

ANALISIS CLUSTER DATA DAFTAR KENDARAAN BERMOTOR MENGUNAKAN ALGORITMA K-MEANS

Mariani Gita Budiarti, Nining Rahaningsih, Raditya Danar Dana

Teknik Informatika, Sekolah Tinggi Manajemen dan Informatika IKMI Cirebon
Jalan Perjuangan No. 10B Karyamulya Kec. Kesambi Kota Cirebon Jawa Barat 45135, Indonesia
marianigita9@gmail.com

ABSTRAK

Dalam era modern ini, kendaraan bermotor telah menjadi sarana transportasi yang sangat umum dan mudah diakses oleh semua orang, menjadi kendaraan dengan pertumbuhan tercepat dan menghasilkan data dalam jumlah besar. Tantangan besar muncul dalam mengelola informasi dari Big Data yang dihasilkan oleh kendaraan bermotor. Penelitian ini memiliki tujuan utama untuk membuat model kelompok di cabang pelayanan yang menunjukkan kemiripan data di Provinsi Jawa Barat. Data daftar ulang kendaraan bermotor dikumpulkan selama 9 tahun dan digunakan sebagai sumber data, mencakup informasi seperti jumlah kendaraan, jenis kendaraan, dan lokasi geografis. Data yang kompleks ini diproses menggunakan algoritma K-Means, dengan langkah awal berupa pra-pemrosesan data, termasuk pemilihan atribut relevan dan penghapusan data tidak valid. Hasil clustering menunjukkan bahwa cluster optimal berada pada cluster 3, dengan nilai Davies-Bouldin Index (DBI) sebesar 0,270, dan titik centroid sebesar 1.287.880.558,278. Keselarasan antara nilai DBI dan titik centroid menegaskan konsistensi pada cluster 3. Temuan ini memberikan informasi berharga untuk pengambilan keputusan lebih efektif dalam pengelolaan dan pengembangan infrastruktur transportasi di Provinsi Jawa Barat. Hasil penelitian ini berpotensi meningkatkan efisiensi dan keselamatan lalu lintas serta memenuhi kewajiban pajak kendaraan bermotor, memberikan dasar kokoh untuk penyusunan kebijakan yang terarah dan berkelanjutan dalam meningkatkan sistem transportasi di wilayah tersebut.

Kata kunci: K-Means, Clustering, Daftar, Kendaraan, Bermotor

1. PENDAHULUAN

Kebaikan bermasa dapat dicapai melalui penerapan tata kelola yang efektif dan pembangunan yang adil di seluruh wilayah. Salah satunya sumber dana pemerintah adalah dari penerimaan pajak [1]. Pada Pasal 285 Undang-Undang Nomor 32 Tahun 2014, pendapatan daerah meliputi pajak daerah, hasil pajak daerah, hasil retribusi daerah, hasil pengelolaan kekayaan daerah yang dipisahkan dan lain-lain PAD yang disahkan, dana transfer, dan lain-lain pendapatan daerah yang sah. Yang di mana pasal demi pasal untuk hasil pajak daerah cukup jelas. Perpindahan kekayaan dari masyarakat ke kas negara untuk membiayai pengeluaran rutin dan surplusnya digunakan untuk investasi publik dikenal sebagai perpajakan. Pajak dapat dijadikan alat untuk membangun suatu negara agar semua orang dapat menikmati kesejahteraan. Kemudian Pasal 1 UU Nomor 11 tahun 2009 kesejahteraan sosial adalah kondisi terpenuhinya kebutuhan material, spiritual, dan sosial warga negara agar dapat hidup layak dan mampu mengembangkan diri, sehingga dapat melaksanakan fungsi sosial.

Dalam beberapa dekade terakhir, perubahan sosial di negara-negara berkembang telah menjadi tema utama. Jumlah kendaraan bermotor berkontribusi pada pendapatan bagi daerah, dapat mengganggu kepatuhan perpajakan, dan akan menimbulkan masalah keamanan. Penelitian ini fokus pada integritas teknologi pengelompokan kendaraan bermotor pada cabang pelayanan. Dengan menggunakan data historis, penelitian ini bertujuan untuk mengetahui pemahaman awal pola kelompok kendaraan bermotor.

Tantangannya meliputi pengembangan algoritma K-Means tidak hanya menyusun data secara efisien menjadi kelompok berdasarkan kemiripan data, tetapi juga pra-pemrosesan data yang teliti, mengingat data bisa berisi kesalahan atau tidak valid. Selanjutnya, menentukan jumlah cluster yang optimal dan interpretasi hasil yang informatif adalah tantangan kritis, karena kesalahan dalam hal ini dapat menghasilkan keputusan yang kurang efektif dan kurang menguntungkan.

