

## IMPLEMENTASI DATA MINING CLUSTERING DATA VALUASI EKSPOR KERTAS INDONESIA MENGGUNAKAN ALGORITMA K-MEANS

Netiya Pooja, Muhandi Saputra, Siti Aisyah, Palma Juanta  
Program Studi Sistem Informasi, Fakultas Teknologi dan Ilmu Komputer, Universitas Prima Indonesia  
Jl. Sekip Kota Medan, Sumatera Utara 20111  
E-mail: palmajuanta@unprimdn.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. Jurnal ini membahas tentang penerapan *data mining clustering* pada nilai ekspor kertas berdasarkan pelabuhan asal di Indonesia menggunakan algoritma *K-Means*. Data yang digunakan dalam penelitian ini adalah data valuasi ekspor kertas dari Januari 2014 sampai Juni 2021 yang diambil dari publikasi Badan Pusat Statistik (BPS). Data diolah dengan melakukan *clustering* dalam 3 *cluster* yaitu *cluster* ekspor dengan nilai tinggi, nilai ekspor sedang dan nilai ekspor rendah. Sehingga nantinya dapat diperoleh informasi kluster pelabuhan-pelabuhan yang menghasilkan transaksi ekspor kertas yang tinggi. Juga akan menghasilkan informasi kluster pelabuhan-pelabuhan yang berada di kategori menengah untuk dikembangkan / diperdayakan oleh badan terkait di pemerintahan maupun perusahaan-perusahaan dibidang manufaktur kertas dan logistic. Proses clustering akan dihitung menggunakan *tools RapidMiner* yang merupakan salah satu *tools* terkemuka dalam bidang *data mining*. Diperoleh data hanya sekitar 10% pelabuhan yang terkategori memiliki transaksi ekspor kertas dengan valuasi tinggi, 3% level sedang dan 87% atau sekitar 27 pelabuhan memiliki catatan transaksi level rendah. Juga ditemukan sejumlah pelabuhan ekspor yang berdekatan secara geografis namun memiliki nilai transaksi yang sangat timpang satu sama lain.

**Kata kunci :** *Clustering, K-Means, Data Mining, Ekspor, Kertas, RapidMiner*

### 1. PENDAHULUAN

Kegiatan ekspor adalah sistem perdagangan dengan cara mengeluarkan barang-barang dari dalam negeri keluar negeri dengan memenuhi ketentuan yang berlaku. Ekspor merupakan total barang dan jasa yang dijual oleh sebuah negara ke negara lain, termasuk diantara barang-barang, asuransi, dan jasa-jasa pada suatu tahun tertentu[1]. Perdagangan internasional akan menciptakan keuntungan dengan memberikan peluang kepada setiap negara untuk mengekspor barang-barang yang diproduksinya menggunakan sumber daya yang langka dari negaranya tersebut. Indonesia adalah salah satu negara pengekspor ke negara-negara maju dan berkembang[2]. Barang-barang yang diekspor oleh Indonesia umumnya terdiri dari dua macam, yaitu minyak bumi dan gas alam (migas) serta juga non-migas.

Objek penelitian kali ini menitik beratkan pada ekspor produk hasil hutan dalam bentuk produk kertas berdasarkan daerah pelabuhan asal di Indonesia. Data yang digunakan dalam penelitian ini berdasarkan dokumen-dokumen keterangan ekspor yang dihimpun oleh Badan Pusat Statistik (BPS) melalui situs <https://www.bps.go.id>. Dalam penelitian ini analisis yang dilakukan menggunakan metode *clustering* dengan Algoritma *K-Means*.

Menurut Putri Elfa Mas'udia mengatakan, Data mining adalah proses menemukan hubungan dalam data yang tidak diketahui oleh pengguna dan menyajikannya dengan cara yang dapat dipahami

sehingga hubungan tersebut dapat menjadi dasar pengambilan keputusan[3]

*Clustering* merupakan salah satu teknik dari salah satu fungsionalitas *data mining*, algoritma *clustering* merupakan algoritma pengelompokan sejumlah data menjadi kelompok-kelompok data tertentu (*cluster*) [4]. Sedangkan *data mining*, sering juga disebut sebagai bagian dari proses KDD (*Knowledge Discovery in Databases*) yang terdiri dari beberapa tahapan seperti pemilihan data, pra pengolahan, transformasi, *data mining*, dan evaluasi hasil [5].

