke subset. Dalam proses ini, tujuannya adalah untuk mendistribusikan kasus (benda, orang, peristiwa, dll.) ke dalam kelompok, sehingga tingkat hubungan antara anggota cluster yang sama kuat dan lemah antara anggota cluster yang berbeda [7][8]. Ada dua jenis metode dalam pengelompokan, yaitu hierarchical clustering dan non-hierarchical clustering. Hierarchical clustering adalah metode pengelompokan data yang bekerja dengan menggabungkan dua atau lebih set data yang memiliki kemiripan atau persamaan. Proses ini dilanjutkan dengan menggabungkan set data berikutnya yang paling dekat, dan berlanjut hingga cluster membentuk semacam pohon di mana ada tingkat atau tingkat yang jelas antara objek yang mirip dengan yang paling mirip. Namun, secara logika, semua data akan membentuk satu cluster pada akhirnya [9][10][11].

2.4. K-Means

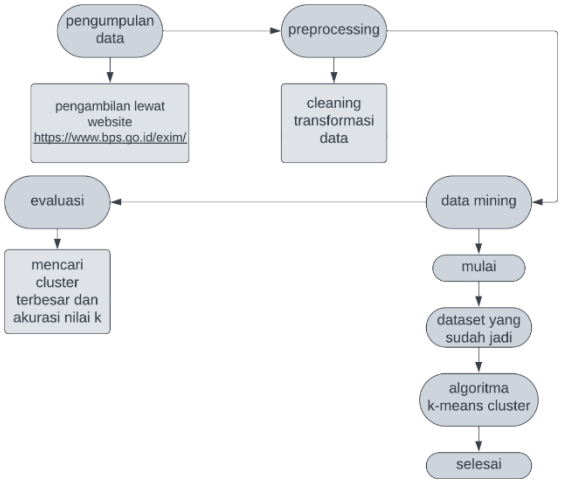
Algoritma K-Means mengelompokkan data berdasarkan centroid, atau pusat klaster, terdekat. Tujuan algoritma ini adalah untuk memaksimalkan kemiripan data dalam satu klaster. dan mengurangi kemiripan data antar klaster. Ukuran kemiripan yang digunakan dalam klaster adalah fungsi jarak, jadi tingkat kemiripan tertinggi dapat dicapai dengan menghitung jarak terkecil antara data dan titik centroid [12].

2.5. Rapid Miner

RapidMiner adalah perangkat lunak yang terlibat dalam pemrosesan data, menggunakan berbagai prinsip dan algoritma yang terkait dengan penambangan data. Dengan memanfaatkan metodologi ini, RapidMiner secara efektif mengekstrak pola yang mungkin disembunyikan dalam data yang ada. Ini dicapai dengan menggabungkan teknik statistik, metodologi kecerdasan buatan, dan strategi database [13]. Tujuan menyeluruh RapidMiner adalah untuk mempercepat pelaksanaan perhitungan data skala besar, sehingga memungkinkan pengguna untuk melakukan perhitungan kompleks yang berkaitan dengan data besar. Operator, yang berfungsi sebagai komponen fundamental RapidMiner, memainkan peran penting dalam memodifikasi data. Dengan menggabungkan operator yang ada dengan node yang diinginkan, pengguna dapat dengan mulus menavigasi data dan memeriksa hasilnya. Akibatnya, hasil yang diperoleh melalui RapidMiner disajikan secara visual, memungkinkan pengguna untuk dengan mudah memahami implikasi data melalui penggunaan grafik.

3. METODE PENELITIAN

Pada metode penelitian ini akan dijelaskan Langkah-langkah metode Knowledge Discovery in Databases (KDD) untuk mengelompokkan data – data yang memiliki cluster tinggi dan rendah. Adapun Langkah-langkahnya dapat dilihat pada Gambar 1.



Gambar 1. Metode Penelitian

3.1. Pengumpulan Data

Tahap pengumpulan data merupakan langkah awal dalam proses penelitian ini. Pada tahap ini, data historis ekspor produk Indonesia dari tahun 2020 hingga 2022 dikumpulkan.