Penelitian ini bertujuan untuk mengisi kesenjangan pengetahuan di bidang pajak kendaraan bermotor dengan mengumpulkan data dari berbagai sumber. Melalui penggunaan clustering, penelitian ini berupaya mengidentifikasi kelompok kendaraan bermotor yang di mana untuk memberikan kontribusi yang komprehensif pada bidang ilmu komputer. Meskipun analisis data historis memberikan informasi yang berharga, tantangan untuk menyajikan informasi yang relevan meningkat seiring dengan perubahan dinamis dalam lingkungan. Oleh karena itu, kemampuan untuk mengakses, menganalisis, dan menafsirkan data sangat penting untuk mengambil tindakan proaktif dalam menanggapi meningkatnya jumlah kendaraan bermotor.

Metode yang digunakan adalah mengumpulkan data Daftar Kendaraan Bermotor selama periode waktu tertentu. Selanjutnya, data ini akan dipersiapkan melalui proses pra-pemrosesan, termasuk pemilihan atribut yang relevan dan penghapusan data yang tidak valid. Algoritma K-Means akan digunakan untuk mengelompokkan kendaraan berdasarkan atribut yang

dipilih, dengan menggunakan metrik jarak antara titik data. Hasil dari pengelompokan ini akan membantu dalam mengidentifikasi pola kendaraan yang memungkinkan tindakan yang lebih terfokus dalam penegakan aturan dan peningkatan kesadaran perpajakan. Dengan demikian, pendekatan ini memanfaatkan teknik analisis data dan clustering untuk memberikan wawasan yang mendalam dalam mengatasi permasalahan yang melibatkan jumlah kendaraan bermotor.

Hasil penelitian ini diharapkan dapat memberikan informasi jumlah kendaraan bermotor setiap tahunnya terhadap kewajiban pajak kendaraan. Implikasi ini dapat mencakup peningkatan kesadaran masyarakat akan kewajiban pajak kendaraan, pengembangan strategi yang lebih efektif untuk meningkatkan kepatuhan pajak, dan memungkinkan pengembangan solusi IT yang lebih cerdas untuk pengelolaan data pajak kendaraan. Oleh karena itu, penelitian ini memberikan dampak positif terhadap praktik, penelitian dan pengembangan teknologi di bidang ilmu komputer.

2. TINJAUAN PUSTAKA

Dalam penelitian sebelumnya, peneliti mendapati setidaknya sepuluh jurnal yang dianalisis sebagai acuan dalam penelitian ini. Perbandingan sebelumnya akan menunjukkan kejelasan yang kuat untuk relevansi dan kontribusi penelitian ini.

Salah satunya pada jurnal [2] menjelaskan beberapa peraturan terkait pajak kendaraan bermotor, seperti Peraturan No. 23 Januari 2011 tentang Perhitungan Pajak Kendaraan Bermotor dan Bea Balik Nama Kendaraan Bermotor, peraturan Daerah Provinsi Jawa Barat Nomor 23, Peraturan Gubernur No. 13 tahun 2011 tentang Pajak Daerah, dan Peraturan Nomor 68 Tahun 2011 tentang Perhitungan Dasar Pengenaan Pajak Kendaraan Bermotor (PKB) dan Bea Balik Nama Kendaraan Bermotor. Jurnal ini juga mencakup metode penelitian, terutama pendataan di Bapenda Jawa Barat dan penyusunan data berdasarkan RT dan RW. Hasil penelitian menunjukkan beberapa permasalahan, seperti kurangnya pemahaman dan kesadaran wajib pajak terhadap edukasi yang diberikan oleh Bapenda. Kesimpulan dan rekomendasi jurnal ini menyebutkan adanya wajib pajak yang belum menyatakan pindah alamat, belum bayar pajak, dan belum melakukan registrasi ulang kendaraan. Jurnal ini juga menyarankan Bapenda Jawa Barat untuk meningkatkan keakuratan data dan melibatkan pegawai yang bertanggungjawab dan amanah dalam berkomunikasi dengan wajib pajak.