Selain itu algoritma *K-Means Clustering* dinilai memiliki teknik yang sangat cepat dalam proses *clustering* untuk mengelompokan data daerah pelabuhan asal.

Di Indonesia jumlah industri kertas dan pulp mencapai 62 perusahaan, dengan kapasitas industri kertas nasional mencapai 12,98 juta ton kertas/tahun. Industri pulp dan kertas telah menduduki peringkat ketujuh sebagai penyumbang devisa terbesar dari sektor nonmigas. Industri pulp dan kertas merupakan salah satu industri yang memegang peranan penting bagi perekonomian Indonesia.

Kelompok produk Kertas, Karton dan Barang Daripadanya (HS 48) adalah salah satu produk yang termasuk dalam 15 produk utama ekspor Indonesia. Selain itu Kelompok produk ini juga salah satu produk yang meningkat ekspornya pada bulan Agustus 2021. Indonesia berada pada peringkat ke-12 sebagai eksportir utama di dunia untuk sektor Kertas,

Karton dan Barang Daripadanya (HS 48) pada tahun 2020 dengan nilai ekspor USD 4,19 Miliar dan pangsa ekspor 2,68%. Pesaing utama Indonesia untuk produk ini adalah RRT dengan pangsa ekspor 13,38%, Jerman dengan pangsa ekspor 12,32% dan Amerika Serikat dengan pangsa ekspor 8,83% [6]

Sehingga dengan adanya pengelompokan data ini diharapkan pemerintah bisa memperoleh pengetahuan yang menjadi informasi penting serta dapat menjadi masukan kepada perusahaan dibidang logistik dan perusahaan produsen kertas mengenai pemetaan pelabuhan asal dengan tingkat valuasi rendah hingga tinggi untuk kegiatan ekspor produk kertas di Indonesia.

Adapun batasan masalah dalam penelitian ini yaitu:

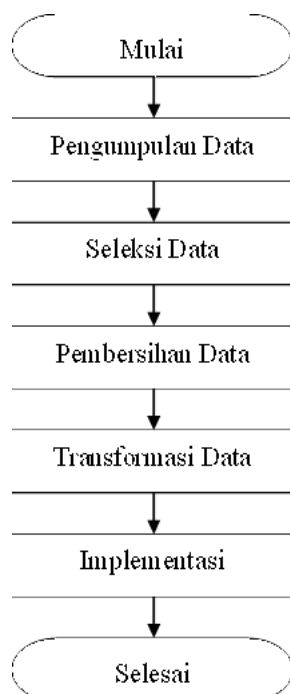
1. Metode yang digunakan adalah K-Means.
2. Pengujian menggunakan program *RapidMiner*.
3. Data yang digunakan adalah data ekspor kertas dari pelabuhan asal Indonesia pada periode Januari 2014 sampai dengan Juni 2021.
4. Produk yang diteliti adalah produk kertas dengan awalan HS Code 48 yang merupakan HS Code untuk produk kertas putih berbagai ukuran mulai dari A4, Folio hingga Sheeter Roll

## 2. ISI PENELITIAN

### 2.1 Metode Penelitian

Penelitian ini menggunakan teknik klasifikasi dengan *K-Means* untuk mengelompokkan data kategorikal hingga menghasilkan kluster yang lebih stabil[7]. *K-Means* merupakan salah satu metode terbaik dan paling populer dalam algoritma clustering dimana *K-Means* mencari partisi yang optimal dari data dengan meminimalkan kriteria jumlah kesalahan kuadrat dengan prosedur iterasi yang optimal[8].

