

**KLASIFIKASI TINGKAT PENJUALAN MOBIL LISTRIK
DI INDONESIA MENGGUNAKAN
ALGORITMA *K-MEANS***



SKRIPSI

Diajukan untuk memenuhi salah satu syarat kelulusan Program Sarjana

FARHAN JIDDAN

NIM: 19200242

LISANNE

NIM: 19200828

MOEHAMAD PANDU

NIM: 19200253

Program Studi Sistem Informasi

Fakultas Teknik dan Informatika

Universitas Bina Sarana Informatika

Jakarta

2024

SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini:

Nama : Farhan Jiddan
NIM : 19200242
Jenjang : Sarjana
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa Tugas Akhir yang telah saya buat dengan judul: **“Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means”**, adalah asli (orsinil) atau tidak plagiat (menjiplak) dan belum pernah diterbitkan/dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun juga. Apabila di kemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa Tugas Akhir yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari **Universitas Bina Sarana Informatika** dicabut/dibatalkan.

Dibuat di : Jakarta
Pada tanggal : 02 Juli 2024

Yang menyatakan,

Anggota :

1. Lisanne

2. Moehamad Pandu


.....

.....




Farhan Jiddan

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya:

Nama : Farhan Jiddan
NIM : 19200242
Jenjang : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa seluruh data, informasi, interpretasi serta pernyataan yang terdapat dalam karya ilmiah Penulis dengan judul “**Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means**” ini, kecuali yang disebutkan sumbernya adalah hasil pengamatan, penelitian, pengelolaan, serta pemikiran saya.

Penulis menyetujui untuk memberikan izin kepada pihak **Universitas Bina Sarana Informatika** untuk mendokumentasikan karya ilmiah saya tersebut secara internal dan terbatas, serta tidak untuk mengunggah karya ilmiah Penulis pada repository Universitas Bina Sarana Informatika.

Penulis bersedia untuk bertanggung jawab secara pribadi, tanpa melibatkan pihak **Universitas Bina Sarana Informatika**, atas materi/isi karya ilmiah tersebut, termasuk bertanggung jawab atas dampak atau kerugian yang timbul dalam bentuk akibat tindakan yang berkaitan dengan data, informasi, interpretasi serta pernyataan yang terdapat pada karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 02 Juli 2024
Yang menyatakan,



Handwritten signature of Farhan Jiddan.

Farhan Jiddan

PERSETUJUAN DAN PENGESAHAN SKRIPSI

Skripsi ini diajukan oleh:

Nama : Farhan Jiddan
NIM : 19200242
Jenjang : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika
Judul Skripsi : Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means

Telah dipertahankan pada periode 2024-1 dihadapan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh Sarjana Komputer (S.Kom) pada Program Sarjana (S1) Program Studi Sistem Informasi di Universitas Bina Sarana Informatika.

Jakarta, 08 Agustus 2024

PEMBIMBING SKRIPSI

Pembimbing I : Agustiena Merdekawati, S.T., M.Kom.

Pembimbing II : Yuri Yuliani, M.Kom.

DEWAN PENGUJI UNIVERSITAS

Penguji I : Bibit Sudarsono, M.Kom.

Penguji II : Ita Dewi Sintawati, ST.,M.Kom



Handwritten signatures of the supervisors and examiners, each on a dotted line.

PEDOMAN PENGGUNAAN HAK CIPTA

Skripsi sarjana yang berjudul “**Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means**” adalah hasil karya tulis asli Farhan Jiddan dan bukan hasil terbitan sehingga peredaran karya tulis hanya berlaku dilingkungan akademik saja, serta memiliki hak cipta. Oleh karena itu, dilarang keras untuk menggandakan baik sebagian maupun seluruhnya karya tulis ini, tanpa seizin penulis.

Referensi kepustakaan diperkenankan untuk dicatat tetapi pengutipan atau peringkasan isi tulisan hanya dapat dilakukan dengan seizin penulis dan disertai ketentuan pengutipan secara ilmiah dengan menyebutkan sumbernya.

Untuk keperluan perizinan pada pemilik dapat menghubungi informasi yang tertera di bawah ini:



Nama : Farhan Jiddan
Alamat : Jl. Annur 2b no.91A Rt.002/010 Jakarta Timur
No. Telp : 083875246757
E-mail : farhanjiddan666@gmail.com

	LEMBAR KONSULTASI SKRIPSI
	UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19200242
 Nama Lengkap : Farhan Jiddan
 Asisten Pembimbing : Agustiena Merdekawati, ST, M.Kom
 Judul Skripsi : Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means

NO	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing
1	01 April 2024	Bimbingan perdana, pengajuan judul dan ACC judul	
2	18 April 2024	Pengajuan bab 1, pengajuan data & olah data yang akan diusulkan	
3	27 April 2024	Revisi bab 1 (Data yang diusulkan dan Roadmap penelitian)	
4	07 Mei 2024	ACC bab 1, dan pengajuan bab 2	
5	18 Mei 2024	Revisi bab 2 (Kerangka Penelitian)	
6	22 Mei 2024	ACC bab 2, dan pengajuan bab 3	
7	08 Juni 2024	Revisi bab 3 (koreksi perhitungan manual)	
8	22 Juni 2024	ACC bab 3, dan pengajuan bab 4	
9	27 Juni 2024	ACC bab 4 dan halaman depan dan belakang skripsi	
10	03 Juli 2024	ACC keseluruhan	

Catatan untuk Dosen Pembimbing.
Bimbingan Skripsi

- Dimulai pada tanggal : 01 April 2024
- Diakhiri pada tanggal : 03 Juli 2024
- Jumlah pertemuan bimbingan : 10

Disetujui oleh,
Dosen Pembimbing


 (Agustiena Merdekawati, ST, M.Kom)



LEMBAR KONSULTASI SKRIPSI

UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19200242
Nama Lengkap : Farhan Jiddan
Asisten Pembimbing : Yuri Yuliani, M.Kom
Judul Skripsi : Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means

NO	Tanggal Bimbingan	Pokok Bahasan	Paraf Asisten Pembimbing
1	23 April 2024	Pengajuan bab 1, data yang digunakan untuk penelitian	
2	29 April 2024	Revisi bab 1, teknik penulisan	
3	13 Mei 2024	ACC bab 1, dan pengajuan bab 2	
4	20 Mei 2024	ACC bab 2	
5	28 Mei 2024	Pengajuan bab 3, revisi struktur penulisan	
6	10 Juni 2024	ACC bab 3, dan pengajuan bab 4	
7	28 Juni 2024	ACC bab 4	
8	04 Juli 2024	ACC keseluruhan	

Catatan untuk Asisten Pembimbing.

Bimbingan Skripsi

- Dimulai pada tanggal : 23 April 2024
- Diakhiri pada tanggal : 04 Juli 2024
- Jumlah pertemuan bimbingan : 8

Disetujui oleh,
Asisten Pembimbing

(Yuri Yuliani, M.Kom)

PERSEMBAHAN

*"Barangsiapa yang menjaga Allah dalam hatinya,
maka Allah akan menjaga dirinya dan segala urusannya."
(Ibnu Qayyim Al-Jauziyyah)*

Dengan mengucapkan syukur kepada Allah S.W.T,
skripsi ini saya persembahkan untuk:

1. Ibu Hayani Terima kasih atas cinta, dukungan dan pengorbanan yang tiada henti. Ibu adalah pilar utama dalam hidup saya, yang selalu memberikan motivasi dan semangat untuk mencapai impian.
2. Kak Fithri, Kak Nina, Kak Fani Terima kasih telah menjadi saudari sekaligus teman terbaik yang selalu ada di saat suka dan duka. Kalian adalah inspirasi dan penguat hati saya.
3. Untuk semua dosen yang telah membimbing saya, Terima kasih atas ilmu, bimbingan dan keteladanan yang telah diberikan. Tanpa kalian, pencapaian ini tidak akan mungkin terjadi.

*Tanpa mereka, saya tidak akan pernah mencapai ini semua
Mereka adalah lentera yang menerangi jalan hidup saya*

UNIVERSITAS

KATA PENGANTAR

Alhamdulillah, Dengan mengucapkan puji syukur kehadiran Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya, sehingga pada akhirnya penulis dapat menyelesaikan tugas ini dengan baik Skripsi pada Program Sarjanaini penulis sajikan dalam bentuk buku yang sederhana Adapun judul Skripsi yang penulis ambil sebagai berikut, “**Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma *K-Means***”.

Tujuan penulisan Skripsi ini dibuat sebagai salah satu syarat kelulusan Sarjana Universitas Bina Sarana Informatika. Sebagai bahan penulisan diambil berdasarkan hasil penelitian (eksperimen), observasi dan beberapa sumber literatur yang mendukung penulisan ini. Penulis menyadari bahwa tanpa bimbingan dan dorongan dari semua pihak, maka penulisan Skripsi ini tidak akan berjalan lancar. Oleh karena itu pada kesempatan ini, ijinilah penulis menyampaikan ucapan terima kasih kepada:

1. Rektor Universitas Bina Sarana Informatika.
2. Dekan Fakultas Teknik dan Informatika
3. Ketua Program Studi Sistem Informasi Universitas Bina Sarana Informatika.
4. selaku Dosen Pembimbing Skripsi
5. selaku Asisten Pembimbing Skripsi
6. Staff / karyawan / dosen di lingkungan Universitas Bina Sarana Informatika.
7. Orang tua tercinta yang telah memberikan dukungan moral maupun spiritual.
8. Rekan-rekan mahasiswa kelas SI-8A.

Serta semua pihak yang terlalu banyak untuk disebut satu persatu sehingga terwujudnya penulisan ini. Penulis menyadari bahwa penulisan Skripsi. ini masih jauh sekali dari sempurna, untuk itu penulis mohon kritik dan saran yang bersifat membangun demi kesempurnaan penulisan di masa yang akan datang.

Akhir kata semoga Skripsi ini dapat berguna bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 2 Juli 2023

Penulis

ABSTRAK

**Farhan Jiddan (19200242), Lisanne (19200828), Moehamad Pandu (19200253),
Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan
Algoritma *K-means***

Mobil listrik adalah salah satu perkembangan terkini yang menonjol. Mobil ini menggunakan motor listrik sebagai penggerak utama dan menggunakan baterai sebagai sumber energi. Permintaan akan barang berkualitas tinggi mendorong perusahaan untuk meningkatkan kualitas produk dan mempertahankan reputasi merek mereka. Dengan begitu Perusahaan dapat mengidentifikasi minat pelanggan dalam membeli mobil listrik dengan melihat statistik penjualannya. *K-Means* adalah untuk membagi data saat ini ke dalam satu atau lebih kelompok atau *cluster*. Adapun metode pengumpulan data yang digunakan pada penelitian ini adalah dengan mengambil data sekunder data web Gaikindo, Metode *Clustering* ini dapat digunakan untuk membantu pihak perusahaan-perusahaan otomotif khususnya penjualan mobil listrik dalam melihat tingkat penjualan mobil mereka dengan perusahaan-perusahaan lain. Berdasarkan hasil pengujian dari cara manual dan dengan *software* menggunakan rapidminer v10.3.001 mendapatkan hasil yang sama dan tepat Yaitu *Cluster 0* jumlah data brand sebanyak 12 dengan presentase 80% (Cukup Laris), *cluster 1* dengan jumlah 1 data dan presentase 6,7% (Sangat Laris), *cluster 2* jumlah 1 data dengan presentase 6,7% (Kurang Laris) dan *cluster 3* jumlah data 1 dengan presentase 6,7% (Laris). Untuk hasil perhitungan menggunakan algoritma *K-means* nilai DBI (*Davies Bouldin Index*) pada penelitian ini menghasilkan nilai sebesar 0,099 Hasil *Davies-Bouldin Index* (DBI) dalam mengukur kualitas *clustering* menghasilkan nilai sebesar 0,099. Yang artinya menunjukkan bahwa hasil *clustering* yang dilakukan sangat baik dan juga bisa memanfaatkan data eksternal seperti tren ekonomi, harga bahan bakar, kebijakan pemerintah terkait kendaraan listrik dan preferensi konsumen untuk memperkaya analisis dan memberikan insight yang lebih mendalam.