Jurnal [3] membahas pentingnya penerimaan pajak kendaraan bermotor dalam meningkatkan pendapatan provinsi dan mendukung pembangunan infrastruktur. Penelitian ini bertujuan menguji pengaruh jumlah kendaraan bermotor dan kendaraan tidak terdaftar terhadap penerimaan pajak kendaraan. Menggunakan data sekunder dari Kantor Pusat

Pengelolaan Pendapatan Daerah Kabupaten Cirebon Sumber 1 tahun 2017-2021 dengan 60 sampel, penelitian menggunakan metode universal sampling dan analisis regresi linier berganda. Hasilnya menunjukkan bahwa jumlah kendaraan berdampak positif terhadap penerimaan pajal, sementara dampak tidak registrasi ulang kendaraan masih belum jelas. Disarankan untuk meningkatkan kesadaran wajib pajak terhadap pembayaran pajak kendaraan. Secara teoritis, disarankan penelitian lebih lanjut di Kabupaten Indramayu dengan penambahan variabel penerimaan pajak kendaraan. Meskipun memberikan informasi mengenai dampak jumlah kendaraan terhadap penerimaan pajak kendaraan, beberapa informasi seperti dampak STNK masih perlu klarifikasi lebih lanjut.

Jurnal [4] mengkaji pengaruh kualitas pelayanan, sanksi perpajakan, dan kondisi pembiayaan terhadap kepatuhan wajib pajak kendaraan bermotor di SAMSAT Kota Padang. Populasi penelitian mencakup 58447 kendaraan terdaftar di kantor SAMSAT tersebut. Metode analisis regresi digunakan dengan variabel terkait Kepatuhan Wajib Pajak (Y), dan variabel bebas Sanksi Administrasi (X1), Tingkat Pendapatan (X2), dan SAMSAT *Drive Thru* (X3). Hasil analisis deskriptif menunjukkan bahwa sanksi administrasi, tingkat pendapatan, dan SAMSAT *Drive Thru* secara bersama-sama berpengaruh signifikan terhadap kepatuhan wajib pajak.

Jurnal [5] bertujuan mengetahui sejauh mana pengaruh pajak, sanksi pajak kendaraan bermotor, kesadaran pajak dan sistem e-Samsat. Hasil penelitian ini menunjukkan bahwa pengetahuan perpajakan, sanksi pajak kendaraan bermotor, kesadaran wajib pajak dan sistem e-Samsat dapat mempengaruhi kepatuhan wajib pajak dalam membayar pajak kendaraan bermotor di Kabupaten Subang. Penelitian ini peneliti menggunakan kerangka teori *Theory of Planned Behavior* (TPB) yang menjelaskan bahwa perilaku atau sikap seseorang disebabkan oleh tiga faktor penentu, yaitu: percaya akan akibat perilaku dan mengevaluasi hasil perilaku, mematuhi, dan kepatuhan dengan peraturan ini. Penelitian ini juga menambahkan variabel independen yaitu sistem e-Samsat.

Jurnal [6] membahas tentang penerapan data mining menggunakan algoritma K-Means *clustering* untuk menentukan strategi promosi mahasiswa baru Universitas Bina Darma Palembang. Penelitian ini bertujuan untuk mengetahui strategi periklanan yang efektif untuk menarik minat calon mahasiswa baru Universitas Bina Darma Palembang. Metode yang digunakan adalah data mining dengan menggunakan algoritma K-Means *clustering*. Hasil penelitian menunjukkan bahwa strategi promosi seperti menyebarkan leaflet, membuat desain yang menarik dan menyebarkan di banyak tempat, disertai dengan promosi dari teman dekat atau saudara merupakan strategi promosi sekolah yang baik untuk menarik

lebih banyak siswa untuk belajar IPA di Universitas Bina Darma.

Jurnal [7] membahas menggunakan data mining dan algoritma K-Means untuk mengklasifikasi gejala Infeksi Saluran Pernafasan Akut (ISPA). Berdasarkan studi literatur sebelumnya, algoritma K-Means terbukti efektif dalam mengklasifikasi data dengan akurasi tinggi. Penelitian ini juga mencakup penggunaan data mining untuk identifikasi gejala penyakit, termasuk ISPA. Kajian literatur melibatkan teori-teori seperti *Knowledge Discovery in Databases* (KDD), Data Mining, Klasifikasi, dan Algoritma K-Means. Penelitian menggunakan data primer yang dikumpulkan melalui observasi, wawancara, dan penelitian dokumen. Data kemudian dipersihkan dan diseleksi menggunakan algoritma K-Means *Clustering* untuk membentuk beberapa *cluster*. Hasil penelitian menegaskan bahwa algoritma K-Means dapat sukses mengklasifikasi gejala ISPA dengan akurasi yang tinggi.

Jurnal [8] Dalam penelitian ini, data mining digunakan untuk memprediksi banjir di Kota Padang dengan memanfaatkan data historis dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). Metode penelitian melibatkan pencarian, pencatatan, konstruksi, analisis, dan penyusunan laporan berdasarkan fakta ilmiah. Dilakukan penelitian lapangan, pembahasan masalah, pengumpulan data, analisis menggunakan metode K-Means *clustering*, implementasi algoritma K-Means, pengujian perangkat lunak, dan evaluasi hasil penelitian. Teknologi data mining dapat memprediksi banjir di Kota Padang dengan menggunakan K-Means *clustering*. Data dikelompokkan menjadi tiga tingkat (tinggi, sedang, rendah) dari situasi iklim Kecamatan Kabupaten Solok Selatan BMKG. Mekanisme peringatan dini melibatkan koordinasi antar BMKG dan Pemerintah untuk memberikan informasi tentang bahaya banjir.