Algoritma *K-Means Clustering* adalah suatu metode penganalisaan data atau metode *data mining* yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem



partisi. *Data mining* merupakan bidang dari beberapa bidang keilmuan yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, *database*, dan visualisasi untuk pengenalan permasalahan pengambilan informasi dari *database* yang besar [9].

Dalam algoritma ini, centroid juga diperlakukan sebagai data objek dalam kluster masing-masing, yang memberikan bantuan untuk menghindari masalah kluster kosong dan memberikan solusi yang optimal dari masalah karena semua data-objek disimpan dalam cluster terdekatnya, tujuan dari penelitiannya adalah peningkatan akurasi, tingkat presisi dan efisiensi algoritma [10].

### 2.2 Prosedur Penelitian

Langkah prosedur yang akan dilakukan dalam penelitian ini dapat dilihat pada *flowchart* berikut.

Gambar 1. Diagram Alur Penelitian

#### 2.2.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data produk kertas dengan awalan HS Code 48 tahun 2014 sampai 2020. Data tersebut dikumpulkan berdasarkan dokumen-dokumen keterangan ekspor impor yang dihimpun oleh Badan Pusat Statistik (BPS) melalui situs <https://www.bps.go.id>.

#### 2.2.2 Seleksi Data

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *clustering* dimulai. Data di seleksi hanya yang berasal dari asal berjenis Pelabuhan, lalu disimpan dalam suatu berkas, terpisah dari basis data utama.

Contoh dari data awal ditemukan sejumlah data bandar udara yang termasuk kedalam data mentah. Lalu penulis melakukan seleksi data dan mengeluarkan data ekspor yang melalui bandar udara.

#### 2.2.3 Pembersihan Data

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *clustering*. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*). Juga dilakukan proses *enrichment*, yaitu proses memperkaya data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk proses *clustering*, seperti data atau informasi eksternal.

Contoh dari data awal ditemukan sejumlah pelabuhan asal seperti pelabuhan Tenau, pelabuhan UjungPandang dan pelabuhan Wonosari yang hanya pernah melakukan ekspor produk kertas kurang dari 3 kali dalam 72 bulan terakhir sehingga dianggap

tidak valid untuk menjadi bagian data yang akan di *clustering*.

### 2.2.4 Transformasi data

Pada tahap ini dilakukan proses perubahan data, tujuannya adalah agar data dapat diolah dengan menggunakan algoritma *K-Means Clustering* dan algoritma ini hanya bisa diolah dengan angka.

### 2.2.5 Implementasi

Pada tahapan ini peneliti melakukan tahap akhir dimana data yang telah valid diinput kedalam aplikasi *RapidMiner*. *RapidMiner* adalah sebuah lingkungan *machine learning data mining*, *text mining* dan *predictive analytics*. Data yang diperoleh diolah dengan menggunakan perhitungan bobot dari tiap indeks. Data akan di bagi ke dalam 3 *cluster* yakni *cluster* tingkat valuasi ekspor tinggi, *cluster* tingkat valuasi ekspor sedang dan *cluster* tingkat valuasi ekspor rendah. Pada tahapan inilah akan dianalisis hasil tiap *cluster*-nya.

## 2.3 Hasil dan Pembahasan

### 2.3.1 Data

Sumber data utama yang digunakan dalam penelitian ini adalah *database* valuasi ekspor (USD) kertas asal Indonesia sejak Januari 2014 hingga Desember 2020 dengan jumlah keseluruhan data ada 469 *record*. Data diunduh dari sistem Badan Pusat Statistik (BPS) dalam bentuk tabel pada *Microsoft Excel* sehingga mempermudah untuk dilakukan pembersihan data atau memfilter data. Tabel data dari laman *web* BPS dapat dilihat pada tabel berikut.