3.2. *Preprocessing*

Tahap preprocessing merupakan langkah krusial dalam proses data mining yang melibatkan penghilangan data yang tidak relevan atau tidak diperlukan serta transformasi data untuk memastikan bahwa data tersebut dapat diolah oleh algoritma. Dalam konteks algoritma K-means clustering, preprocessing mencakup pembersihan data, seperti menghilangkan noise dan outliers, Hal ini penting karena K-means clustering mengandalkan pengukuran jarak antar titik data, dan perbedaan skala dapat mempengaruhi hasil clustering. Selain itu, tahap ini juga dapat meliputi konversi data kategorikal menjadi numerik, Dengan preprocessing yang efektif, data akan lebih siap untuk dianalisis yang akan meningkatkan kualitas dan akurasi dari hasil clustering.

3.3. Data mining

Penerapan algoritma k-means clustering untuk Menentukan nilai centeroid dan mencari cluster yang terbaik untuk data yang diperoleh dan menentukan cluster tertinggi dan terendah.

3.4. Evaluasi

Tahap ini melibatkan proses pengukuran akurasi dari algoritma K-means clustering. Evaluasi dilakukan dengan mencari nilai Davies-Bouldin Index (DBI) dari hasil analisis data mining K-means clustering. DBI adalah metrik evaluasi yang digunakan untuk mengukur kualitas pengelompokan dalam data mining dan machine learning. Nilai DBI yang lebih rendah menunjukkan bahwa pengelompokan lebih baik, karena menunjukkan bahwa anggota dalam satu cluster lebih mirip satu sama lain dibandingkan dengan anggota di cluster lain. Dengan demikian, melalui evaluasi ini, kita dapat menentukan sejauh mana efektivitas dan akurasi dari metode K-means clustering dalam mengelompokkan data ekspor.

4. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan metode Knowledge Discovery in Databases (KDD) dalam lima tahapan utama, yaitu pengumpulan data, preprocessing, transformasi, data mining, dan evaluasi hasil. Proses ini dilakukan dengan menggunakan bahasa pemrograman Python dan Google Collab.

4.1. Pengumpulan Data

Proses penelitian ini dimulai dengan tahap pengumpulan data, yang merupakan fondasi penting untuk analisis selanjutnya. Data ekspor yang digunakan dalam penelitian ini bersumber dari situs resmi pemerintahan Indonesia, yaitu Badan Pusat Statistik Indonesia, yang dikelola oleh Direktorat Bea Cukai. Data yang dikumpulkan mencakup informasi ekspor dari periode 2020-2022, yang merupakan populasi penelitian. Pencarian dan pengumpulan data ini dilakukan pada tanggal 26 September 2023, memastikan bahwa data yang diperoleh adalah yang terbaru dan relevan untuk penelitian. Informasi ini akan digunakan untuk menganalisis pola ekspor Indonesia dan mengidentifikasi kelompok atau cluster produk berdasarkan volume ekspor melalui metode K-Means clustering.

	Kode HS dan Deskripsi	Nilai Ekspor (US $)	Nilai Impor (US $)
0	Payung, tongkat, dan bagiannya	2.081353e+05	3.431766e+06
1	Alas kaki	1.656885e+09	2.382227e+08
2	Aluminium dan barang daripadanya	1.791579e+08	5.515115e+08
3	Ampas dan sisa industri makanan	5.044465e+08	1.102209e+09
4	Bahan anyaman nabati	1.241890e+08	3.618180e+05

Gambar 2. Data Mentah

4.2. *Preprocessing*

Tahap *preprocessing* data melibatkan serangkaian langkah yang bertujuan untuk mengubah data mentah menjadi data yang dapat dianalisis dan dipahami oleh algoritma. Langkah-langkah ini mencakup penggabungan dataset, konversi data menjadi float, dan perhitungan selisih antara ekspor dan impor. Proses ini membantu dalam mengurangi noise, standarisasi data, dan mengubah data menjadi bentuk yang lebih terstruktur dan bersih dari informasi yang tidak relevan, sehingga memudahkan pemrosesan dan pemahaman lebih lanjut terhadap algoritmanya. Pertama, penggabungan dataset dilakukan untuk mengintegrasikan data dari berbagai sumber menjadi satu set data yang konsisten dan komprehensif. Selanjutnya, data diubah menjadi float, yang merupakan proses penting dalam analisis data. Dalam Python, fungsi float() digunakan untuk mengubah integer menjadi float, dan fungsi int() digunakan untuk mengubah float menjadi integer. Selain itu, perhitungan selisih antara ekspor dan impor dilakukan untuk menghasilkan data yang relevan dalam konteks perdagangan internasional. Selisih ini

dihitung dengan rumus "Ekspor - Impor". Dengan melakukan langkah-langkah preprocessing ini, data dapat diolah menjadi bentuk yang lebih terstruktur, bersih dari informasi yang tidak relevan, dan siap untuk analisis lebih lanjut. Hal ini membantu mengurangi noise dan melakukan standarisasi data sehingga memudahkan pemrosesan dan pemahaman lebih lanjut terhadap algoritmanya.