Jurnal [9] membahas tentang penggunaan algoritma K-Means dan optimasi metode *elbow* untuk mengelompokkan data pengguna narkoba. Penelitian ini bertujuan untuk membuktikan kinerja *elbow* untuk menghasilkan *cluster* yang optimal pada data pengguna narkoba menggunakan algoritma K-Means sebagai metode pengelompokan data. Hasil pengujian memperlihatkan bahwa metode *elbow* dapat menghasilkan *cluster* yang optimal pada data pengguna narkoba. Algoritma K-Means merupakan model yang menggunakan *centroid* untuk membuat *cluster*. *Centroid* merupakan titik tengah dari suatu *cluster*. Secara umum algoritma K-Means memiliki langkah-langkah seperti menentukan objek ke dalam *cluster* yang memiliki *centroid* terdekat.

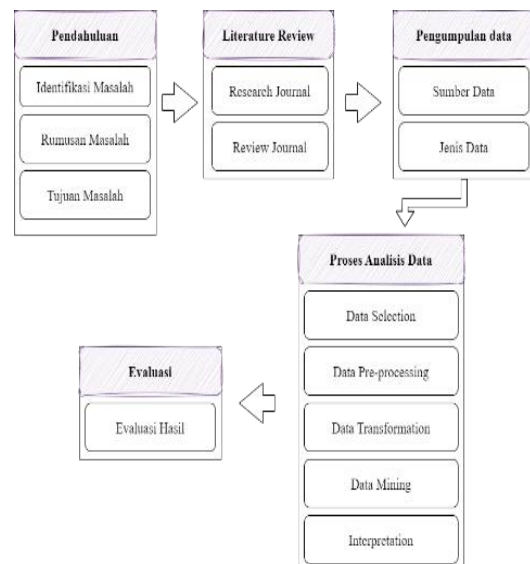
Jurnal [10] dengan membahas pengelompokkan data komponen Indeks Pembangunan Manusia di Provinsi Jawa Tengah pada tahun 2018 silam menggunakan algoritma K-Means dengan metode *elbow*. Penelitian ini membantu pemerintah dalam mempertimbangkan pengambilan kebijakan pada

wilayah Kabupaten/Kota di Jawa Tengah berdasarkan variabel-variabel IPM. Metodologi penelitian menunjukkan bahwa terdapat dua *cluster* dengan karakteristik yang berbeda. *Cluster* 1 memiliki komponen pembentukan IPM lebih rendah dibandingkan dengan *cluster* 2. Penelitian ini memberikan saran untuk menggunakan alternatif lain selain metode *elbow* dalam mencari validitas *cluster* terbaik dan menggunakan bilangan acak untuk menentukan *cluster* awal pada perhitungan manual K-Means di *Microsoft Excel*.

Jurnal [11] membahas tentang penggunaan model data mining *clustering* untuk melakukan pengelompokan data. Salah satu algoritma yang digunakan adalah K-Means. Proses menentukan banyak kelompok yang ideal dilakukan dengan menggunakan teknik evaluasi model yaitu *Davies-Bouldin Index*, metode *elbow*, dan metode *silhouette*. Hasil pengelompokan dapat digunakan untuk memetakan *dataset* berdasarkan pada kategori tertentu.

3. METODE PENELITIAN

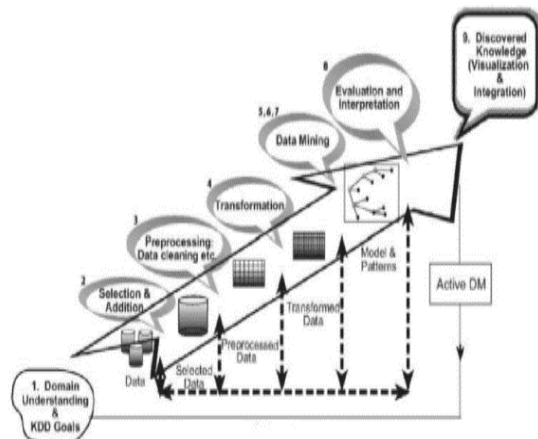
Dalam penelitian ini menggunakan metode pendekatan kuantitatif eksperimen yang melibatkan penggunaan teknik kontrol eksperimental. Teknik kontrol eksperimental adalah metode yang memungkinkan untuk mengungkapkan suatu masalah dengan menganalisis pengumpulan data secara sistematis. Tujuan pendekatan ini adalah untuk memperoleh gambaran yang lebih jelas dan rinci mengenai situasi atau fenomena yang menjadi fokus penelitian. Adapun tahapan penelitian tersebut seperti Gambar 1. Tahapan Metode Penelitian dibawah ini.