Tabel 1. Data dari Badan Pusat Statistik

Pelabuhan	Date	Nilai Transaksi Perdagangan dalam US Dollar (USD)					
		Year_2014	Year_2015	Year_2016	Year_2017	Year_2018	Year_2019
ACHMAD YANI	1	5317	4602	6375	6504	4211	8417
ACHMAD YANI	2	5943	4262	5343	5604	5216	2558
ACHMAD YANI	3	5317	2149	3495	4195	8672	7690
ACHMAD YANI	4	2616	3696	2814	4188	8822	8822
ACHMAD YANI	5	13288	3474	3025	2267	7480	8430
ACHMAD YANI	6	5962	2924	3180	3597	8139	8514
ACHMAD YANI	7	3982	4936	8817	3955	4972	5287
ACHMAD YANI	8	8808	5187	2176	28888	3422	4148
ACHMAD YANI	9	5237	1782	3813	7229	4167	2923
ACHMAD YANI	10	5218	6782	6430	5870	8090	3095
ACHMAD YANI	11	6939	2788	4881	8116	3987	5247
ACHMAD YANI	12	5822	3818	8818	3122	8627	8115
ADI SUKPTO	1	833	0	0	0	0	0
ADI SUKPTO	2	0	196	47	19	0	0
ADI SUKPTO	3	0	0	2860	19	0	0
ADI SUKPTO	4	218	188	389	226	0	0
ADI SUKPTO	5	174	0	381	78	0	0
ADI SUKPTO	6	1182	398	798	190	0	0
ADI SUKPTO	7	2844	0	0	0	1880	0
ADI SUKPTO	8	86	0	0	0	0	0
ADI SUKPTO	9	0	0	0	0	0	0
ADI SUKPTO	10	0	0	0	0	438	2280
ADI SUKPTO	11	482	0	0	148	0	0
ADI SUKPTO	12	0	0	0	0	783	0
ALUJUNG	1	5216	14338	2949	20861	8738	7626
ALUJUNG	2	1787	7382	1918	14574	7783	4973
ALUJUNG	3	3728	8886	11876	11986	7083	7136
ALUJUNG	4	11485	3622	7238	6190	8922	8658
ALUJUNG	5	9784	4485	10336	1004	3513	2658
ALUJUNG	6	6532	5848	10388	8778	8831	3032
ALUJUNG	7	12887	5083	6397	7883	7782	18122

### 2.3.2 Penyeleksian Data

Langkah selanjutnya adalah peneliti melakukan penyeleksian data dengan mengeluarkan data yang tidak valid atau tidak lengkap. Lalu peneliti juga menyeleksi dengan hanya menggunakan satu *variable* 'nama\_pelabuhan' sebagai label dan satu *variable* lain bernama 'valuasi\_expor\_usd' sebagai *variable* data. Penulis

juga melakukan seleksi dengan hanya menggunakan data pelabuhan asal yang memiliki rata-rata nilai ekspor diatas USD \$ 20.000.

Untuk memudahkan proses pembacaan data dari aplikasi *RapidMiner* maka data yang telah di seleksi ini kemudian di pindahkan ke dalam tabel basisdata MySQL, seperti pada gambar berikut.