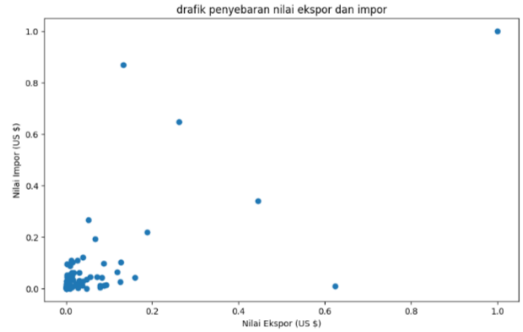
```
Hasil Konversi ke NumPy Array:
[[' Payung, tongkat, dan bagiannya' 208135.28 3431766.0]
 ['Alas kaki' 1656884827.12 238222685.0]
 ['Aluminium dan barang daripadanya' 179157915.41 551511463.0]
 ['Ampas dan sisa industri makanan' 504446477.87 1102208976.0]
 ['Bahan anyaman nabati' 124189039.34 361818.0]
 ['Bahan bakar mineral' 13156661994.990002 9071160390.0]
 ['Bahan kimia anorganik' 400012564.5 549223617.0]
 ['Bahan kimia organik' 890981784.66 1742839627.0]
 ['Bahan peledak, korek api, dan kembang api' 1439982.57 34814447.0]
 ['Barang anyaman' 32995796.42 156467.0]
 ['Barang dari batu, semen, asbes, atau mika' 37740991.65 135216496.0]
 ['Barang dari besi dan baja' 350968455.89 996324284.0]
 ['Barang dari bulu unggas, bunga artifisial, dan wig' 111474807.42000002
  17109495.0]
```

Gambar 3 Hasil Preprocessing

4.3. Data Mining

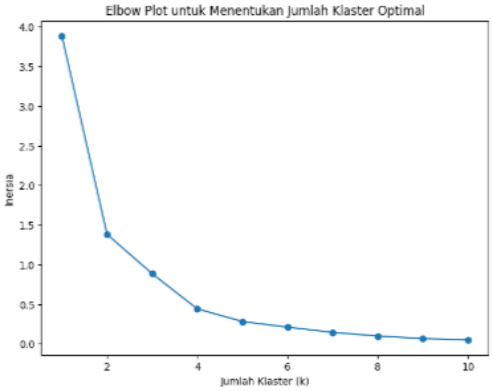
Data mining merupakan proses analisis data yang kompleks, terstruktur, dan tidak terstruktur untuk menemukan pola, atribut, relasi, dan keanekaragaman lain untuk mendapatkan informasi yang diperlukan. Penelitian ini peneliti menggunakan salah satu metode data mining yaitu K-means clustering. K-means clustering adalah metode pengelompokan data kedalam cluster. Metode ini sangat berguna dalam mengelompokkan data dengan atribut kategorikal dan dalam penanganan data yang besar.

Dalam konteks penelitian data ekspor impor, K-means cluster akan mengelompokkan data setiap kelas berdasarkan centroidnya. Setalah dataset selesai di preprocessing dataset siap diolah ke algortim K-means clustering. Dalam proses clustering yang harus di perhatikan adalah nilai centroid dalam setiap cluternya. Agar memudahkan penjelasan menggenai K-means cluster langkah pertama yang harus dicari yaitu nilai K. Sebelum mencari nilai K dilakukan penskalaan fitur menggunakan MinMaxScaler untuk memastikan bahwa sebaran data lebih akurat dan siap untuk proses clustering. Penskalaan ini penting karena K-Means bergantung pada pengukuran jarak antar titik data, dan perbedaan skala dapat mempengaruhi hasil klasterisasi hasilnya dapat dilihat pada Gambar 4.



Gambar 4. Hasil Sebaran Data

Gambar 4 menunjukan sebaran data yang sudah dikelola dari nilai ekspor dan nilai impor sebaran data ini menunjukan data banyak berkumpul pada nilai 0.0. Langkah selanjutnya menentukan jumlah kluster yang tepat untuk memilih jumlah klaster yang tepat yang direpresentasikan menggunakan grafik elbow pada Gambar 5.