Gambar 1. Tahapan Metode Penelitian

Pada metode penelitian akan melalui tahapan *Knowledge Discovery in Database* (KDD). Dimana proses data untuk mengekstrak semua yang dianggap dengan spesifikasi ukuran dan batas, data yang digunakan secara bersama dengan preprocessing yang

dibutuhkan. Kegiatan penelitian terdiri 5 tahapan untuk melakukan penelitian yang terdiri dari data selection, data preprocessing, data transformation, data mining, dan interpretation. Berdasarkan Gambar 2. Proses KDD dibawah ini dapat dijelaskan sebagai berikut.



Gambar 2. Proses KDD [12]

- Data Selection**
Data yang pertama kali dilakukan pada proses seleksi untuk menentukan attributes yang akan digunakan.
- Data Pre-processing**
Kemudian setelah dilakukannya seleksi terhadap data tersebut maka dilakukannya proses cleaning data untuk menghilangkan missing value, atau data yang tidak konsisten dan tidak relevansi. Dan dilakukan integrasi data untuk atribut-atribut yang mengidentifikasi entitas yang unik.
- Data Transformation**
Selanjutnya, tahap ini menyesuaikan data dengan algoritma yang akan digunakan dalam Data Mining.
- Data Mining**
Tahapan ini, data akan diproses untuk mendapatkan informasi atau model tertentu menggunakan metode-metode dalam Data Mining.
- Interpretation**
Tahapan terakhir, tahapan penilaian untuk hasil dari proses Data Mining apakah sudah memenuhi target yang ditentukan atau belum, kemudian dari KDD ini adalah agar pengetahuan yang diperoleh dapat memberikan manfaat kepada penggunaannya [13].

4. HASIL DAN PEMBAHASAN

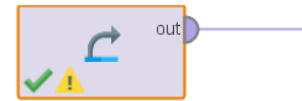
Data daftar kendaraan bermotor dibuat dalam dataset dengan format .xlsx (excel) untuk bisa dianalisis software yang dipakai RapidMiner Studio.

4.1. Data selection

Import data, pada aplikasi *RapidMiner* fitur *import data* berfungsi sebagai memasukan data yang berada di *computer* pengguna ke dalam proses

RapidMiner. Operator ini dapat digunakan untuk memuat data dari *spreadsheet Microsoft Excel*. Dari parameter *Retrieve* bisa dilihat informasi pada Tabel 1. Statistik Dataset Daftar Kendaraan Bermotor.

Retrieve Data Daftar Kendaraan Bermotor



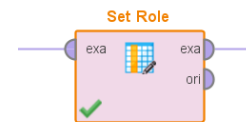
Gambar 3. Retrieve Daftar Kendaraan Bermotor

Tabel 1. Statistik Dataset Daftar Kendaraan Bermotor

No.	Uraian	Keterangan
1.	Record	4590
2.	Special Attributes	0
3.	Reguler Attributes	11
4.	Attributes:	
	Id.	Integer, missing value 0
	Kode_provinsi	Integer, missing value 0
	Nama_provinsi	Nominal, missing value 0
	Kode_kabupaten_kota	Integer, missing value 0
	Nama_kabupaten_kota	Nominal, missing value 0
	Cabang_pelayanan	Nominal, missing value 0
	Jenis_kendaraan	Nominal, missing value 0
	Fungsi_kendaraan	Nominal, missing value 0
	Jumlah_kendaraan	Integer, missing value 0
	satuan	Nominal, missing value 0
	Tahun	Integer, missing value 0

4.2. Data Pre-processing

Kemudian operator *Set Role* fitur ini berfungsi untuk menandakan *attribute* pada baris pertama dengan penamaan agar tidak diikutsertakan hitung dan mengubah.



Gambar 4. Operator Set Role

Parameter yang digunakan pada saat *Set Role* adalah sebagai Tabel 2. Tabel Parameter *Set Role*

Tabel 2. Parameter Set Role

No.	Uraian	Keterangan
1.	Attribute Filter Type	subset
2.	Selected attribute	cabang_pelayanan

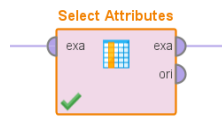
Berikut hasil setelah melakukan *Set Role* pada Tabel 3. Statistik *Set Role*.