Gambar 2. Data yang telah di impor ke DB MySQL

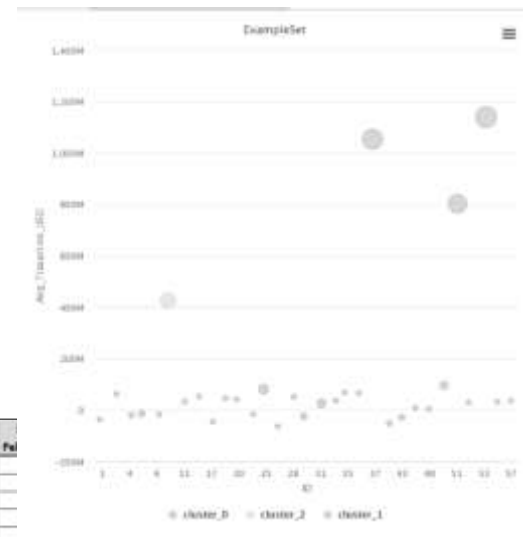
### 2.3.3 Transformasi Data

Pada tahap ini dilakukan proses perubahan data, agar data dapat diolah dengan menggunakan algoritma k-means clustering. Data yang berupa non-numerik dilakukan proses inisiasi ke dalam bentuk numerik. Namun jika data yang dimiliki sudah dalam bentuk *numeric* maka tidak diperlukan inisiasi. Atribut yang digunakan pada penelitian ini adalah 'Nama\_Pelabuhan', Tahun dan 'NilaiEkspor\_USD'. Atribut yang akan diubah adalah Nama\_Pelabuhan menjadi Kode\_Pelabuhan berupa angka numerik. Proses inisiasi nama pelabuhan ke bentuk kode inisial dapat dilihat di tabel berikut :

Tabel 2. Data Atribut Kode Pelabuhan

### 2.3.4 Pengolahan Data

Pada tahapan ini dilakukan pemodelan data menggunakan hasil dari transformasi data seperti tabel 2, adapun metode yang dipakai pada penelitian ini adalah metode *clustering* dengan



Nama Pelabuhan	Kode	Nilai
ACHMAD YANI	1	36
ADI SUKPTO	2	37
ASAFUFI	3	38
BATU AMPAR	4	39
BELOWAN	5	40
BENGLALUJUNG	6	41
BIRIP KASSI	7	42
BUATAP	8	43
ENTEDING	9	44
GEDE BAGE-BANDUNG (PTT)	10	45
HALIM PERDANA KUSUMA	11	46
HANG NAGIM	12	47
HANGAJUDIN	13	48
HUSEIN SASTRANEGARA	14	49
LANGKABALANG	15	50
JAKARTA / PASAR DAN	16	51
JAMBI	17	52
JARAPURA	18	53
JARAPURA / SENTANI	19	54
JUANJA-SURABAYA	20	55
KABU'PAHAU	21	56
KARANGALI	22	57
KIDAP	23	58
KUALA NAMU INTERNATIONAL AIRPORT	24	59
KUALA TUNGKAL	25	60
LINGSAR TARAKAN	26	61
LORAM	27	62
MANDOWAJI	28	63
MEGAN / POLONGA	29	64
MELAK	30	65
PEKAN BARU	36	
PERAWANG SURABAYA	37	
PONTIANAK	38	
SAM RATULANGI	39	
SEKELANG	40	
SEPPINGGAR	41	
SM. BADARUDDIN	42	
SOEKARNO-HATTA	43	
SOKO/TERRE/ADA LUMAJANG	44	
SURABAYA GUMILANG	45	
SURABAYA (PTT)	46	
TANJUNG BALAI ASAHAN	47	
TANJUNG BALAI KAPIMUN	48	
TANJUNG BATU, BULU	49	
TANJUNG EMAS	50	
TANJUNG PERAK	51	
TANJUNG PINANG	52	
TANJUNG PRIOK	53	
TANJUNG UJIAN	54	
TENALI	55	
TUMANGKANG	56	
WONOSARI	57	

menggunakan algoritma *K-Means*. Data yang sudah dikumpulkan, diseleksi dan ditransformasi akan dilakukan pengolahan dari data tersebut dengan

menggunakan metode *clustering*. Metode ini bekerja dengan melakukan pengelompokan data yang memiliki kesamaan karakteristik pada setiap data. Dari hasil perhitungan diperoleh nilai *centroid* sebagai berikut:

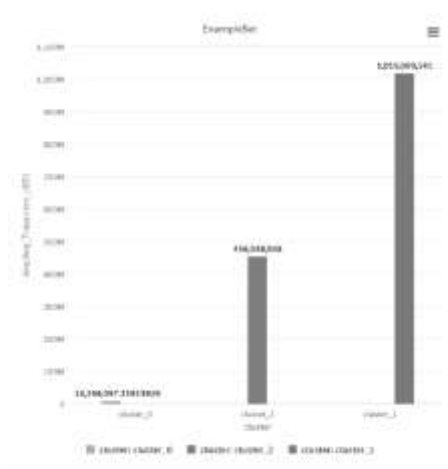
Tabel 3. Nilai *centroid* masing-masing *cluster*.