Kata Kunci : Mobil Listrik, *K-Means*, *Clustering*

DAFTAR ISI

PEDOMAN PENGGUNAAN HAK CIPTA.....	v
PERSEMBAHAN.....	viii
KATA PENGANTAR.....	ix
ABSTRAK.....	x
DAFTAR ISI.....	xi
PENDAHULUAN.....	2
1.1 Latar Belakang Masalah.....	2
1.2 Tujuan Penelitian.....	4
1.3 Penelitian Terdahulu.....	4
1.4 <i>Roadmap</i> dan Pembaruan.....	6
BAB II.....	8
METODOLOGI PENELITIAN.....	8
2.1 Kerangka Penelitian atau Tahap Penelitian.....	8
2.2 Instrumen Penelitian.....	10
2.2.1 Dokumentasi.....	11
2.2.2 Observasi.....	12
2.3 Metode Pengumpulan Data, Populasi dan Sampel.....	12
2.4 Metode Analisa Data.....	13
BAB III.....	14
HASIL DAN PEMBAHASAN.....	14
3.1 Perumusan Masalah dan Studi Litelatur.....	14
3.2 Pencarian Data.....	14
3.3 Pemilihan Data.....	15
3.4 Pengolahan Data Awal (Preprocessing).....	15
3.5 Perhitungan <i>Data Mining</i> Dengan Manual.....	17
A. Hitung Manual.....	17
B. Implementasi rapid miner.....	28
3.6 Evaluasi Hasil.....	31
3.7 Kesimpulan.....	40
BAB IV.....	43
PENUTUP.....	43
4.1 Kesimpulan.....	43
4.2 Saran.....	45
BAB V.....	46

REFERENSI	46
DAFTAR RIWAYAT HIDUP	49
LEMBAR KONSULTASI BIMBINGAN.....	50
SURAT KETERANGAN RISET	52
BUKTI HASIL PENGECEKAN PLAGIARISME	53

DAFTAR GAMBAR

Gambar I. 1 Roadmap dan pembaruan.....	7
Gambar II. 1 Kerangka Penelitian atau Tahap Penelitian	8
Gambar III. 1 Website Gaikindo.....	15
Gambar III. 2 Tampilan Awal RapidMiner.....	28
Gambar III. 3 Desain Operator Algoritma K-means.....	29
Gambar III. 4 Tampilan pencarian lokasi file	29
Gambar III. 5 Tampilan pemilihan sheet	30
Gambar III. 6 Tahap pemberian anotasi id	30
Gambar III. 7 Pemberian anotasi id	31
Gambar III. 8 Hasil Cluster Model	31
Gambar III. 9 Hasil Cluster Model	32
Gambar III. 10 Hasil ExampleSet (Clustering).....	32
Gambar III. 11 Hasil Performance (Clustering).....	34
Gambar III. 12 Penyebaran data Bulan dari X1 Sampai X12	34
Gambar III. 13 Hasil Davies Bouldin Index	35

DAFTAR TABEL

Tabel I. 1 Penelitian terdahulu	5
Tabel III. 1 Data yang telah diseleksi	16
Tabel III. 2 Data Normalisasi Max Min	18
Tabel III. 3 Cluster Awal	19
Tabel III. 4 Hasil Data Ke Cluster Pada Iterasi ke 1	20
Tabel III. 5 Hasil Centroid Baru	22
Tabel III. 6 Hasil Centroid Pada Iterasi ke 2.....	22
Tabel III. 7 Hasil Data Ke Cluster Pada Iterasi ke 2 (Terakhir)	22
Tabel III. 8 Titik Pusat Akhir Dari Proses Clustering.....	24
Tabel III. 9 Hasil Jumlah Cluster	35
Tabel III. 10 Hasil Clustering secara manual dan menggunakan rapidminer	37
Tabel III. 11 Analisis Penjualan Mobil Listrik	39
Tabel III. 12 Hasil ramalan bulan Januari hingga Desember 2024.....	40
Tabel IV. 1 Pengelompokkan Cluster Berdasarkan Brand	44

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Di era modern yang dipenuhi dengan kemajuan teknologi, dampak kemajuan ini sangat terasa di berbagai aspek kehidupan manusia, termasuk dalam industri otomotif. Mobil listrik merupakan salah satu perkembangan terkini yang menonjol. Mobil ini menggunakan motor listrik sebagai penggerak utama dan menggunakan baterai sebagai sumber energi. Mobil listrik memiliki banyak keuntungan, termasuk efisiensi energi dan efek positif terhadap lingkungan. Karena setiap merek memiliki karakteristik tertentu yang membedakan produknya dari yang lain, merek tersebut dapat mempengaruhi minat pembeli. Permintaan akan barang berkualitas tinggi mendorong perusahaan untuk meningkatkan kualitas produk dan mempertahankan reputasi merek mereka. Dengan begitu Perusahaan dapat mengidentifikasi minat pelanggan dalam membeli mobil listrik dengan melihat statistik penjualannya.

Penjualan mobil listrik, juga dikenal sebagai mobil bertenaga listrik (BEV), di Indonesia mencatatkan angka sebanyak 17.062 unit pada periode Januari hingga Desember 2023, melonjak sebesar 65,22% pertahun dari 10.327 unit pada tahun sebelumnya. Data ini dikumpulkan oleh Gabungan Industri Kendaraan Bermotor Indonesia (Gaikindo) (Dimas & Kartika, 2024). Menurut data terbaru dari Gabungan Industri Kendaraan Bermotor Indonesia (Gaikindo), penjualan mobil listrik mencapai 13.873 unit dari Januari hingga November 2023, naik dari 10.183 unit pada periode yang sama tahun sebelumnya. Namun, penjualan ini masih hanya 1,50 persen dari total penjualan mobil yang mencapai 920.518 unit (Benediktus, 2024). Hal ini menunjukkan peningkatan signifikan dalam penggunaan kendaraan listrik di

Indonesia. Ini menunjukkan kesadaran yang semakin meningkat terhadap pentingnya penggunaan kendaraan listrik dan kemajuan dalam infrastruktur yang mendukung kendaraan ramah lingkungan.

Ketidakjelasan tentang seberapa besar atau kecilnya pembelian mobil listrik di Indonesia disebabkan oleh kurangnya informasi yang jelas tentang waktu pembelian dan jumlah mobil listrik yang dibeli. Sulit untuk mendapatkan pemahaman menyeluruh tentang tren pembelian mobil listrik ramah lingkungan di pasar otomotif Indonesia jika tidak ada informasi yang jelas tentang kapan dan seberapa banyak yang dibeli. Memahami dinamika pasar dan membuat keputusan yang dibuat oleh konsumen tentang pembelian mobil listrik menjadi sulit jika kita tidak memberikan konteks yang cukup tentang faktor-faktor yang memengaruhi pembelian mobil listrik, seperti insentif pemerintah. Misalnya, insentif pemerintah dapat berdampak besar pada harga dan daya tarik mobil listrik dibandingkan kendaraan konvensional. Di sisi lain, ketersediaan infrastruktur pengisian daya yang memadai adalah komponen penting dalam membuat keputusan pembelian.

Permintaan akan barang berkualitas tinggi semakin meningkat, yang mendorong perusahaan untuk berusaha lebih baik untuk meningkatkan kualitas produk mereka dan mempertahankan reputasi brand mereka. Setiap perusahaan, terutama dalam industri otomotif, terus bersaing untuk meningkatkan penjualan mobil kepada pelanggan mereka dengan menggunakan citra brand dan kualitas produksinya. Untuk melakukan ini, perusahaan dapat menggunakan analisis grafik penjualan untuk memahami faktor-faktor yang mempengaruhi minat pelanggan dalam membeli mobil (Hasyim et al., 2022).

Konsumen mungkin ragu untuk menggunakan mobil listrik karena perlu mengisi daya lebih sering atau khawatir akan kehabisan daya di tengah perjalanan (Anshor & Rahardjo, 2023).

Dengan teknologi yang sebanding, mobil listrik mungkin lebih mahal pada awalnya daripada mobil berbahan bakar minyak (Anshor & Rahardjo, 2023).

Mobil listrik memiliki beberapa kekurangan, termasuk biaya tinggi, jarak tempuh yang terbatas dan waktu pengisian baterai yang lama (Adha, 2020).

1.2 Tujuan Penelitian

Berdasarkan latar belakang masalah yang telah jabarkan, tujuan penelitian ini adalah sebagai berikut :

1. Mengetahui hasil pengelompokan tingkat pembelian mobil listrik di Indonesia dengan metode *K-means*.
2. Bagaimana hasil kualitas pengelompokan atau *clustering* yang didapatkan?
3. Mengidentifikasi musim atau periode tertentu dalam setahun di mana penjualan mobil dari brand tertentu cenderung tinggi atau rendah.

1.3 Penelitian Terdahulu

Penelitian Terdahulu dapat memberikan dampak signifikan dalam ilmu pengetahuan. Oleh karena itu, untuk memperluas referensi, beberapa penelitian sebelumnya diperlukan untuk mengumpulkan literasi.

Dalam penelitian ini literasi yang digunakan adalah jurnal dengan metode *K-means*, metode itu berfungsi untuk mengelompokan data mobil listrik dari situs Gaikindo yang diharapkan dapat membantu penelitian ini. Berikut adalah penelitian terdahulu yang dijabarkan dalam tabel dibawah ini.

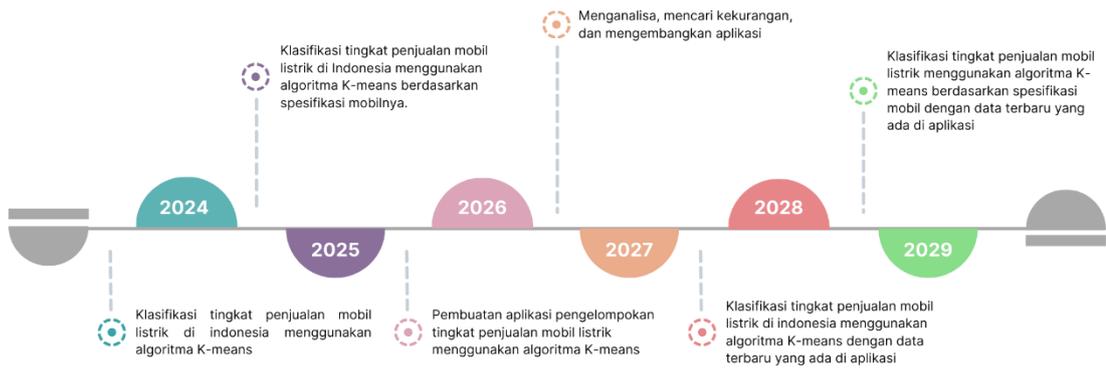
Tabel I. 1
 Penelitian terdahulu

No	Judul Jurnal dan Peneliti	Tahun dan Tempat Penelitian	Metode Penelitian	Objek Penelitian	Hasil Penelitian
1	<p>Penerapan <i>Data Mining</i> Untuk Prediksi Penjualan Mobil Menggunakan Metode <i>K-Means Clustering</i></p> <p>Peneliti :</p> <ul style="list-style-type: none"> - Sufajar Butsianto - Nindi Tya Mayangwulan 	2020, Indonesia	<i>K-means</i>	<p>GAIKINDO sebanyak 900 data hasil penjualan mobil per bulan selama 5 tahun</p>	<p>Setelah menggunakan metode <i>K-Means Clustering</i>, tiga <i>cluster</i> terbentuk. <i>Cluster 0</i> memiliki 235 anggota dan memiliki presentase 26% dikategorikan Laris; <i>Cluster 1</i> memiliki 604 anggota dan memiliki presentase 67% dikategorikan Kurang Laris; dan <i>Cluster 2</i> memiliki 61 anggota dan memiliki presentase 7% dikategorikan Paling Laris.</p>
2	<p>Klasifikasi Data Penjualan Menggunakan Algoritma <i>K-Means</i> Dan <i>Analytic Hierarchy Process</i></p> <p>Peneliti :</p>	2022, Indonesia	<i>K-means</i>	<p>Data 10 swalayan di wilayah Jakarta selama tahun 2020 dan tahun 2021</p>	<p>Menggunakan algoritma <i>K-Means</i> dengan k2 pada literasi ke 3 dengan rasio 0,04926 menunjukkan hasil terbaik.</p>

	<ul style="list-style-type: none"> - Bibit Sudarsono - Umi Faddillah - Ayuni Asistiyasari - Yosep Nuryaman 				
3	<p>Implementasi <i>Data Mining</i> Menggunakan Algoritma <i>K- Means Clustering</i> Dalam Analisis Penjualan Produk</p> <p>Peneliti :</p> <ul style="list-style-type: none"> - Talitha Tiara Alifa - Rini Astuti - Fadhil Muhamad Basysyar 	2024, Indonesia	<i>K-means</i>	Data yang menjadi fokus penelitian adalah data stok fashion pria dari bulan Januari hingga April 2022	Sesuai dengan perhitungan manual, masing-masing klaster memiliki anggota kelompok klaster. Misalnya, <i>Cluster_0</i> di RapidMiner memiliki 207 produk fashion pria yang laris, <i>Cluster_1</i> memiliki 210 produk fashion pria yang biasa dan <i>Cluster_3</i> memiliki 83 produk fashion pria yang kurang laris.

1.4 Roadmap dan Pembaruan

Penelitian ini akan mengklasifikasikan tingkat pembelian mobil listrik di Indonesia dengan menggunakan algoritma *K-means* dan juga bertujuan untuk mendapatkan hasil tingkat akurasi dari penggunaan algoritma *K-means*. Di sisi lain, penelitian ini dapat memperoleh data tentang tingkat pembelian mobil listrik di Indonesia dari hasil algoritma *K-means*. Berikut adalah *roadmap* 5 tahun kedepan berdasarkan penelitian ini :



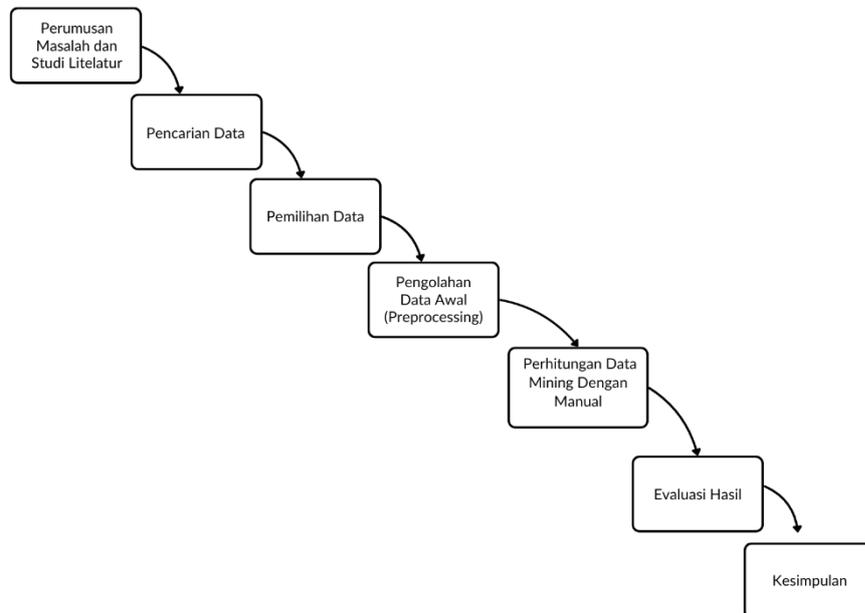
Gambar I. 1
Roadmap dan pembaruan

BAB II

METODOLOGI PENELITIAN

2.1 Kerangka Penelitian atau Tahap Penelitian

Kerangka penelitian merupakan konseptualisasi yang terdiri dari sejumlah konsep, teori, metode dan pendekatan yang digunakan, kerangka penelitian berfungsi sebagai struktur komprehensif yang membahas landasan teoritis dan metodologis penelitian serta mengarahkan langkah-langkah praktis yang diperlukan untuk menyusun, melaksanakan dan menganalisis data untuk mencapai tujuan penelitian.