Tabel 3. Statistik Set Role

No.	Uraian	Keterangan
1.	Record	4590
2.	Special Attributes	1
3.	Reguler Attributes	10
4.	Attributes:	
	Id.	Integer, missing value 0
	Kode_provinsi	Integer, missing value 0
	Nama_provinsi	Nominal, missing value 0
	Kode_kabupaten_kota	Integer, missing value 0
	Nama_kabupaten_kota	Nominal, missing value 0

No.	Uraian	Keterangan
	Cabang_pelayanan	Nominal, missing value 0
	Jenis_kendaraan	Nominal, missing value 0
	Fungsi_kendaraan	Nominal, missing value 0
	Jumlah_kendaraan	Integer, missing value 0
	satuan	Nominal, missing value 0
	Tahun	Integer, missing value 0

Select attribute dari keseluruhan atribut yang berjumlah 11 hanya akan dipakai beberapa atribut saja yang relevan, pada operator select attributes di bagian parameter dilakukannya program sehingga bisa menampilkan atribut yang dibutuhkan saja seperti type include attributes yang artinya memilih semua kolom yang ada pada dataset tersebut, kemudian attribute filter type menggunakan a subset dimana pada parameter ini bisa memilih atribut mana saja yang akan digunakan untuk proses selanjutnya dilakukan



Gambar 5. Operator Select Attributes

Parameter yang digunakan pada saat *Select Attributes* adalah sebagai Tabel 4. Tabel Parameter *Select Attributes*

Tabel 4. Parameter Select Attributes

No.	Uraian	Keterangan
1.	Attribute Filter Type	A subset
2.	Selected attribute	cabang_pelayanan, jenis_kendaraan, fungsi_kendaraan, jumlah_kendaraan

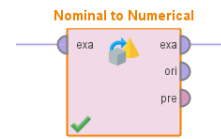
Berikut hasil setelah melakukan *Set Role* pada Tabel 4. Statistik *Select Attributes*.

Tabel 5. Statistik Select Attributes

No.	Uraian	Keterangan
1.	Record	4590
2.	Special Attributes	1
3.	Reguler Attributes	4
4.	Attributes:	
	cabang_pelayanan	Nominal, missing value 0
	Jenis_kendaraan	Nominal, missing value 0
	Fungsi_kendaraan	Nominal, missing value 0
	Jumlah_kendaraan	Integer, missing value 0

4.3. Data Transformation

Bisa dilihat dari data tersebut diantaranya jenis kendaraan dan fungsi kendaraan bukanlah numeric sehingga akan dilakukannya transformasi data seperti berikut.



Gambar 6. Operator Nominal to Numerical

Parameter yang digunakan pada saat *Nominal to Numerical* adalah sebagai Tabel 6. Tabel Parameter *Nominal to Numerical*

Tabel 6. Parameter Nominal to Numerical

No.	Uraian	Keterangan
1.	Attribute Filter Type	subset
2.	Selected attribute	Jenis_kendaraan, fungsi_kendaraan

Berikut hasil setelah melakukan *Nominal to Numerical* pada Tabel 7. Statistik *Nominal to Numerical*.

Tabel 7. Statistik Nominal to Numerical.

No.	Uraian	Keterangan
1.	Record	4590
2.	Special Attributes	1
3.	Reguler Attributes	4
4.	Attributes:	
	Cabang_pelayanan	Polynominal, missing value 0
	Jenis_kendaraan	Numeric, missing value 0
	Fungsi_kendaraan	Numeric, missing value 0
	Jumlah kendaraan	Integer, missing value 0

4.4. Data Mining

Pada tahap ini akan dilakukan untuk menentukan sebuah cluster mana yang akan digunakan dalam Daftar Kendaraan Bermotor tersebut.



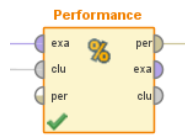
Gambar 7. Operator K-Means

K-Means clustering dilakukan sebanyak 9 kali dan dimulai dengan parameter k angka 2 sampai 10. Dan hasil pembacaan operator K-Means clustering dengan parameter default didapat informasi sebagai berikut pada Tabel 8. Hasil K-Means Clustering.

Tabel 8. Hasil K-Means Clustering

No.	Cluster	Measures_type	Main_Criterion	Centroid_distance
1.	2	bregmanDivergance	Davies_Bouldin	5391457188.758
2.	3	BregmanDivergance	Davies_Bouldin	1287880558.278
3.	4	BregmanDivergance	Davies_Bouldin	761127566.779
4.	5	BregmanDivergance	Davies_Bouldin	415808172.960
5.	6	BregmanDivergance	Davies_Bouldin	361252980.953
6.	7	BregmanDivergance	Davies_Bouldin	223952778.197
7.	8	BregmanDivergance	Davies_Bouldin	174826225.994
8.	9	BregmanDivergance	Davies_Bouldin	162835823.776
9.	10	BregmanDivergance	Davies_Bouldin	130304541.210

Selanjutnya menggunakan operator *Performance Cluster Distance* untuk mengetahui nilai *Davies-Bouldin Index (DBI)*.