Attribute	cluster_0	cluster_1	cluster_2
Trasaction_USD	\$ 10.266.097	\$ 1.019.009.541	\$ 456.048.048

Dari nilai *centroid* tersebut data kemudian dapat dikelompokkan ke masing-masing *cluster*.

Tabel 4. Hasil *clustering* berdasarkan nilai *centroid*

Cluster	AvgMonthly Trasaction (USD)	ID Pelabuhan	Nama Pelabuhan
Cluster_1	\$ 1.175.871.720	53	TANJUNG PRIOK
Cluster_1	\$ 1.011.271.807	37	PERAWANG, RIAU
Cluster_1	\$ 869.885.096	51	TANJUNG PERAK
Cluster_2	\$ 456.048.048	8	BUATAN PORT, RIAU
Cluster_0	\$ 107.618.595	25	KUALA TUNGKAL
Cluster_0	\$ 66.223.759	31	MUSI RIVER/BOOM BARU
Cluster_0	\$ 61.083.800	50	TANJUNG EMAS
Cluster_0	\$ 14.370.236	30	MERAK
Cluster_0	\$ 12.594.176	5	BELAWAN
Cluster_0	\$ 3.810.119	43	SOEKARNO-HATTA
Cluster_0	\$ 3.001.100	40	SEKUPANG
Cluster_0	\$ 2.696.150	4	BATU AMPAR
Cluster_0	\$ 1.513.026	46	SURABAYA
Cluster_0	\$ 1.472.367	32	NGURAH RAI
Cluster_0	\$ 1.184.256	20	JUANDA-SURABAYA
Cluster_0	\$ 373.808	27	LOBAM
Cluster_0	\$ 180.502	57	WONOSARI
Cluster_0	\$ 159.666	35	PANJANG
Cluster_0	\$ 144.246	48	TANJUNG BALAI KARIMUN
Cluster_0	\$ 129.856	54	TANJUNG UBAN
Cluster_0	\$ 117.318	36	PEKAN BARU
Cluster_0	\$ 100.000	28	MANOKWARI
Cluster_0	\$ 95.722	3	ATAPUPU
Cluster_0	\$ 65.613	1	ACHMAD YANI
Cluster_0	\$ 46.485	23	KUJANG
Cluster_0	\$ 44.272	12	HANG NADIM
Cluster_0	\$ 41.080	17	JAMBI
Cluster_0	\$ 37.453	11	HALIM P.K.
Cluster_0	\$ 31.248	52	TANJUNG PINANG
Cluster_0	\$ 27.999	6	BENOA/LOLOAN
Cluster_0	\$ 21.774	18	JAYAPURA



Gambar 3. Hasil *clustering* dalam bentuk Bar Chart  
Gambar 4. Hasil *clustering* dalam bentuk Scatter Chart

Berdasarkan pengujian yang telah dilakukan pada *RapidMiner Studio* dengan melihat

pengelompokkan *cluster* pada tabel sebelumnya, maka kita bisa mendapatkan hasil sebagai berikut:

1. Di-Indonesia kita memiliki 3 pelabuhan ekspor produk Paper dengan nilai transaksi rata-rata mencapai range \$ 850.000.000 hingga \$1.175.000.000 perbulan selama 6 tahun terakhir.
2. Sedangkan di kategori menengah atau Pelabuhan yang berpotensi dikembangkan ke cluster 1 adalah Buatan Port, Riau. Pelabuhan yang terletak di Kabupaten Siak ini berpotensi untuk bisa terus meningkatkan valuasi ekspor kertas-nya karena merupakan Pelabuhan asal ekspor yang utamanya mengelola produk kertas
3. Sedangkan untuk 27 pelabuhan lain-nya berada di *cluster* 0 atau yang nilai rata-rata valuasi per bulan-nya cukup rendah. Data ini menjadi petunjuk penting bagi pemangku kepetingan di pemerintahan maupun di perusahaan bidang *logistic* untuk melihat kendala yang terjadi dan untuk bisa meningkatkan jumlah ekspor dari pelabuhan-pelabuhan di *cluster* ini.