Sumber : Penelitian

Gambar II. 1
Kerangka Penelitian atau Tahap Penelitian

Berdasarkan Kerangka Penelitian atau Tahap Penelitian diatas, maka masing-masing uraiannya adalah sebagai berikut :

1. Perumusan Masalah dan Studi Litelatur

Pada tahap ini dilakukan proses mendefinisikan lalu dilakukan proses pengumpulan, membaca dan analisis literatur yang berkaitan dengan subjek penelitian.

Pada penelitian ini perumusan masalahnya sebagai berikut :

- a. Bagaimana pengelompokan tingkat pembelian mobil listrik di Indonesia dengan Metode *K-means*?
- b. Bagaimana hasil kualitas pengelompokkan atau *clustering* yang didapatkan?
- c. Kapan musim atau periode tertentu dalam setahun di mana penjualan mobil dari brand tertentu cenderung tinggi atau rendah?

2. Pencarian Data

Dalam tahap ini dilakukan proses pencarian data penjualan mobil listrik pada *web* Gaikindo.

3. Pemilihan Data

Pada tahap ini dilakukan proses pemilihan data yang diambil dari *website*, sehingga data yang diperoleh relevan dan akurat untuk mendukung tujuan penelitian.

4. Pengolahan Data Awal (*Preprocessing*)

Pada tahap preprocessing terdapat 4 proses yang dapat dilakukan yaitu :

- a. Data *Understanding*
- b. Data *Selection*
- c. Data *Cleaning*
- d. Data *Transformation*

Pada tahap ini dilakukan seleksi data untuk mengambil data yang sesuai untuk kebutuhan penelitian.

5. Perhitungan *Data Mining* Dengan Manual

Pada tahap ini dilakukan implementasi *data mining* dengan menggunakan *K-means clustering* terhadap data yang diambil dari *website*, sehingga permasalahan pada penelitian ini dapat terselesaikan dan dianalisis secara mendalam dan juga dilakukan perhitungannya menggunakan excel.

6. Evaluasi Hasil

Pada tahap ini dilakukan evaluasi hasil terhadap data yang diambil dari *website*, sehingga keakuratan dan efektivitas temuan penelitian dapat dinilai dengan cermat.

7. Kesimpulan

Pada tahap ini, penarikan kesimpulan dilakukan berdasarkan analisis data yang diambil dari *website*, sehingga hasil penelitian dapat dirumuskan dan dipahami secara menyeluruh.

2.2 Instrumen Penelitian

Pada dasarnya, instrumen penelitian adalah alat yang digunakan untuk mengumpulkan data penelitian. Teori yang digunakan sebagai dasar instrumen penelitian dan tujuan pengukuran disesuaikan. Instrumen penelitian dibuat untuk satu tujuan penelitian dan tidak dapat digunakan untuk tujuan penelitian lain. Akibatnya, peneliti harus merancang instrumen mereka sendiri. Karena tujuan dan mekanisme kerja yang berbeda dari setiap teknik penelitian, susunan instrumen dari setiap penelitian tidak selalu sama dengan susunan instrumen dari penelitian lainnya. Data yang dikumpulkan melalui instrumen tertentu akan dijelaskan, dilampirkan, atau digunakan untuk menguji hipotesis penelitian. Selain membuat instrumen sendiri

untuk mengumpulkan data dalam suatu penelitian, instrumen yang telah tersedia, juga dikenal sebagai instrumen baku, dapat digunakan secara langsung untuk mengumpulkan data variabel penelitian. Ini karena instrumen tersebut telah melalui serangkaian uji kualitas dan layak digunakan untuk mengumpulkan data variabel tertentu. sesuai dengan teori yang digunakan sebagai dasar dan struktur variabel yang akan diukur dalam sebuah penelitian (Widodo et al., 2023).

Instrumen penelitian memainkan peran penting dalam proses penelitian karena mereka berfungsi sebagai alat untuk mengumpulkan data yang diperlukan oleh penelitian. Dengan menggunakan instrumen penelitian, peneliti akan mengetahui sumber daya data yang akan diteliti, jenis data yang akan dikumpulkan, metode pengumpulan data, instrumen yang digunakan dan langkah-langkah yang diambil untuk menyusun instrumen penelitian. Dengan adanya instrumen penelitian maka peneliti juga akan mengetahui validitas, reliabilitas, tingkat kesukaran daya pembeda dan tingkat kesalahan daya pembeda (Widodo et al., 2023).

Instrumen penelitian yang digunakan dalam penelitian ini adalah dokumentasi, dengan cara mencari data terkait penjualan mobil pada *web* Gaikindo lalu memilah data tersebut hingga menyisakan data penjualan mobil listrik saja.

2.2.1 Dokumentasi

Istilah "Dokumentasi Instrumen Penelitian" mengacu pada informasi tertulis. Peneliti dapat menggunakan instrumen ini untuk mengumpulkan data dengan menyelidiki objek tertulis seperti buku, majalah, catatan harian, artefak, video dan sebagainya. instrumen ini biasanya digunakan dalam penelitian seperti bukti sejarah, landasan hukum undang-undang, dll (Widodo et al., 2023).

2.2.2 Observasi

Kemampuan manusia untuk menggunakan semua panca inderanya dan menggunakan fungsi mata, salah satu panca indera utama, untuk mengumpulkan data atau informasi dikenal sebagai observasi (M. Makbul, 2021).

2.3 Metode Pengumpulan Data, Populasi dan Sampel

1. Metode Pengumpulan Data

Peneliti dapat mengumpulkan data menggunakan berbagai metode. Metode pengumpulan data dapat dilakukan secara terpisah dari metode analisis data, atau bahkan dianggap sebagai bagian penting dari analisis data itu sendiri. Data yang dikumpulkan dalam penelitian akan digunakan untuk menguji hipotesis atau menjawab pertanyaan yang diajukan dalam rumusan masalah dan kemudian akan dijadikan dasar untuk membuat kesimpulan atau keputusan (M. Makbul, 2021).

Data sekunder merupakan data yang dapat diperoleh atau dikumpulkan dari berbagai sumber yang sudah ada. Data sekunder dapat dikategorikan menjadi dua yaitu :

- a. Data internal adalah data yang berasal dari lingkungan sendiri, seperti data dari penelitian sebelumnya atau data rumah sakit (*research design*).
- b. Data yang berasal dari sumber luar, seperti publikasi, instansi dan lembaga ilmiah, disebut data eksternal (Widodo et al., 2023).

Adapun metode pengumpulan data yang digunakan pada penelitian ini adalah dengan mengambil data sekunder data *web* Gaikindo.

2. Populasi dan Sampel

Populasi dapat didefinisikan sebagai keseluruhan bagian dari penelitian,

termasuk subjek dan objek yang memiliki karakteristik tertentu. Oleh karena itu, populasi dapat mencakup semua anggota kelompok, baik itu manusia, binatang, peristiwa, atau benda yang tinggal bersama di suatu tempat secara terencana (Amin et al., 2023).

Sumber data penelitian adalah sampel, yang merupakan bagian dari populasi. Dengan kata lain, sampel adalah sebagian dari populasi untuk menunjukkan populasi secara keseluruhan (Amin et al., 2023).

Adapun data yang digunakan adalah data penjualan mobil listrik dari tahun 2019 hingga tahun 2023, dengan isi 15 brand yang diantaranya adalah BMW, Honda, Wuling, Toyota, Seres, Nissan, Neta, Morris Garage, Mitsubishi Motors, Mini, Mercedes Benz, Mazda, Lexus, Kia dan Hyundai.

2.4 Metode Analisa Data

Analisis data adalah proses mencari dan menyusun data dari catatan lapangan, dokumentasi dan wawancara (Waruwu, 2023).

Metode Analisa data yang digunakan pada penelitian ini adalah dengan menggunakan algoritma *K-means*.

Tujuan dari metode *clustering* data non-hierarki yang dikenal sebagai *K-Means* adalah untuk membagi data saat ini ke dalam satu atau lebih kelompok atau *cluster*. Metode ini memungkinkan untuk menggabungkan kelompok data dengan atribut yang serupa ke dalam kelompok yang sama dan kelompok data dengan atribut yang berbeda ke dalam kelompok yang berbeda (Nuryaman et al., 2022).

BAB III

HASIL DAN PEMBAHASAN

3.1 Perumusan Masalah dan Studi Litelatur

Untuk mengarahkan fokus penelitian dan menetapkan tujuan yang jelas, perumusan masalah adalah langkah awal dalam proses penelitian atau penyelesaian proyek. Langkah ini melibatkan identifikasi, definisi dan penyempurnaan masalah yang akan dipecahkan.

Pada penelitian ini perumusan masalahnya sebagai berikut :

- a. Bagaimana pengelompokan tingkat pembelian mobil listrik di Indonesia dengan Metode *K-means*?
- b. Bagaimana hasil kualitas pengelompokkan atau *clustering* yang didapatkan?
- c. Kapan musim atau periode tertentu dalam setahun di mana penjualan mobil dari brand tertentu cenderung tinggi atau rendah?

Pengumpulan, peninjauan dan analisis berbagai sumber literatur yang relevan dengan topik penelitian dikenal sebagai studi literatur. Tujuan dari studi ini adalah untuk mendapatkan pemahaman yang lebih baik tentang apa yang sudah diketahui tentang topik penelitian, menemukan kelemahan dalam penelitian dan memberikan konteks untuk penelitian yang akan datang. Pada penelitian ini studi litelatur yang dilakukan dengan membaca berbagai sumber baik jurnal, artikel dan buku.

3.2 Pencarian Data

Sebelum dilakukan pemilahan data dan pemrosesan data untuk mengetahui hasil yang diinginkan, diperlukan pencarian data yang sesuai dengan tujuan penelitian,

data yang dibutuhkan pada penelitian ini merupakan data penjualan mobil listrik dari tahun 2019 sampai dengan 2023, oleh karena itu peneliti melakukan pencarian data melalui internet.

3.3 Pemilihan Data

Setelah pencarian data dari berbagai *website* di internet ditemukan data yang cocok untuk penelitian ini yaitu data dari *website* Gaikindo <https://www.gaikindo.or.id/>, data yang diambil merupakan data penjualan mobil listrik dari tahun 2019 hingga tahun 2023.



Sumber : <https://www.gaikindo.or.id/>

Gambar III. 1
Website Gaikindo

3.4 Pengolahan Data Awal (Preprocessing)

1. Data Understanding

Ada beberapa atribut (bulan), bulan ini menyatakan penjualan perbulan selama setahun dengan jumlah atribut sebanyak 12 dan 15 brand mobil tersebut adalah Lexus, BMW, Honda, Wuling, Toyota, Seres, Nissan, Neta, Morris Garage, Mitsubishi Motors, Mini, Mercedes-Benz PC, Mazda, Kia, Hyundai. Terdapat juga 3 kategori mobil dengan kategori mobil sedan, 4x2 dan 4x4, selain

itu ada 7 tipe cc mobil yaitu, CC < 1.500 [G/D], CC 1.501 - 3.000 [G], CC > 3.001 [G] 2.501 [D], CC 1.501 - 2.501 [G/D], CC 1.501 - 2.500 [G/D], CC 1.501 - 3.000 [G] / 2.500 [D], CC > 3.001 [G], terdapat juga 8 jenis *origin country* yang terdiri dari Jerman, Korea, Jepang, Indonesia, Thailand, India, China, United Kingdom.

2. Data Selection

Proses memilih data dengan menghapus atribut yang tidak diperlukan dan tidak ada hubungannya dengan penelitian, data yang dihapus pada penelitian ini yaitu data kategori mobil, data tipe cc mobil dan data origin country. Atribut yang digunakan yaitu atribut brand dan atribut bulan.

3. Data Cleaning

Proses penghapusan data yang merupakan data duplikasi, data yang tidak inkonsisten, pada penelitian ini data yang dihapus adalah data dari brand mobil yang terduplikasi, untuk data penjualan mobil, angka penjualan yang sama tidak mengalami penghapusan dikarenakan penjualan bisa saja mengalami penurunan serta kenaikan penjualan.

4. Data Transformation

Proses akhir data yang sudah siap untuk diolah dengan cara manual atau menggunakan rapidminer untuk mendapatkan hasil yang diinginkan, berikut adalah data yang sudah diseleksi.