Gambar 8. Operator Performance

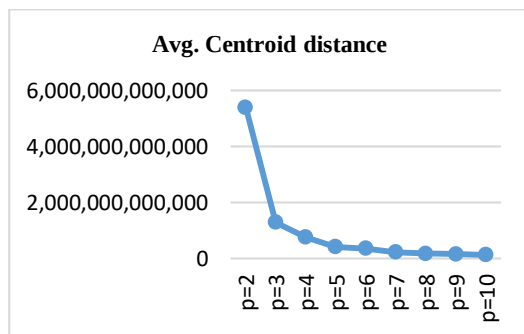
Berikut hasil dari operator *Performance* pada Tabel 9. Nilai DBI tiap *cluster*.

Tabel 9. Nilai DBI tiap cluster

No.	K	DBI
1.	2	0.475
2.	3	0.270
3.	4	0.429
4.	5	0.421
5.	6	0.461
6.	7	0.446
7.	8	0.473
8.	9	0.463
9.	10	0.455

4.5. Interpretation

Dari hasil yang didapat dalam pengoperasian RapidMiner. Penilaian optimasi cluster untuk dataset Daftar Kendaraan Bermotor pada Tabel 8. Hasil K-Means *Clustering* menggunakan metode Elbow. Untuk mendapatkan perbandingan sudut antar cluster yang membentuk siku. Hasil perhitungan untuk semua k uji ditunjukkan pada Gambar 9. Identifikasi titik *elbow*.



Gambar 9. Identifikasi titik elbow

Dan disimpulkan bahwa optimal untuk dataset Daftar Kendaraan Bermotor berdasarkan metode Elbow menunjukkan pada $k=3$ dengan grafik siku. Sedangkan optimasi cluster menggunakan Davies Bouldin Indeks dari sejumlah k uji ditunjukkan pada Tabel 9. Nilai DBI tiap *cluster*. Terlihat bahwa nilai DBI terendah ada di $k=3$ yaitu sebesar 1287880558.278. Sehingga menurut perhitungan Davies Bouldin Indeks yang merupakan cluster optimal adalah tiga cluster.

5. KESIMPULAN DAN SARAN

Dengan melihat hasil penelitian yang telah dilakukan, maka dapat ditarik kesimpulan. Hasil penelitian menunjukkan bahwa analisis *cluster* Daftar Kendaraan Bermotor di Provinsi Jawa Barat mengidentifikasi tiga *cluster* berbeda, memberikan wawasan tentang pola cabang pelayanan. *Cluster_0* kategori sebagai rendah dengan beragam kendaraan, *cluster_1* adalah cluster tinggi dengan fokus pada sepeda motor, dan *cluster_2* sebagai *cluster* sedang dengan kendaraan pribadi. *Cluster* tiga dianggap sebagai *cluster* optimal dengan nilai DBI 0,270 dan titik *centroid* 1.287.880.558,278. Konsistensi ini menunjukkan bahwa *cluster* tiga adalah pilihan terbaik dalam menganalisis Daftar Kendaraan Bermotor, menggambarkan karakteristik serupa di cabang pelayanan tersebut. Analisis *cluster* K-Means menghasilkan tiga *cluster* yang mewakili kategori cabang pelayanan (rendah, sedang, tinggi). *Cluster* "rendah" mencakup kendaraan dengan atribut rendah, sementara "sedang" dan "tinggi" menunjukkan tingkat klasifikasi yang baik. Hasil ini dapat menjadi dasar untuk pengambilan keputusan terkait pengelolaan infrastruktur transportasi dan pemenuhan kewajiban pajak. Saran untuk penelitian lebih lanjut mencakup proses *clustering* terhadap Kendaraan Tidak Melakukan Daftar Ulang (KTMDU) untuk membandingkan jumlah kendaraan dan tingkat kepatuhan pajak. Disarankan juga memusatkan fokus pada satu cabang pelayanan untuk analisis yang lebih mendalam. (Memuat kesimpulan yang diperoleh dan saran-saran untuk penelitian selanjutnya (jika ada). Tuliskan Kesimpulan dan saran dalam 1 paragraf.)