### 3. KESIMPULAN

Berdasarkan pembahasan sebelumnya, maka dapat diambil kesimpulan bahwa :

1. Dari total 57 pelabuhan asal di Indonesia yang melakukan kegiatan ekspor produk Kertas dapat kita kelompok-kan menjadi 3 *cluster* utama yaitu:
  - *Cluster* 1 terdapat 3 pelabuhan yang terkategori sebagai *cluster* 1 yang memiliki rata-rata nilai penjualan ekspor kertas perbulan sangat tinggi selama 6 tahun terakhir.
  - *Cluster* 2 terdapat 1 pelabuhan yang dikategorikan sebagai *cluster* 2 dengan rata-rata nilai penjualan ekspor kertas perbulan stabil diangka menengah.
  - *Cluster* 0 terdapat 27 pelabuhan lainnya yang termasuk kedalam cluster 0 atau rata-rata bernilai rendah selama 6 tahun terakhir. Dan ada 26 pelabuhan lainnya yang tereliminasi saat proses penyeleksian data dikarenakan sejumlah *factor* seperti data yang tidak valid, pelabuhan telah tidak aktif.
2. Pengujian menggunakan aplikasi *Rapidminer* mampu membantu untuk mendapatkan hasil pengelompokan yang akurat, berimbang dan cepat.

### 4. PENUTUP

Demikianlah hasil penelitian mengenai Implementasi Data Mining Clustering Data Valuasi Ekspor Kertas Indonesia Dengan Algoritma *K-Means* yang telah dibuat. Peneliti berharap hasil dari pengolahan data yang telah dilakukan dapat memberikan solusi kepada pihak manufaktur produk kertas agar dapat mengetahui potensi pelabuhan asal ekspor yang dapat di perdayakan untuk menunjang kegiatan ekspor kertas dari Indonesia.

## DAFTAR PUSTAKA

- [1] Adrian Sutedi. Hukum Ekspor Impor, Cetakan 1, (2014), Hlm.7
- [2] Halwani, Hendra. 2005. Ekonomi Internasional dan Globalisasi Ekonomi. Bogor. Ghalia Indonesia
- [3] P. N. Malang, “Klasifikasi Tugas Akhir Untuk Menentukan Dosen Pembimbing Menggunakan Naïve Bayes Classifier (Nbc),” vol. 7, pp. 135–140, 2015.
- [4] Anjelita, M., Windarto, A. P., Wanto, A., & Saifullah, S. (2019, August). Analisis Metode K-Means pada Kasus Ekspor Barang Perhiasandan Barang Berharga Berdasarkan Negara Tujuan. InSeminar Nasional Sains dan Teknologi Informasi (SENSASI), vol.2, no.1.
- [5] A. M. Siregar dan A. Puspabhuana, Data Mining: Pengolahan Data Menjadi Informasi dengan RapidMiner, CV Kekata Group, 2017.
- [6] Badan Pengkajian dan Pengembangan Perdagangan, Kajian Kinerja Perdagangan Ekspor Impor 2021 (2021, Oktober).
- [7] T. Suprawoto, “Klasifikasi Data Mahasiswa Menggunakan Metode K-Means untuk Menunjang Pemilihan Strategi Pemasaran,” J. Inform. dan Komput., vol. 1, no. 1, pp. 12–18, 2016
- [8] E. Muningsih and S. Kiswati, “Penerapan Metode K-Means Untuk Clustering Produk Online Shop Dalam Penentuan Stok Barang,” J. Bianglala Inform., vol. 3, no. 1, pp. 10– 17, 2015.
- [9] Susanto, “Implementasi Data Mining untuk Memprediksi Masa Studi Mahasiswa menggunakan Algoritma C4.5”, Jurnal Media Infotama Vol 11 No. 2, 2015
- [10] J. Wang, “An improved K-Means clustering algorithm,” Technology, vol. 5, no. m, pp. 4–6, 2011.