Tabel III. 1
Data yang telah diseleksi

BRAND	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
LEXUS	38	90	172	241	173	272	155	302	287	265	269	389
BMW	0	1	31	71	74	44	103	42	140	71	135	100
HONDA	0	0	0	0	0	0	0	0	0	490	488	549
WULING	39	147	457	460	469	218	299	1303	2741	2507	2479	4638
TOYOTA	208	1012	641	459	487	509	591	670	519	485	616	828
SERES	0	0	0	0	0	0	0	0	0	0	67	44
NISSAN	163	107	92	110	124	121	64	135	185	195	117	118
NETA	0	0	0	0	0	0	0	0	0	0	100	81
MORRIS GARAGE	0	0	2	0	0	0	36	16	25	4	17	1
MITSUBISHI MOTORS	2	5	10	13	11	9	1	0	1	2	14	6
MINI	0	0	17	6	14	2	26	26	31	10	29	56
MERCEDES-BENZ PC	1	3	10	21	17	13	5	29	23	34	8	7
MAZDA	0	0	0	0	0	0	26	23	23	28	15	10
KIA	0	0	12	8	2	0	19	1	1	7	8	57
HYUNDAI	339	342	666	904	1172	1055	525	765	1095	1423	1110	818

3.5 Perhitungan *Data Mining* Dengan Manual

A. Hitung Manual

1. Normalisasi Data

Data pada tabel III. 1 belum bisa secara langsung dilakukan pemrosesan dengan algoritma *K-Means*. Karena terdapat besaran angka yang terlalu jauh antara data satu dengan data yang lainnya. Perbedaan jarak atau besaran angka yang terlalu jauh ini dapat menyulitkan dalam proses pengelompokkan. Sehingga solusi yang dapat digunakan adalah dengan memperkecil besaran nilai penjualan tiap bulannya. Metode yang digunakan dalam penelitian ini adalah transformasi data menggunakan normalisasi. Yaitu Min-Max Normalization : Normalisasi min-max merupakan metode normalisasi yang mengubah nilai pada data secara linear, yaitu mengubah nilai maksimum menjadi 1 dan nilai minimum menjadi 0. Normalisasi min-max dapat mengurangi nilai-nilai yang terlalu besar pada fitur yang memiliki pengaruh terbesar pada proses klasifikasi. Min-Max normalization merupakan metode normalisasi dengan melakukan transformasi linier terhadap data asli sehingga menghasilkan keseimbangan nilai perbandingan antar data saat sebelum dan sesudah proses. (Reynaldo et al., 2020) Metode ini dapat menggunakan rumus sebagai berikut:

$$d' = \frac{d - \min(p)}{\max(p) - \min(p)}$$

Sumber : (Reynaldo et al., 2020)

Keterangan:

d' = Nilai hasil normalisasi

d = Nilai sebelum normalisasi

p = Keseluruhan nilai pada suatu atribut

$\max(p)$ = Nilai maksimum pada suatu atribut

min(p) = Nilai minimum pada suatu atribut

Sebagai contoh :

Berdasarkan Tabel III.1 jumlah minimal penjualan, yaitu 0. Dengan maksimum penjualan yaitu 4638. Nilai awal data pada brand Lexus, pada bulan Januari adalah 38. Sehingga perhitungannya adalah :

$$Normalisasi = \frac{(38 - 0)}{(4638 - 0)} = 0,00819 = 0,008$$

Nilai data pada brand Lexus dibulan Februari adalah 90. Sehingga perhitungannya adalah :

$$Normalisasi = \frac{(90 - 0)}{(4638 - 0)} = 0,01940 = 0,019$$

Berikut hasil keseluruhan Normalisasi data dari Tabel III.1 adalah sebagai berikut :

Tabel III. 2
Data Normalisasi Max Min

Brand	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
LEXUS	0,008	0,019	0,037	0,052	0,037	0,059	0,033	0,065	0,062	0,057	0,058	0,084
BMW	0,000	0,000	0,007	0,015	0,016	0,009	0,022	0,009	0,030	0,015	0,029	0,022
HONDA	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,106	0,105	0,118
WULING	0,008	0,032	0,099	0,099	0,101	0,047	0,064	0,281	0,591	0,541	0,534	1,000
TOYOTA	0,045	0,218	0,138	0,099	0,105	0,110	0,127	0,144	0,112	0,105	0,133	0,179
SERES	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,014	0,009
NISSAN	0,035	0,023	0,020	0,024	0,027	0,026	0,014	0,029	0,040	0,042	0,025	0,025
NETA	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,022	0,017
MORRIS GARAGE	0,000	0,000	0,000	0,000	0,000	0,000	0,008	0,003	0,005	0,001	0,004	0,000
MITSUBISHI MOTORS	0,000	0,001	0,002	0,003	0,002	0,002	0,000	0,000	0,000	0,000	0,003	0,001
MINI	0,000	0,000	0,004	0,001	0,003	0,000	0,006	0,006	0,007	0,002	0,006	0,012
MERCEDES-BENZ PC	0,000	0,001	0,002	0,005	0,004	0,003	0,001	0,006	0,005	0,007	0,002	0,002
MAZDA	0,000	0,000	0,000	0,000	0,000	0,000	0,006	0,005	0,005	0,006	0,003	0,002
KIA	0,000	0,000	0,003	0,002	0,000	0,000	0,004	0,000	0,000	0,002	0,002	0,012
HYUNDAI	0,073	0,074	0,144	0,195	0,253	0,227	0,113	0,165	0,236	0,307	0,239	0,176

Pada tabel III.2 atribut bulan di inisialisasikan dengan X1, X2, X3 dan seterusnya.

Dengan keterangan sebagai berikut :

X1 : Bulan Januari

X2 : Bulan Februari

X3 : Bulan Maret

X4 : Bulan April

- X5 : Bulan Mei
- X6 : Bulan Juni
- X7 : Bulan Juli
- X8 : Bulan Agustus
- X9 : Bulan September
- X10 : Bulan Oktober
- X11 : Bulan November
- X12 : Bulan Desember

2. Menentukan Jumlah *Cluster* Secara Random

Pada penelitian ini memilih 4 *cluster* secara acak pada baris dataset. Data yang di pilih sebagai berikut :

Tabel III. 3
Cluster Awal

No	Brand	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
1	BMW	0,004	0,004	0,006	0,009	0,007	0,008	0,008	0,010	0,013	0,020	0,023	0,025
2	WULING	0,008	0,032	0,099	0,099	0,101	0,047	0,064	0,281	0,591	0,541	0,534	1,000
3	TOYOTA	0,045	0,218	0,138	0,099	0,105	0,110	0,127	0,144	0,112	0,105	0,133	0,179
4	HYUNDAI	0,073	0,074	0,144	0,195	0,253	0,227	0,113	0,165	0,236	0,307	0,239	0,176

Setelah itu dilakukan perhitungan jarak tiap data ke pusat *cluster*.

3. Menghitung Jarak Tiap Data Kepusat *Cluster*

Setiap data akan diukur jaraknya dengan tiap *cluster* awal menggunakan rumus euclidean distance, sebagai berikut :

$$d_{Euclidean}(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

Sumber : (Rohmah et al., 2021)

Keterangan:

d (x,y) : jarak antara data pada titik x dan y

x : titik data objek

y : titik data centroid

i : jumlah atribut data

Dalam penelitian ini menggunakan 15 variabel, sehingga perhitungan jarak *euclidean* dilakukan sebanyak 15 kali pada masing-masing variabel. Berikut contoh perhitungan jarak *euclidean* pada variabel X1, baris pada brand Lexus sebagai berikut :

a) Dari data ke-1 ke pusat *cluster1* :

Data₁ (0,008; 0,019; 0,037; 0,052; 0,037; 0,059; 0,033; 0,065; 0,062.....),

Centroid₁ (0,004; 0,004; 0,006; 0,009; 0,007; 0,008.....)

$$\sqrt{(0,008 - 0,004)^2 + (0,019 - 0,004)^2 + \dots + (0,058 - 0,023)^2 + (0,084 - 0,025)^2}$$

$$= 0,137$$

Hasil perhitungan *euclidean* dari data ke-1 ke pusat *cluster 1* adalah 0,137.

b) Dari data ke-1 ke pusat *cluster2* :

Data₁ (0,008; 0,019; 0,037; 0,052; 0,037; 0,059; 0,033; 0,065; 0,062.....),

Centroid₂ (0,008; 0,032; 0,099; 0,099; 0,101; 0,047; 0,064.....)

$$\sqrt{(0,008 - 0,008)^2 + (0,019 - 0,032)^2 + \dots + (0,058 - 0,534)^2 + (0,084 - 1,000)^2}$$

$$= 1,2797$$

Hasil perhitungan *euclidean* dari data ke-1 ke pusat *cluster 2* adalah 1,2798

Tabel III. 4

Hasil Data Ke *Cluster* Pada Iterasi ke 1

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	C1	C2	C3	C4	JARAK TERDEKAT	CLUSTER
1	0,008	0,019	0,037	0,052	0,037	0,059	0,033	0,065	0,062	0,057	0,058	0,084	0,137	1,279789	0,308441242	0,5142	0,137	1
2	0,000	0,000	0,007	0,015	0,016	0,009	0,022	0,009	0,030	0,015	0,029	0,022	0,027	1,37983	0,414037438	0,6275	0,027	1
3	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,106	0,105	0,118	0,153	1,271467	0,394251189	0,5852	0,153	1
4	0,008	0,032	0,099	0,099	0,101	0,047	0,064	0,281	0,591	0,541	0,534	1,000	1,386	0	1,148297871	1,0175	0,000	2
5	0,045	0,218	0,138	0,099	0,105	0,110	0,127	0,144	0,112	0,105	0,133	0,179	0,422	1,148298	0	0,3666	0,000	3
6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,014	0,009	0,036	1,417956	0,452784717	0,6735	0,036	1
7	0,035	0,023	0,020	0,024	0,027	0,026	0,014	0,029	0,040	0,042	0,025	0,025	0,064	1,359009	0,375953455	0,5879	0,064	1
8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,022	0,017	0,032	1,409451	0,447792363	0,6689	0,032	1
9	0,000	0,000	0,000	0,000	0,000	0,000	0,008	0,003	0,005	0,001	0,004	0,000	0,042	1,424576	0,454431513	0,6750	0,042	1
10	0,000	0,001	0,002	0,003	0,002	0,000	0,000	0,000	0,000	0,000	0,003	0,001	0,043	1,427045	0,456198422	0,6764	0,043	1
11	0,000	0,000	0,004	0,001	0,003	0,000	0,006	0,006	0,007	0,002	0,006	0,012	0,032	1,413161	0,446038115	0,6674	0,032	1
12	0,000	0,001	0,002	0,005	0,004	0,003	0,001	0,006	0,005	0,007	0,002	0,002	0,037	1,420489	0,450065551	0,6684	0,037	1
13	0,000	0,000	0,000	0,000	0,000	0,000	0,006	0,005	0,005	0,006	0,003	0,002	0,038	1,421347	0,452725082	0,6724	0,038	1
14	0,000	0,000	0,003	0,002	0,000	0,000	0,004	0,000	0,000	0,002	0,002	0,012	0,038	1,419033	0,452079639	0,6741	0,038	1
15	0,073	0,074	0,144	0,195	0,253	0,227	0,113	0,165	0,236	0,307	0,239	0,176	0,639	1,017533	0,366642878	0,0000	0,000	4

Berikut data hasil perhitungan euclidean pada dataset, yaitu :

Setelah menempatkan data ke dalam pusat *cluster* terdekat, dengan cara menentukan *cluster* didapat dari nilai c_1 , c_2 , c_3 dan c_4 yang mendekati 0. Setelah itu penentuan *cluster* didapat dari jarak terdekat terdapat pada *cluster* berapa. Seperti contoh, pada tabel III.4 baris 1, memiliki nilai c_1 , c_2 , c_3 , c_4 yang mendekati 0 adalah c_1 yaitu bernilai 0, oleh karena itu termasuk dalam *cluster*.

4. Mendapatkan Pusat *Cluster* Baru

Proses perhitungan kembali dilakukan yaitu dengan menggunakan centroid baru dari iterasi pertama yang dihitung dari nilai rata-rata tiap kelompok *cluster*. Untuk centroid baru didapatkan dari jumlah seluruh data sebuah atribut pada satu centroid dibagi dengan jumlah data dan berlaku untuk seterusnya untuk semua atribut centroid.

1. Sebagai contoh untuk atribut brand pada centroid pertama :

$$\begin{aligned} &= \frac{\text{Jumlah seluruh brand pada cluster 1 (c1)}}{\text{banyak data pada cluster pertama (c1)}} \\ &= \frac{0,043 + 0,044 + 0,075 + 0,102 + 0,089 + \dots + 0,154 + 0,238 + 0,273 + 0,304}{12} \\ &= 0,004; 0,004; 0,006; 0,009; 0,007; 0,008; 0,008; 0,010; 0,013; 0,020; 0,023; 0,025 \end{aligned}$$

2. Sebagai contoh untuk atribut brand pada centroid kedua :

$$\begin{aligned} &= \frac{\text{Jumlah seluruh brand pada cluster 2 (c2)}}{\text{banyak data pada cluster kedua (c2)}} \\ &= \frac{0,008 + 0,032 + 0,099 + 0,099 + 0,101 + \dots + 0,591 + 0,541 + 0,534 + 1,000}{1} \end{aligned}$$

=

0,008; 0,032; 0,099; 0,099; 0,101; 0,047; 0,064; 0,281; 0,591; 0,541; 0,534; 1,000

Tabel III. 5
Hasil Centroid Baru

Centroid	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
C1	0,004	0,004	0,006	0,009	0,007	0,008	0,008	0,010	0,013	0,020	0,023	0,025
C2	0,008	0,032	0,099	0,099	0,101	0,047	0,064	0,281	0,591	0,541	0,534	1,000
C3	0,045	0,218	0,138	0,099	0,105	0,110	0,127	0,144	0,112	0,105	0,133	0,179
C4	0,073	0,074	0,144	0,195	0,253	0,227	0,113	0,165	0,236	0,307	0,239	0,176

5. Menghitung Kembali Data Dengan Pusat *Cluster* Yang Baru (Iterasi 2)

Proses selanjutnya dengan menggunakan centroid baru dilakukan pengulangan untuk mendapatkan hasil dengan nilai rata-rata yang tidak berubah atau konvergen.