DAFTAR PUSTAKA

- [1] D. Efriyenty, "... Pajak Terhadap Kepatuhan Wajib Pajak Orang Pribadi Dalam Membayar Pajak Kendaraan Bermotor Di Kota Batam: Sanksi Perpajakan, Pemahaman perpajakan ...," *J. Akunt. Bareleng*, 2019, [Online]. Available: <http://ejournal.upbatam.ac.id/index.php/jab/article/view/1244>
- [2] S. Pentanurbowo, "Upaya Meningkatkan Potensi Pajak Kendaraan Bermotor Melalui Penelusuran Kendaraan Bermotor Tidak Melakukan Daftar Ulang (KTMDU) Di Kelurahan ...," *Fair Value: Jurnal Ilmiah Akuntansi dan* journal.ikopin.ac.id, 2021. [Online]. Available: <https://www.journal.ikopin.ac.id/index.php/fairvalue/article/download/739/343>
- [3] N. Tania, "PENGARUH JUMLAH KENDARAAN BERMOTOR DAN KENDARAAN TIDAK MELAKUKAN DAFTAR ULANG TERHADAP PENERIMAAN PAJAK KENDARAAN ...," *J. Proaksi*, 2022, [Online]. Available: <https://garuda.kemdikbud.go.id/documents/detail/3225927>
- [4] P. O. Puteri, E. Syofyan, and E. Mulyani, "Analisis pengaruh sanksi administrasi, tingkat

- pendapatan, dan sistem samsat drive thru terhadap kepatuhan wajib pajak kendaraan bermotor,” *J. Eksplor. Akunt.*, 2019, [Online]. Available: <http://jea.ppj.unp.ac.id/index.php/jea/article/view/163>
- [5] A. Winasari, “Pengaruh Pengetahuan, Kesadaran, Sanksi, Dan Sistem E-SAMSAT Terhadap Kepatuhan Wajib Pajak Kendaraan Bermotor Di Kabupaten Subang.(Studi Kasus ...,” *Prism. (Platform Ris. Mhs. Akuntansi)*, 2020, [Online]. Available: <http://ojs.stiesia.ac.id/index.php/prisma/article/view/362>
- [6] M. R. Alhapizi, M. Nasir, and I. Effendy, “Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Universitas Bina Darma Palembang,” *J. Softw. ...*, 2020, [Online]. Available: <https://journal-computing.org/index.php/journal-sea/article/view/10>
- [7] S. Marpaung and Z. Aidha, “Perilaku Manusia Silver Terhadap Keluhan Kesehatan di kecamatan Medan Halvetia,” *Heal. Inf. J. ...*, 2023, [Online]. Available: <https://myjurnal.poltekkes-kdi.ac.id/index.php/hijp/article/view/932>
- [8] I. Nozomi, “Penerapan Data Mining Untuk Peringatan Dini Banjir Menggunakan Metode Klastering K-Means (Studi Kasus Kota Padang),” *J. Sains Inform. Terap.*, 2023, [Online]. Available: <https://rcf-indonesia.org/jurnal/index.php/jsit/article/view/165>
- [9] A. Winarta and W. J. Kurniawan, “Optimasi cluster k-means menggunakan metode elbow pada data pengguna narkoba dengan pemrograman python,” *JTIK (Jurnal Tek. ...)*, 2021, [Online]. Available: <https://jurnal-backup.kaputama.ac.id/index.php/JTIK/article/view/466>
- [10] R. Y. Sari, H. Oktavianto, and ..., “Algoritma K-Means Dengan Metode Elbow Untuk Mengelompokkan Kabupaten/Kota Di Jawa Tengah Berdasarkan Komponen Pembentuk Indeks Pembangunan ...,” *J. Smart ...*, 2022, [Online]. Available: <http://jurnal.unmuhjember.ac.id/index.php/JST/article/view/6928>
- [11] M. Sholeh and K. Aeni, “Perbandingan Evaluasi Metode Davies Bouldin, Elbow dan Silhouette pada Model Clustering dengan Menggunakan Algoritma K-Means,” *STRING (Satuan Tulisan Ris. dan ...)*, 2023, [Online]. Available: <https://journal.lppmunindra.ac.id/index.php/STRING/article/view/16388>
- [12] L. Markusoski, I. Zdravkoski, and ..., “KNOWLEDGE DISCOVERY DATABASES (KDD) PROCESS IN DATA MINING,” *Int. ...*, 2019, [Online]. Available: <https://eprints.uklo.edu.mk/3082/1/Paper Ljupce Markusheski PhD.docx>
- [13] M. Orisa, “Optimasi Cluster pada Algoritma K-Means,” *Pros. SENIATI*, vol. 6, no. 2, pp. 430–437, 2022, doi: 10.36040/seniati.v6i2.5034.