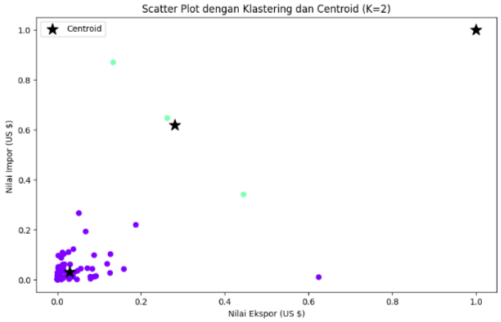
Gambar 5. Grafik Elbow

Dari Gambar 5 dapat ditarik kesimpulan untuk menentukan jumlah klaster yang optimal, digunakan metode Elbow Plot, yang menggambarkan hubungan antara jumlah klaster dan inersia (sum of squared distances to the nearest cluster center). Dari grafik Elbow Plot yang disajikan, terlihat bahwa titik 'siku' terbaik adalah pada jumlah klaster 2, yang menunjukkan bahwa ini adalah jumlah klaster yang paling sesuai untuk data yang dianalisis maka ditetapkan nilai klusternya 2.

```
[[0.02952893 0.03122866]
 [0.28022464 0.6183742 ]
 [1.         1.        ]]
```

Gambar 6. Hasil Pembagian Kluster

Selanjutnya melatih data yang telah diskala untuk mencetak nilai centroid dari pusat masing masing klaster yang nantinya akan memberikan informasi mengenai lokasi tengah pada klaster. visualisasi data dengan scatter plot dengan warna berdasarkan klaster hasil dari algoritma K-means pusat klaster ditandai dengan bintang hitam dapat dilihat pada Gambar 7.



Gambar 7. Hasil Algoritma K-Means

Plot ini membantu memahami bagaimana algoritma K-Means mengelompokkan data dalam ruang fitur dua dimensi dimana terdapat 2 claster claster 0 memiliki jumlah yang paling mendominasi dan klaster 1 terdapat 4 anggota.

Hasil klasterisasi ini kemudian dievaluasi menggunakan Davies Bouldin Index (DBI), yang mengukur rata-rata 'kesamaan' antara klaster dengan klaster yang paling mirip. Nilai DBI yang lebih rendah menunjukkan pembagian klaster yang lebih baik.

4.4. Evaluasi

Evaluasi model machine learning adalah proses mengukur seberapa baik model yang dilatih dapat melakukan prediksi atau klasifikasi pada data baru. Tujuannya adalah untuk menilai performa model sebelum diimplementasikan, sehingga diketahui apakah model sudah cukup baik atau perlu perbaikan. Untuk melihat melihat performa klaster yang bagus penulis menggunakan Davies Bouldin Index (DBI) yang bertujuan untuk melihat jarak dari klaster, semakin kecil jarak antara klaster maka semakin bagus.

Table 1. DBI

Jumlah K	Nilai DBI
K2	0.5452
K3	0.3665
K4	0.3944
K5	0.5966
K6	0.4388
K7	0.5031
K8	0.4551
K9	0.3674
K10	0.3841

Langkah awal yang dilakukan peneliti secara garis besar adalah melakukan persiapan dataset Kemudian peneliti membuatnya menjadi empat model yang dibagi berdasarkan ratio dari pembagian data latih dan data uji untuk mengetahui dari peforma tiap-tiap model peneliti sudah menjelaskan proses implementasinya dari awal sampai akhir,selanjutnya pembahasan mengenai implementasi tersebut. Dari hasil penelitian terdapat 9 percobaan dalam pemilihan jumlah klaster, dari 9 percobaan tersebut akan dilihat BDI terkecilnya karena nilai DBI yang kecil dapat meningkatkan performa dari algoritma K-Means clustering hasil dari percobaan tersebut terdapat pada table 1.