Dimana perhitungan pengulangan pada penelitian ini dengan centroid baru berhenti pada perhitungan iterasi ke 2. Dimana centroid terakhir yang tidak mengalami perubahan dapat dilihat pada tabel III.6 dibawah ini :

Tabel III. 6
Hasil Centroid Pada Iterasi ke 2

Centroid	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
C1	0,004	0,004	0,006	0,009	0,007	0,008	0,008	0,010	0,013	0,020	0,023	0,025
C2	0,008	0,032	0,099	0,099	0,101	0,047	0,064	0,281	0,591	0,541	0,534	1,000
C3	0,045	0,218	0,138	0,099	0,105	0,110	0,127	0,144	0,112	0,105	0,133	0,179
C4	0,073	0,074	0,144	0,195	0,253	0,227	0,113	0,165	0,236	0,307	0,239	0,176

Berikut dibawah ini merupakan hasil dari iterasi terakhir :

Tabel III. 7
Hasil Data Ke *Cluster* Pada Iterasi ke 2 (Terakhir)

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	C1	C2	C3	C4	JARAK TERDEKAT	CLUSTER
1	0,008	0,019	0,037	0,052	0,037	0,059	0,033	0,065	0,062	0,057	0,058	0,084	0,137	1,279769	0,308441	0,5142	0,137	1
2	0,000	0,000	0,007	0,015	0,016	0,009	0,022	0,009	0,030	0,015	0,029	0,022	0,027	1,37983	0,414037	0,6275	0,027	1
3	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,106	0,105	0,118	0,153	1,271467	0,394251	0,5852	0,153	1
4	0,008	0,032	0,099	0,099	0,101	0,047	0,064	0,281	0,591	0,541	0,534	1,000	1,386	0	1,148298	1,0175	0,000	2
5	0,045	0,218	0,138	0,099	0,105	0,110	0,127	0,144	0,112	0,105	0,133	0,179	0,422	1,148298	0	0,3666	0,000	3
6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,014	0,009	0,036	1,417956	0,452785	0,6735	0,036	1
7	0,035	0,023	0,020	0,024	0,027	0,026	0,014	0,029	0,040	0,042	0,025	0,025	0,064	1,359009	0,375953	0,5879	0,064	1
8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,022	0,017	0,032	1,409451	0,447792	0,6689	0,032	1
9	0,000	0,000	0,000	0,000	0,000	0,000	0,008	0,003	0,005	0,001	0,004	0,000	0,042	1,424576	0,454432	0,6750	0,042	1
10	0,000	0,001	0,002	0,003	0,002	0,002	0,000	0,000	0,000	0,000	0,003	0,001	0,043	1,427045	0,456198	0,6764	0,043	1
11	0,000	0,000	0,004	0,001	0,003	0,000	0,006	0,006	0,007	0,002	0,006	0,012	0,032	1,413161	0,446038	0,6674	0,032	1
12	0,000	0,001	0,002	0,005	0,004	0,003	0,001	0,006	0,005	0,007	0,002	0,002	0,037	1,420489	0,450066	0,6684	0,037	1
13	0,000	0,000	0,000	0,000	0,000	0,000	0,006	0,005	0,005	0,006	0,003	0,002	0,038	1,421347	0,452725	0,6724	0,038	1
14	0,000	0,000	0,003	0,002	0,000	0,000	0,004	0,000	0,000	0,002	0,002	0,012	0,038	1,419033	0,45208	0,6741	0,038	1
15	0,073	0,074	0,144	0,195	0,253	0,227	0,113	0,165	0,236	0,307	0,239	0,176	0,639	1,017533	0,366643	0,0000	0,000	4

6. Pengujian dengan Davies Bouldin Index

Untuk mendapatkan nilai Davies Bouldin Index, terlebih dahulu hitung nilai *Sum of square within-cluster*, *Sum of square between-cluster* dan *Ratio*. Untuk melakukan perhitungan Davies Bouldin Index data yang digunakan adalah data yang sudah terbentuk pada *clustering* akhir dan titik pusat (centroid) *cluster* yang terakhir seperti pada tabel III.6.

a. Hitung SSWi (*Sum of Squared Within-cluster*)

Menghitung jarak antara setiap titik data dengan centroid *cluster* masing-masing, kemudian menjumlahkan kuadrat dari jarak-jarak tersebut.

$$SSWi = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_j)$$

Sumber : (Syamfithriani et al., 2023)

Sebagai contoh :

a). Data ke 1 (*Cluster 1*):

$$\text{Data}_{1,C_1} = (0.008-0,004)^2 + (0.019-0,004)^2 + \dots + (0.058-0.023)^2 + (0.084-0.025)^2$$

$$\sqrt{(0.008 - 0,004)^2 + (0.019 - 0,004)^2 + \dots + (0.058 - 0.023)^2 + (0.084 - 0.025)^2}$$
$$= 0.137$$

Hitung SSW untuk setiap data dan kemudian di jumlahkan :

Sebagai contoh:

b). Total SSW untuk *Cluster 1* :

SSW

$$= \frac{0,137 + 0,027 + 0,153 + 0,036 + 0,064 + 0,032 + 0,042 + +0,038 + 0,057}{12}$$

$$SSW_1 = 0,057$$

Lakukan hal yang sama untuk *Cluster 2, 3 dan 4*.

Didapatkan nilai, sebagai berikut :

$$SSW2 = 0$$

$$SSW3 = 0$$

$$SSW4 = 0$$

Setelah mengetahui nilai SSW maka selanjutnya dilakukan perhitungan *Sum of square between-cluster* (SSB). Untuk menghitung nilai SSB maka diperlukan centroid terakhir pada iterasi terakhir. Berikut tabel III.8 centroid terakhir yang didapat dari iterasi terakhir.

Tabel III. 8
Titik Pusat Akhir Dari Proses *Clustering*

Centroid	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
C1	0,004	0,004	0,006	0,009	0,007	0,008	0,008	0,010	0,013	0,020	0,023	0,025
C2	0,008	0,032	0,099	0,099	0,101	0,047	0,064	0,281	0,591	0,541	0,534	1,000
C3	0,045	0,218	0,138	0,099	0,105	0,110	0,127	0,144	0,112	0,105	0,133	0,179
C4	0,073	0,074	0,144	0,195	0,253	0,227	0,113	0,165	0,236	0,307	0,239	0,176

b. Hitung SSB_{ij} (*Sum of Squared Between-cluster*)

$$SSB_{i,j} = d(c_i, c_j)$$

Sumber : (Syamfithriani et al., 2023)

Menghitung jarak antara setiap pasangan centroid.

1) Sebagai contoh, untuk SSB_{1,2} (centroid 1 dan centroid 2) :

$$SSB_{1,2}$$

$$= \sqrt{(0,004 - 0,008)^2 + (0,004 - 0,032)^2 + \dots + (0,006 - 0,099)^2 + (0,025 - 1,000)^2}$$

Hitung SSB untuk setiap pasangan centroid dan kemudian jumlahkan:

$$SSB_{1,2} = 1,39$$

Lakukan hal yang sama untuk pasangan centroid lainnya (C1,C3; C1,C4; C2,C3; C2,C4; C3,C4).

Didapatkan hasil SSB sebagai berikut :

SSB	1	2	3	4
1	0	1,386	0,422	0,639
2	1,386	0	1,148	1,018
3	0,422	1,148	0	0,367
4	0,639	1,018	0,367	0

$$SSB_{1,2} = 1,3862$$

$$SSB_{1,3} = 0,4219$$

$$SSB_{1,4} = 0,6391$$

$$SSB_{2,3} = 1,1483$$

$$SSB_{2,4} = 1,0175$$

$$SSB_{3,4} = 0,3666$$

c. Hitung Rij (Ratio of SSW_i to SSB_{ij})

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}}$$

Sumber : (Syamfithriani et al., 2023)

Sebagai contoh, untuk $R_{1,2}$:

$$R_{1,2} = \frac{SSW_1 + SSW_2}{SSB_{1,2}}$$

$$R_{1,2} = \frac{0,057 + 0}{1,39}$$

$$= 0,04$$

Didapatkan hasil Ratio keseluruhan sebagai berikut :

R	1	2	3	4	R MAX
1	0	0,041	0,134	0,088	0,134
2	0,041	0	0	0	0,0408
3	0,134	0	0	0	0,134
4	0,088	0	0	0	0,0884

$$R_{1,2} = 0,0408$$

$$R_{1,3} = 0,134$$

$$R_{1,4} = 0,0884$$

$$R_{2,3} = 0$$

$$R_{2,4} = 0$$

$$R_{3,4} = 0$$

d. Menghitung DBI (Davies-Bouldin Index)

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j})$$

Sumber : (Syamfithriani et al., 2023)

$$DBI = \frac{0,134 + 0,041 + 0,134 + 0,088}{4}$$

$$= 0,099$$

Dengan menggunakan DBI suatu cluster akan dianggap memiliki hasil clustering yang optimal jika memiliki DBI minimal (Syamfithriani et al., 2023).

7. Nilai Persentase

Nilai persentase ini sebagai penentu tingkat persentase hasil perhitungan data pada tiap *clusster*-nya. Berikut rumus untuk menghitung nilai persentase:

$$persentase = \frac{\text{Jumlah data yang sesuai}}{\text{Jumlah keseluruhan data}} \times 100\%$$

Sumber : (Butsianto & Mayangwulan, 2020)

a. Hasil persentase pada *cluster*1, sebagai berikut :

Data yang terdapat pada *cluster* 1 berjumlah 12 data, maka nilai persentasenya sebagai berikut :

$$persentase = \frac{12}{15} \times 100\% = 80\%$$

b. Hasil persentase pada *cluster* 2, sebagai berikut :

Data yang terdapat pada *cluster* 2 berjumlah 1 data, maka nilai persentasenya sebagai berikut :

$$persentase = \frac{1}{15} \times 100\% = 6,7\%$$

c. Hasil persentase pada *cluster* 3, sebagai berikut :

Data yang terdapat pada *cluster* 3 berjumlah 1 data, maka nilai persentasenya sebagai berikut :

$$persentase = \frac{1}{15} \times 100\% = 6,7\%$$

d. Hasil persentase pada *cluster* 4, sebagai berikut :

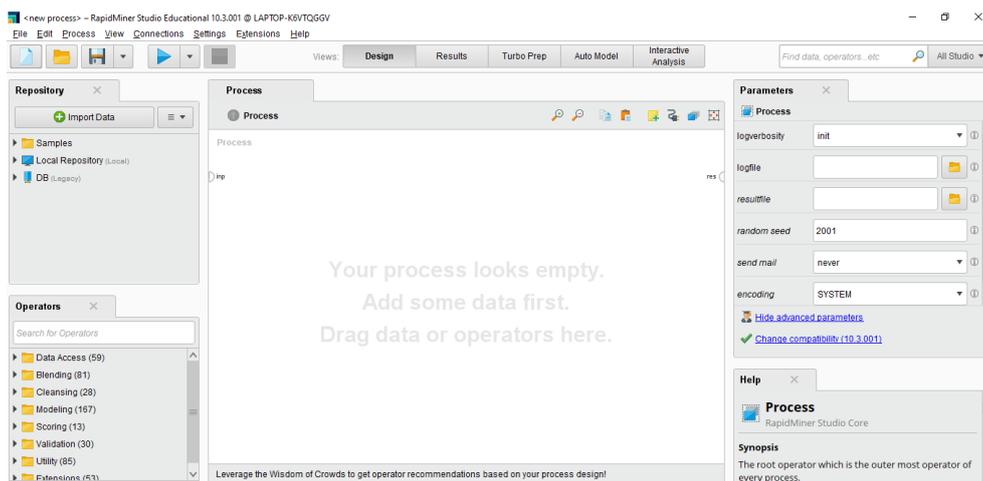
Data yang terdapat pada *cluster* 4 berjumlah 1 data, maka nilai persentasenya sebagai berikut :

$$persentase = \frac{1}{15} \times 100\% = 6,7\%$$

B. Implementasi rapid miner

Pada Implementasi Dan Pengujian ini, peneliti akan menggunakan sebuah software RapidMiner Studio Versi 10.3.001. Dengan pengujian data menggunakan software ini kita akan membandingkan bagaimana hasil pengolahan data secara manual dengan hasil pengolahan data menggunakan sebuah software. Dalam melakukan import data pada aplikasi RapidMiner Studio v.10.3.001 terdiri dari 6 tahap, adapun langkah-langkahnya sebagai berikut :

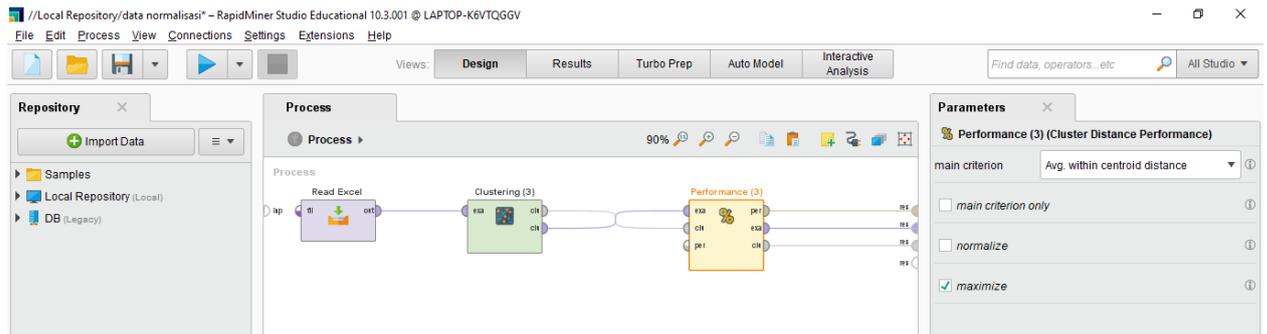
1. Buka aplikasi *RapidMiner*, tampilan halaman seperti dibawah ini:



Gambar III. 2
Tampilan Awal *RapidMiner*

2. Buat design operator di jendela *design*, tampilan seperti gambar III. 3. Operator yang digunakan dalam penelitian ini adalah :

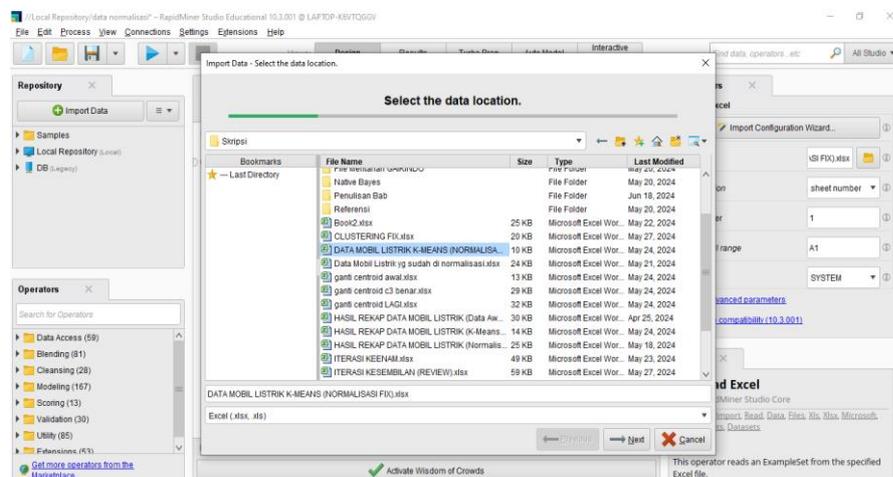
- a) *Read Excel* : Digunakan untuk membaca data yang digunakan untuk penelitian. Data yang dibaca dalam bentuk *Microsoft Excel*.
- b) *Clustering* : digunakan untuk implementasi algoritma *k-means*.
- c) *Cluster distance performance* : digunakan untuk mengevaluasi kinerja metode pengelompokkan berbasis centroid/*cluster*.



Gambar III. 3
Desain Operator Algoritma *K-means*

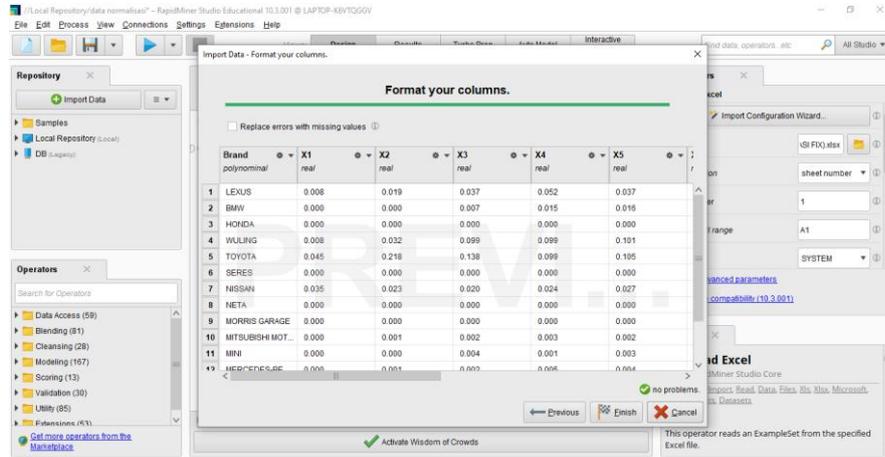
3. Mencari lokasi file Tahap pertama adalah mencari lokasi file yang sudah dibuat sebelumnya dengan format *.xlsx* atau *.xls*. Pada pengujian ini data yang akan diuji disimpan dengan nama DATA MOBIL LISTRIK *K-MEANS* (NORMALISASI FIX).*xlsx* lalu pilih

4. Memilih *sheet*



Gambar III. 4
Tampilan pencarian lokasi file

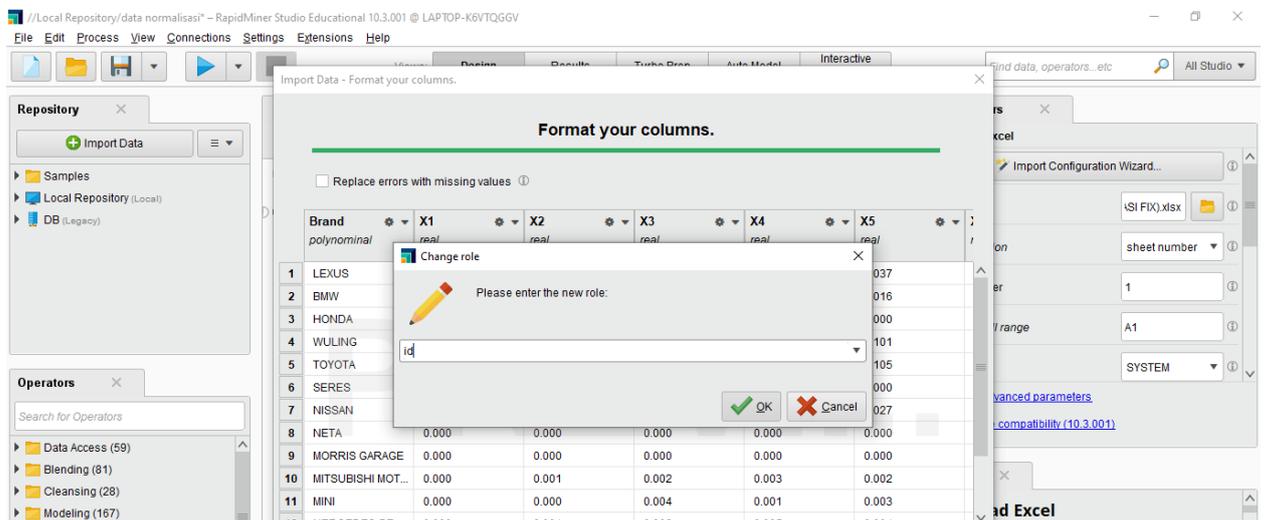
Kemudian pilih next maka lanjut ke tahap berikutnya memilih sheet yang berisikan sumber data pada Microsoft Excel yang digunakan, seperti terlihat pada gambar III. 5.



Gambar III. 5
Tampilan pemilihan sheet

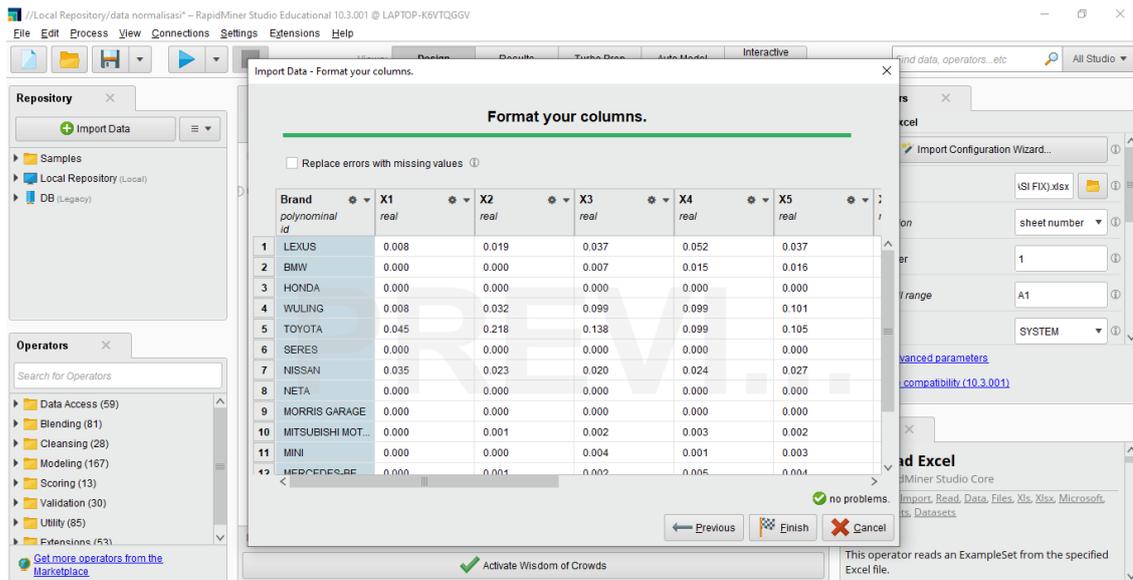
5. Pemberian anotasi

Tahap ini merupakan pemberian anotasi id pada kolom brand. Kemudian klik tombol Ok.



Gambar III. 6
Tahap pemberian anotasi id

Setelah berhasil dirubah, kolom brand berubah menjadi warna biru. Kemudian klik finish.



Gambar III. 7
Pemberian anotasi id

- Langkah selanjutnya menjalankan proses algoritma *K-means*. Klik pada gambar untuk menjalankan proses atau menampilkan hasil. Maka proses akan menghasilkan 3 bagian menu, yaitu *Cluster Model*, *Example Set* dan *Performance Vector (Performance)*. Berikut hasil *Cluster Model* :



Gambar III. 8
Hasil *Cluster Model*

3.6 Evaluasi Hasil

- Langkah selanjutnya menjalankan proses algoritma *K-means*.

Klik pada gambar untuk menjalankan proses atau menampilkan hasil. Maka proses akan menghasilkan 3 bagian menu, yaitu *Cluster Model*, *Example Set* dan *Performance Vector (Performance)*. Berikut hasil *Cluster Model* :



Gambar III. 10
Hasil *Cluster Model*

Pada gambar III. 9 Hasil *cluster* dalam penelitian ini menggunakan *RapidMiner*, menghasilkan kelompok *cluster 0* terdiri dari 12 data/brand, kelompok *cluster 1* terdiri dari 1 data/brand, kelompok *cluster 2* terdiri dari 1 data/brand dan kelompok *cluster 3* terdiri dari 1 data/brand, dengan keseluruhan jumlah data/brand yaitu 15. Selanjutnya brand mana saja yang menempati sebuah *cluster*, terdapat dalam gambar III.10.

Row No. ↑	Brand	cluster	X1	X2	X3	X4	X5	X6	X7
1	LEXUS	cluster_0	0.008	0.019	0.037	0.052	0.037	0.059	0.03
2	BMW	cluster_0	0	0	0.007	0.015	0.016	0.009	0.02
3	HONDA	cluster_0	0	0	0	0	0	0	0
4	WULING	cluster_1	0.008	0.032	0.099	0.099	0.101	0.047	0.06
5	TOYOTA	cluster_3	0.045	0.218	0.138	0.099	0.105	0.110	0.12
6	SERES	cluster_0	0	0	0	0	0	0	0
7	NISSAN	cluster_0	0.035	0.023	0.020	0.024	0.027	0.026	0.01
8	NETA	cluster_0	0	0	0	0	0	0	0
9	MORRIS GAR...	cluster_0	0	0	0	0	0	0	0.00
10	MITSUBISHI ...	cluster_0	0	0.001	0.002	0.003	0.002	0.002	0
11	MINI	cluster_0	0	0	0.004	0.001	0.003	0	0.00
12	MERCEDES-...	cluster_0	0	0.001	0.002	0.005	0.004	0.003	0.00
13	MAZDA	cluster_0	0	0	0	0	0	0	0.00
14	KIA	cluster_0	0	0	0.003	0.002	0	0	0.00
15	HYUNDAI	cluster_2	0.073	0.074	0.144	0.195	0.253	0.227	0.11

Gambar III. 9
Hasil *ExampleSet (Clustering)*

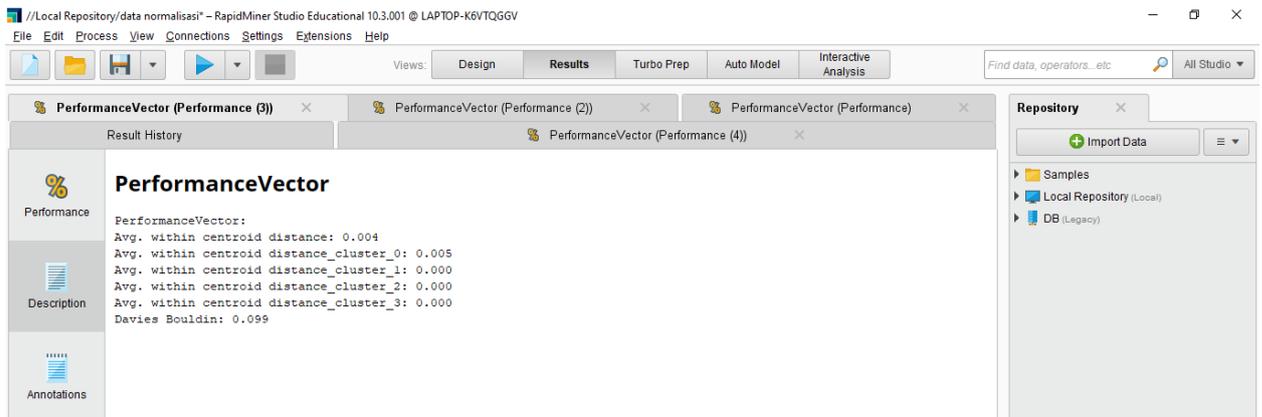
Pada gambar III. 9, merupakan data brand yang menempati suatu kelompok *cluster*, berikut penjelasannya:

- a) Kelompok *Cluster 0*, yaitu Brand Lexus dengan penjualan tertinggi X12 dan terendah pada X1, Honda dengan penjualan tertinggi X12 dan

terendah pada X1 hingga X9, Nissan dengan penjualan tertinggi X10 dan terendah pada X7, Bmw dengan penjualan tertinggi X9 dan terendah pada X1 serta X2, Seres dengan penjualan tertinggi X11 dan terendah pada X1 hingga X10, Neta dengan penjualan tertinggi X11 dan terendah pada X1 hingga X10, Morris Garage dengan penjualan tertinggi X7 dan terendah pada X1 hingga X6, Mitsubishi Motors dengan penjualan tertinggi X4 serta X11 dan terendah pada X1, Mini dengan penjualan tertinggi X12 dengan penjualan terendah pada X1 dan X2, Mercedes-Benz Pc dengan penjualan tertinggi X10 dan terendah pada X1, Mazda dengan penjualan tertinggi X7 dan X10 dengan penjualan terendah pada X1 hingga X6, dan Kia dengan penjualan tertinggi X12 dan penjualan terendah pada X1, X2, X5, X6, X8 dan X9.