Setelah melakukan performa pengujian pada nilai DBI maka di dapartkan K terbaik dalam cluster dari nilai ekspor dan impor adalah K 3 dengan nilai DBI 0.3665 dengan mengevaluasi perhitungan manual yaitu:

DBI=1/3 (0.105372+0.029528+0.69494) = 0.3665

Jadi Hasil dari cluster terbaik menunjukkan pembagian menjadi tiga cluster utama, yaitu cluster tingkat ekspor rendah (C0), cluster tingkat ekspor dan impor sedang (C1), dan cluster tingkat ekspor tinggi (C2). Dalam rinciannya, terdapat 99 data dalam cluster (C0), 3 data dalam cluster (C1), dan 1 data dalam cluster (C2). Pembagian ini memberikan gambaran tentang variasi tingkat ekspor dan impor pada kelompok data yang diamati, memfasilitasi pemahaman yang lebih mendalam terkait karakteristik masing-masing cluster. Dari perhitungan yang telah dilakukan menjelaskan nilai DBI terbaik terdapat Pada klaster 3 maka penulis menggunakan Klaster 3 dalam melakukan klaterisasi dalam penelitian ini adapun hasil klaster dari penelitan ini akan di paparkan di table 2

Table 2. Hasil Clustering

Kode HS dan Deskripsi	Nilai Ekspor (US $)	Nilai Impor (US $)	kluster
Payung, tongkat, dan bagiannya	208135,28	3431766	0
Alas kaki	1656884827	238222685	0
Aluminium dan barang daripadanya	179157915,4	551511463	0
Ampas dan sisa industri makanan	504446477,9	1102208976	0
Bahan anyaman nabati	124189039,3	361818	0
Bahan bakar mineral	13156661995	9071160390	2
Bahan kimia anorganik	400012564,5	549223617	0
Bahan kimia organik	890981784,7	1742839627	0
Bahan peledak, korek api, dan kembang api	1439982,57	34814447	0
Barang anyaman	32995796,42	156467	0
Barang dari batu, semen, asbes, atau mika	37740991,65	135216496	0
Barang dari besi dan baja	350968455,9	996324284	0
Barang dari bulu unggas, bunga artifisial, dan wig	111474807,4	17109495	0
Barang dari kulit samak	270002412,6	109886999	0
Barang fotografi atau sinematografi	376554,36	14667771	0
Barang tekstil jadi lainnya	42256461,01	91195913	0
Berbagai barang buatan pabrik	66823821,89	128737821	0
Berbagai barang logam tidak mulia	28454140,23	245508441	0
Berbagai makanan olahan	402940299,2	275221742	0
Berbagai produk kimia	1669035363	922417102	0
Besi dan baja	5851282232	3086535233	1
Biji dan buah mengandung minyak	103280750,3	402244466	0

Kode HS dan Deskripsi	Nilai Ekspor (US $)	Nilai Impor (US $)	kluster
Bijih logam, terak, dan abu	2091554188	380973342	0
Binatang hidup	18905042,04	160120779	0
Buah-buahan	224685285,4	551597250	0
Daging hewan	4996010,14	252790334	0

Tabel 2 menunjukkan hasil klasterisasi dari penelitian ini. Setiap baris dalam tabel mewakili jenis produk ekspor dengan nilai ekspor dan impor yang sesuai dalam dolar AS, serta klaster yang ditugaskan oleh model K-Means. Klaster ditandai dengan angka 0, 1, dan 2. Klaster 0 mencakup produk seperti "Payung, tongkat, dan bagiannya", "Alas kaki", "Aluminium dan barang daripadanya", dan lainnya. Klaster ini tampaknya mencakup produk dengan nilai ekspor dan impor yang relatif rendah atau moderat. Klaster 1 mencakup produk seperti "Besi dan baja", yang memiliki nilai ekspor dan impor yang cukup tinggi dibandingkan dengan produk lain dalam tabel. Klaster 2 produk seperti "Bahan bakar mineral", yang memiliki nilai ekspor dan impor yang sangat tinggi dibandingkan dengan produk lain dalam tabel.

Dengan demikian, tabel ini memberikan gambaran tentang bagaimana produk ekspor Indonesia dikelompokkan berdasarkan nilai ekspor dan impor mereka. Produk dengan nilai ekspor dan impor yang serupa dikelompokkan dalam klaster yang sama.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil menerapkan metode data mining K-Means clustering untuk menganalisis pola ekspor Indonesia selama periode 2020 hingga 2022. Hasil analisis menunjukkan pembentukan tiga cluster yang mencerminkan tingkat ekspor barang: rendah (C0) dengan 99 data, sedang (C1) dengan 3 data, dan tinggi (C2) dengan 1 data. Dengan nilai DBI terkecil sebesar 0.3665 untuk cluster tiga, penelitian ini memberikan pemahaman yang lebih mendalam tentang pola ekspor dan impor, serta dapat dijadikan dasar untuk pengambilan keputusan yang lebih tepat dalam industri ekspor.