- b) Kelompok *Cluster 1*, yaitu Brand Wuling dengan penjualan tertinggi pada X12 dan terendah pada X1.
- c) Kelompok *Cluster 2*, yaitu Brand Hyundai dengan penjualan tertinggi ada pada X10 dan penjualan terendah pada X1.
- d) Kelompok *Cluster 3*, yaitu Brand Toyota dengan penjualan tertinggi pada X2 dan terendah pada X1.

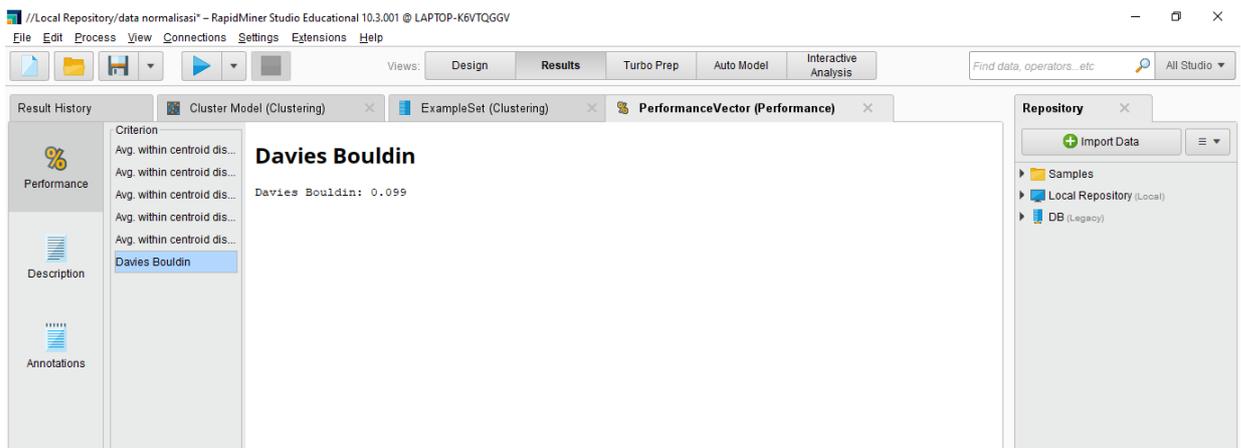
Dengan perolehan nilai performance tiap *Cluster* sebagai berikut :



Gambar III. 11
Hasil *Performance* (*Clustering*)

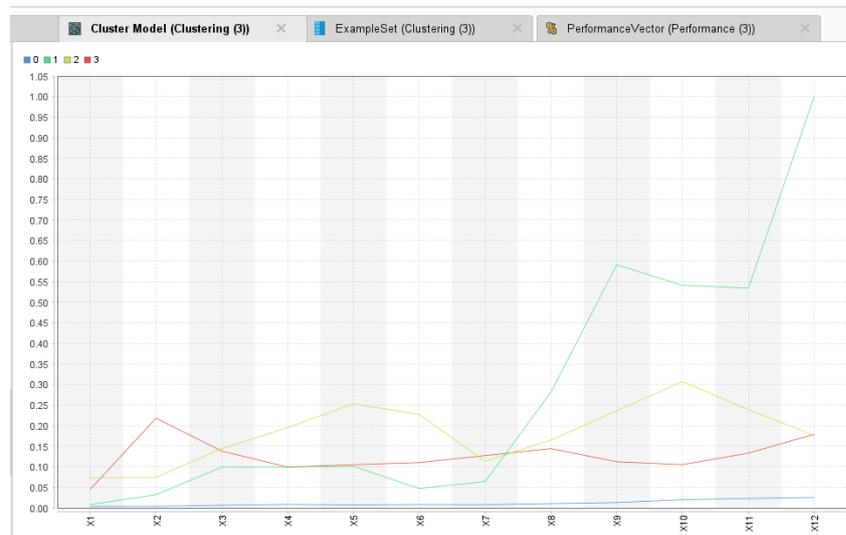
Berdasarkan gambar III. 11 diatas, rata-rata *Performance* keseluruhan sebesar 0,004, rata-rata *performance cluster* 0 sebesar 0,005, rata-rata *performance cluster* 1 sebesar 0,000, rata-rata *performance cluster* 2 sebesar 0,000 dan rata-rata *performance cluster* 3 sebesar 0,000.

Kemudian untuk nilai davies bouldin, semakin kecil nilai davies bouldin index yang dihasilkan maka semakin baik pengelompokkan *cluster* yang terbentuk. Hasil perhitungan menggunakan algorithm *K-means* menunjukkan nilai davies bouldin index sebesar 0,099.



Gambar III. 12
Hasil Davies Bouldin Index

Gambar III.13 dibawah ini merupakan penyebaran data dari bulan X1 sampai dengan X12, yaitu:



Gambar III. 13
Penyebaran data Bulan dari X1 sampai X12

Hasil Cluster model

Hasil jumlah *cluster* (pengelompokkan) pada data brand mobil listrik memiliki jumlah yang sama antara *Microsoft Excel* dengan menggunakan *Rapid Miner*. Untuk hasilnya dapat dilihat pada tabel III.9 dibawah ini:

Tabel III. 9
Hasil Jumlah Cluster

Cluster Model			
Microsoft Excel		RapidMiner	
Cluster	Jumlah Brand	Cluster	Jumlah Brand
Cluster 1	12	Cluster 0	12
Cluster 2	1	Cluster 1	1
Cluster 3	1	Cluster 2	1
Cluster 4	1	Cluster 3	1

Hasil dari table diatas menunjukkan bahwa pengelompokkan *cluster* memiliki hasil yang sama, dimana hasil diatas didapatkan dari *Microsoft Excel* dan *Rapid Miner*.

Hasil pengelompokkan *cluster*

Pada tabel III. 10 yang memiliki tingkat penjualan tertinggi, yaitu kelompok *Cluster 2* pada *Microsoft Excel* dan *Cluster 1* pada *Rapid Miner* dengan Brand yang sama yaitu Wuling. Untuk nama peng-*Clusteran* memang berbeda antara menggunakan *Microsoft Excel* dengan *RapidMiner*, namun yang terpenting adalah hasil kelompok variabel yang sama. Setelah melakukan pengelompokkan dan pengujian yang dilakukan secara manual dan juga pengujian menggunakan *RapidMiner*.

- 1) Diperoleh nilai untuk semua penjualan pada semua brand yang terdapat pada *cluster 1* sebagai berikut :

$$= 38 + 90 + 172 + 241 + 173 + \dots + 8 + 57 =$$

7.618 penjualan

- 2) Total penjualan pada brand Wuling, karena hanya terdapat pada 1 brand saja pada *cluster 2* ini,

$$= 39 + 147 + 457 + \dots + 2507 + 2479 + 4638$$

= 15.757 penjualan

- 3) Total penjualan pada brand Toyota, karena hanya terdapat pada 1 brand saja pada *cluster 3* ini,

$$= 208 + 1012 + 641 + \dots + 485 + 616 + 828$$

$$= 7.025 \text{ penjualan}$$

- 4) Total penjualan pada brand Hyundai, karena hanya terdapat pada 1 brand saja pada *cluster* 4 ini,

$$= 339 + 342 + 666 + \dots + 1423 + 1110 + 818$$

$$= 10.214 \text{ penjualan}$$

Didapatkan hasil untuk penjualan pada *Cluster* 1, dengan penjualan sebesar 7.618 unit. *Cluster* 2 sebesar 15.757 unit. *Cluster* 3 sebesar 7.025 unit. Dan *Cluster* 4 sebesar 10.214 unit. Maka diperoleh tingkat *Cluster* sebagai berikut, *Cluster* 2 sebagai kategori tingkat penjualan Sangat Laris yaitu dengan jumlah penjualan 15.757 unit. *Cluster* 3 sebagai kategori Kurang laris, *Cluster* 1 sebagai kategori Cukup Laris dan *Cluster* 3 termasuk kategori Laris karena memiliki total penjualan terendah yaitu sebesar 1,515 unit. Berdasarkan hasil analisa, persentase untuk tiap *cluster* adalah (80%) untuk *Cluster* 1, (6,7%) untuk *Cluster* 2, (6,7%) untuk *Cluster* 3 dan (6,7%) pada *Cluster* 4. Untuk hasil lebih jelasnya dapat dilihat pada tabel dibawah ini:

Tabel III. 10
Hasil *Clustering* secara manual dan menggunakan rapidminer

<i>Cluster Model</i>				Persentase	Keterangan Tingkat Penjualan
<i>Microsoft Excel</i>		<i>RapidMiner</i>			
<i>Cluster</i>	Brand	<i>Cluster</i>	Brand		
Cluster 1 (C ₁)	Lexus	Cluster 0 (C ₀)	Lexus	80%	Cukup Laris
	Honda		Honda		
	Nissan		Nissan		
	Bmw		Bmw		
	Seres		Seres		
	Neta		Neta		
	Morris Garage		Morris Garage		
	Mitsubishi Motors		Mitsubishi Motors		

	Mini Mercedez- Benz Pc Mazda Kia		Mini Mercedez- Benz Pc Mazda Kia		
<i>Cluster 2</i> (C ₂)	Wuling	<i>Cluster 1</i> (C ₁)	Wuling	6,7%	Sangat Laris
<i>Cluster 3</i> (C ₃)	Toyota	<i>Cluster 3</i> (C ₃)	Toyota	6,7%	Kurang Laris
<i>Cluster 4</i>	Hyundai	<i>Cluster 2</i> (C ₂)	Hyundai	6,7%	Laris
Total	12		12	100%	

Peramalan data penjualan tahun 2024

Metode *least square* digunakan untuk memperoleh informasi mengenai keadaan peramalan produksi di masa yang akan datang dengan mengacu pada data-data historis yang telah ada sebelumnya (Prawiro et al., 2022). data yang digunakan adalah data penjualan mobil listrik tahun 2023, dengan total data sebanyak 70 data dan memiliki atribut bulan sebanyak 12.

$$Y = a + bX$$

$$a = \frac{\sum Y}{n}$$

$$b = \frac{\sum XY}{\sum X^2}$$

Sumber : (Prawiro et al., 2022)

Dimana:

Y = nilai *trend* (*forecast/proyeksi/peramalan*)

α = konstanta

b = slope/kecondongan

X = waktu

Berdasarkan rumus diatas, maka data penjualan mobil listrik dapat dianalisis sebagai berikut :

Tabel III. 11
Analisis Penjualan Mobil Listrik

Bulan	Penjualan Unit (Y)	Waktu (x)	Penjualan dikali Waktu (XY)	x ²
Januari	328	-11	-3608	121
Februari	780	-9	-7020	81
Maret	1386	-7	-9702	49
April	1564	-5	-7820	25
Mei	1904	-3	-5712	9
Juni	1677	-1	-1677	1
Juli	1310	1	1310	1
Agustus	1590	3	4770	9
September	2285	5	11425	25
Oktober	2691	7	18837	49
November	2800	9	25200	81
Desember	4311	11	47421	121
Jumlah	22626	0	73424	572

Untuk mencari α (konstanta) dilakukan perhitungan sebagai berikut :

$$a = \frac{\sum Y}{n} = \frac{22626}{12} = 1885,5$$

Untuk mencari b (kecondongan) dilakukan perhitungan sebagai berikut :

$$b = \frac{\sum XY}{\sum X^2} = \frac{73424}{572} = 128,3636364$$

Setelah mendapatkan hasil α (konstanta) dan b (kecondongan) dapat dilakukan peramalan penjualan mobil listrik di Indonesia pada tahun 2024.