Penelitian ini juga menunjukkan bahwa pemanfaatan data mining dalam konteks ekspor barang dapat membantu perusahaan dalam mengidentifikasi produk ekspor yang paling diminati di pasar internasional dan yang kurang bersaing. Untuk penelitian selanjutnya, disarankan untuk memperluas cakupan analisis terhadap faktor-faktor lain yang memengaruhi ekspor, seperti harga, tujuan ekspor, atau karakteristik produk, guna mendapatkan wawasan yang lebih komprehensif dalam mendukung pengambilan keputusan di industri ekspor. Selain itu, penelitian juga dapat mempertimbangkan penggunaan metode data mining lainnya untuk perbandingan dan validasi hasil. Dalam konteks aplikasi praktis, hasil analisis ini dapat diaplikasikan dalam strategi bisnis dan kebijakan pemerintah yang relevan, seperti pengembangan produk ekspor yang lebih kompetitif, peningkatan efisiensi dan efektivitas ekspor, serta perencanaan strategi ekspor yang lebih luas.

DAFTAR PUSTAKA

[1] A. P. Windarto, "Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method," *Techno.Com*, vol. 16, no. 4, pp. 348–357, 2017, doi: 10.33633/tc.v16i4.1447.

[2] A. Perdana Windarto Program Studi Sistem Informasi and S. Tunas Bangsa Jln Jendral Sudirman Blok, "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Application of Data mining on Fruit Exports by Destination Country Using K-Means Clustering," *IJCCS ISSNTechno.COM*, vol. 16, no. 4, pp. 348–35778, 2017, [Online]. Available: https://www.bps.go.id.

[3] D. Zatira, T. N. Sari, and M. D. Apriani, "Perdagangan Internasional Terhadap Pertumbuhan Ekonomi Indonesia," *J. Ekon.*, vol. 11, no. 1, p. 88, 2021, doi: 10.35448/jequ.v11i1.11277.

[4] N. Pooja, M. Saputra, S. Aisyah, and P. Juanta, "Implementasi Data Mining Clustering Data Valuasi Ekspor Kertas Indonesia Menggunakan Algoritma K-Means," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 5, no. 2, pp. 86–90, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2372.

[5] H. Prastiwi, Jeny Pricilia, and Errissya Rasywir, "Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering," *J. Inform. Dan Rekayasa Komputer(JAKAKOM)*, vol. 2, no. 1, pp. 141–148, 2022, doi: 10.33998/jakakom.2022.2.1.34.

[6] H. O. L. Wijaya, C. Yuliansyah, and Armanto, "IMPLEMENTASI ASOSIASI RULE MINING PADA DATA TRANSAKSI PENJUALAN MENGGUNAKAN ALGORITMA APRIORI," *Technologia*, vol. 13, no. 1, pp. 30–35, 2022.

[7] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.

[8] Yulia and M. Silalahi, "Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means," *Indones. J. Comput. Sci.*, vol. 10, no. 1, 2021, doi: 10.33022/ijcs.v10i1.3008.

[9] E. Juliana, V. N. Aleyda, and Y. Yuliana, "Penerapan Metode Clustering K-Means Untuk

Membantu Menentukan Tingkatan Status Daerah Dampak Covid-19," *J. Mediat.*, vol. 4, no. 3, p. 112, 2021, doi: 10.26858/jmtik.v4i3.23698.

[10] W. Sudrajat, I. Cholid, and J. Petrus, "Penerapan Algoritma K-Means Clustering untuk Pengelompokan UMKM Menggunakan Rapidminer," *J. JUPITER*, vol. 14, no. 1, pp. 27–36, 2022.

[11] S. Sajidah, R. Herdiana, and D. Solihudin, "Segmentasi Pelanggan Salon Nuii Beauty Glow Menggunakan K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 558–566, 2023, doi: 10.36040/jati.v7i1.6333.

[12] L. Pujiastuti, M. Wahyudi, and S. Solikhun, "Penerapan Data Mining Dalam Mengelompokkan Data Impor Tembaga Menurut Negara Asal Menggunakan Algoritma K-Means," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 2, no. 0, pp. 424–431, 2020, [Online]. Available: http://tunasbangsa.ac.id/seminar/index.php/senaris/article/view/191

[13] S. M. Dewi, A. P. Windarto, and D. Hartama, "Penerapan Datamining Dengan Metode Klasifikasi Untuk Strategi Penjualan Produk Di Ud.Selamat Selular," *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 3, no. 1, pp. 617–621, 2019, doi: 10.30865/komik.v3i1.1669.