Berikut adalah salah satu contoh hasil peramalan berdasarkan hasil yang sudah dibuat:

Hasil peramalan bulan Januari 2024

Karena banyak data (bulan) adalah 12 yang dimana adalah angka genap, maka X perbulan ditambahkan 2, sehingga X pada bulan Januari 2024 adalah 13.

$$Y = a + bX$$

$$Y = 1885,5 + 128,3636364 (X)$$

$$Y = 1885,5 + 128,3636364 (13)$$

$$Y = 3.554$$

Berikut adalah hasil keseluruhan ramalan pada bulan Januari hingga Desember tahun 2024 :

Tabel III. 12
Hasil ramalan bulan Januari hingga Desember 2024

Bulan	Waktu (x)	Penjualan Unit (Y)
Januari	13	3.554
Februari	15	3.811
Maret	17	4.068
April	19	4.324
Mei	21	4.581
Juni	23	4.838
Juli	25	5.095
Agustus	27	5.351
September	29	5.608
Oktober	31	5.865
November	33	6.122
Desember	35	6.378

3.7 Kesimpulan

Dari hasil penelitian yang telah dilakukan oleh peneliti, maka didapatkan kesimpulan sebagai berikut:

- 1) Mengacu pada tabel III.10, Kelompok Cluster 0, yaitu Brand Lexus dengan penjualan tertinggi X12 dan terendah pada X1, Honda dengan penjualan

tertinggi X12 dan terendah pada X1 hingga X9, Nissan dengan penjualan tertinggi X10 dan terendah pada X7, Bmw dengan penjualan tertinggi X9 dan terendah pada X1 serta X2, Seres dengan penjualan tertinggi X11 dan terendah pada X1 hingga X10, Neta dengan penjualan tertinggi X11 dan terendah pada X1 hingga X10, Morris Garage dengan penjualan tertinggi X7 dan terendah pada X1 hingga X6, Mitsubishi Motors dengan penjualan tertinggi X4 serta X11 dan terendah pada X1, Mini dengan penjualan tertinggi X12 dengan penjualan terendah pada X1 dan X2, Mercedes-Benz Pc dengan penjualan tertinggi X10 dan terendah pada X1, Mazda dengan penjualan tertinggi X7 dan X10 dengan penjualan terendah pada X1 hingga X6, dan Kia dengan penjualan tertinggi X12 dan penjualan terendah pada X1, X2, X5, X6, X8 dan X9. Kelompok Cluster 1, yaitu Brand Wuling dengan penjualan tertinggi pada X12 dan terendah pada X1. Kelompok Cluster 2, yaitu Brand Hyundai dengan penjualan tertinggi ada pada X10 dan penjualan terendah pada X1. Kelompok Cluster 3, yaitu Brand Toyota dengan penjualan tertinggi pada X2 dan terendah pada X1. Dapat disimpulkan bahwa rata-rata penjualan tertinggi ada pada X10, X11 dan X12 atau dengan kata lain yaitu, pada Bulan Oktober, November dan Desember. Sedangkan untuk penjualan terendah terjadi pada X1, X2 dan X6, atau pada Bulan Januari, Februari dan Juni.

- 2) Metode *clustering* dengan menggunakan algoritma *K-Means* dapat digunakan untuk mengelompokkan data penjualan mobil listrik berdasarkan jumlah penjualannya selama 5 tahun yaitu dari tahun 2019-2023 dari 154 data awal. Di dapatkan 4 *Cluster*, yaitu *Cluster 0* (Cukup Laris), *Cluster 1* (Sangat Laris), *Cluster 3* (Kurang Laris) dan *Cluster 2* (Laris). Sehingga pihak perusahaan dengan penjualan mobil listrik yang masuk dalam *Cluster 3* dapat merancang

dan menerapkan strategi pemasaran yang lebih efektif untuk meningkatkan penjualan mobil listrik mereka.

- 3) Metode *Clustering* ini dapat digunakan untuk membantu pihak perusahaan-perusahaan otomotif khususnya penjualan mobil listrik dalam melihat tingkat penjualan mobil mereka dengan perusahaan-perusahaan lain. Berdasarkan hasil pengujian dari cara manual dan dengan software menggunakan rapidminer v10.3.001 mendapatkan hasil yang sama dan tepat Yaitu *Cluster 0* jumlah data brand sebanyak 12 dengan presentase 80% (Cukup Laris), *cluster 1* dengan jumlah 1 data dan presentase 6,7% (Sangat Laris), *cluster 2* jumlah 1 data dengan presentase 6,7% (Laris) dan *cluster 3* jumlah data 1 dengan presentase 6,7% (Kurang Laris). Untuk hasil perhitungan menggunakan algoritma *K-means* nilai DBI (Davies Bouldin Index) pada penelitian ini menghasilkan nilai sebesar 0,099.
- 4) Hasil ramalan dengan metode Metode *least square* menunjukkan bahwa pada bulan Januari sampai dengan Desember tahun 2024 mengalami kenaikan disetiap bulannya.

BAB IV

PENUTUP

4.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan oleh penulis, maka dapat ditarik kesimpulan sebagai berikut :

1. Penelitian ini bertujuan untuk mengelompokkan tingkat pembelian mobil listrik di Indonesia pada tahun 2019-2023.
2. Hasil Davies-Bouldin Index (DBI) dalam mengukur kualitas klustering menghasilkan nilai sebesar 0,099. Yang artinya menunjukkan bahwa hasil *clustering* yang dilakukan sangat baik.
3. Hasil persentase yang didapat dalam pengolahan penelitian menggunakan *Microsoft Excel* ini adalah:
 - Cluster 1* yang berjumlah 12 data, maka nilai persentase 80%
 - Cluster 2* yang berjumlah 1 data, maka nilai persentase 6,7%
 - Cluster 3* yang berjumlah 1 data, maka nilai persentase 6,7%
 - Cluster 4* yang berjumlah 1 data, maka nilai persentase 6,7%
4. Hasil rata-rata *performance* yang didapat dalam pengolahan penelitian menggunakan *RapidMiner* adalah:
 - Cluster 0* memiliki nilai rata-rata *performance* sebesar 0,005
 - Cluster 1* memiliki nilai rata-rata *performance* sebesar 0,000
 - Cluster 2* memiliki nilai rata-rata *performance* sebesar 0,000
 - Cluster 3* memiliki nilai rata-rata *performance* sebesar 0,000
5. Hasil pengelompokan brand pada penelitian ini adalah sebagai berikut :

Tabel IV. 1
 Pengelompokkan *Cluster* Berdasarkan Brand

<i>Cluster Model</i>				Presentasi	Keterangan Tingkat Penjualan
<i>Microsoft Excel</i>		<i>RapidMiner</i>			
<i>Cluster</i>	Brand	<i>Cluster</i>	Brand		
Cluster 1 (C ₁)	Lexus	Cluster 0 (C ₀)	Lexus	80%	Kurang Laris
	Honda		Honda		
	Nissan		Nissan		
	Bmw		Bmw		
	Seres		Seres		
	Neta		Neta		
	Morris Garage		Morris Garage		
	Mitsubishi Motors		Mitsubishi Motors		
	Mini		Mini		
	Mercedez-Benz Pc		Mercedez-Benz Pc		
	Mazda		Mazda		
Kia	Kia				
Cluster 2 (C ₂)	Wuling	Cluster 1 (C ₁)	Wuling	6,7%	Sangat Laris
Cluster 3 (C ₃)	Toyota	Cluster 3 (C ₃)	Toyota	6,7%	Laris
Cluster 4	Hyundai	Cluster 2 (C ₂)	Hyundai	6,7%	Cukup Laris
Total	12		12	100%	

6. Hasil ramalan dengan metode Metode least square menunjukkan bahwa pada bulan Januari sampai dengan Desember tahun 2024 mengalami kenaikan disetiap bulannya

4.2 **Saran**

Penelitian selanjutnya dapat dikembangkan dengan menggunakan metode analisis yang beragam, seperti : Random Forest, atau Neural Networks untuk mendapatkan perspektif yang lebih komprehensif mengenai faktor-faktor yang mempengaruhi penjualan mobil listrik. Dan juga bisa memanfaatkan data eksternal : seperti tren ekonomi, harga bahan bakar, kebijakan pemerintah terkait kendaraan listrik dan preferensi konsumen untuk memperkaya analisis dan memberikan insight yang lebih mendalam. Diharapkan penelitian selanjutnya dapat memberikan pemahaman yang lebih komprehensif dan strategi yang lebih efektif untuk meningkatkan penjualan mobil listrik.

BAB V

REFERENSI

- Putra, A. S., Anubhakti, D., & Hin, L. L. (2023). Penerapan Algoritma K-Nearest Neighbor Untuk Analisis Sentimen Pengguna Twitter Terhadap Kebijakan Subsidi Kendaraan Listrik. *Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, 2(2), 736–744. <http://senafti.budiluhur.ac.id/index.php/senafti/article/view/871%0Ahttps://senafti.budiluhur.ac.id/index.php/senafti/article/download/871/549>
- Anggraiwan, Y., & Siregar, B. (2022). Klasifikasi Harga Mobil Menggunakan Metode Decision Tree Algoritma C4.5. *Computatio : Journal of Computer Science and Information Systems*, 6(2), 70–79. <https://doi.org/10.24912/computatio.v6i2.19994>
- Aryanti, P. G., & Santoso, I. (2023). Analisis Sentimen Pada Twitter Terhadap Mobil Listrik Menggunakan Algoritma Naive Bayes. *IKRA-ITH Informatika : Jurnal Komputer Dan Informatika*, 7(2), 133–137. <https://journals.upi-yai.ac.id/index.php/ikraith-informatika/article/view/2821>
- Martinus, A., & Pradiani, T. (2023). Pengaruh Brand Image dan Harga Terhadap Minat Beli Mobil Listrik Melalui Gaya Hidup Sebagai Variabel Intervening Di Area Kota Malang Jawa Timur *PENDAHULUAN Menurut Simbolon et al ., (2022) fenomena kendaraan listrik yang semakin berkembang membuat pemer.* 7(2).
- Zahra, N., & Akbar, A. (2024). Analisis Sales Forecasting Kendaraan Mobil Listrik Model Battery Electric Vehicle Di Indonesia (Metode Least Square). *Jurnal Education and Development*, 12(1), 197–205. <https://journal.ipts.ac.id/index.php/ED/article/view/5727>
- Caswadi, C., Dienwati, N., Dwilestari, G., Fathurrohman, F., & Tohidi, E. (2024). Penerapan Algoritma Naïve Bayes Dan K-Nearest Neighbor Untuk Analisis Sentimen Youtube Mengenai Intensif Mobil Listrik. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6), 3851–3857. <https://doi.org/10.36040/jati.v7i6.8252>
- Karimah, A., Dwilestari, G., & Mulyawan, M. (2024). Analisis Sentimen Komentar Video Mobil Listrik Di Platform Youtube Dengan Metode Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 767–737. <https://doi.org/10.36040/jati.v8i1.8373>
- Ernawati, N. W., Satya Kumara, I. N., & Setiawan, W. (2023). Perbandingan Metode Klasifikasi Support Vector Machine Dan Naïve Bayes Pada Analisis Sentimen Kendaraan Listrik. *Jurnal SPEKTRUM*, 10(3), 106. <https://doi.org/10.24843/spektrum.2023.v10.i03.p12>

- Muhammad Zaynurroyhan, Asriyanik, & Agung Pambudi. (2023). Perbandingan TF-IDF dengan Count Vectorization Dalam Content-Based Filtering Rekomendasi Mobil Listrik. *Explore IT: Jurnal Keilmuan Dan Aplikasi Teknik Informatika*, 15(1), 8–15. <https://doi.org/10.35891/explorit.v15i1.3829>
- Butsianto, S., & Mayangwulan, N. T. (2020). Penerapan *Data Mining* Untuk Prediksi Penjualan Mobil Menggunakan Metode *K-Means Clustering*. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 3(3), 187–201. <https://doi.org/10.32672/jnkti.v3i3.2428>
- Rohmah, A., Sembiring, F., & Erfina, A. (2021). Implementasi Algoritma *K-Means Clustering* Analysis untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: SMK Yaspim Gegerbitung). *Seminar Nasional Sistem Informasi Dan Manajemen Informatika*, 290–298. <https://sismatik.nusaputra.ac.id/index.php/sismatik/article/view/32>
- Sari, M. A. P., & Chotijah, U. (2022). Pengelompokan Anggota Divisi Himpunan Mahasiswa Jurusan Pada Universitas “Xyz” Dengan Metode *K-Means Clustering*. *Antivirus : Jurnal Ilmiah Teknik Informatika*, 16(1), 52–62. <https://doi.org/10.35457/antivirus.v16i1.2139>
- Reynaldo, J., Adikara, P. P., & Wihandika, R. C. (2020). Analisis Sentimen Mengenai Produk Toyota Avanza Menggunakan Metode Learning Vector Quantization Versi 3 (LVQ 3) dengan Seleksi Fitur Chi Square, Lexicon *Teknologi Informasi Dan ...*, 4(3), 830–839.
- Tiara Alifa, T., Astuti, R., & Muhamad Basysyar, F. (2024). Implementasi *Data Mining* Menggunakan Algoritma *K-Means Clustering* Dalam Analisis Penjualan Produk. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 602–607. <https://doi.org/10.36040/jati.v8i1.8340>
- Widodo, S., Ladyani, F., Asrianto, L. O., Rusdi, Khairunnisa, Lestari, S. M. P., Wijayanti, D. R., Devriany, A., Hidayat, A., Dalfian, Nurcahyati, S., Sjahriani, T., Armi, Widya, N., & Rogayah. (2023). *Buku Ajar Metode Penelitian*.
- Nuryaman, Y., Sudarsono, B., Faddillah, U., & Asistyasari, A. (2022). Klasifikasi Data Penjualan Menggunakan Algoritma *K-Means* Dan Analytic Hierarchy Process. *Jurnal Larik: Ladang Artikel Ilmu Komputer*, 2(1), 23–28. <https://doi.org/10.31294/larik.v2i1.1371>
- Waruwu, M. (2023). Pendekatan Penelitian Pendidikan: Metode Penelitian Kualitatif, Metode Penelitian Kuantitatif dan Metode Penelitian Kombinasi (Mixed Method). *Jurnal Pendidikan Tambusai*, 7(1), 2896–2910.
- Amin, N. F., Garancang, S., & Abunawas, K. (2023). KONSEP UMUM POPULASI DAN SAMPEL DALAM PENELITIAN. *Jurnal Pilar*, 14(1), 15–31.
- Anshor, A. H., & Rahardjo, S. B. (2023). Analisis Resiko Transformasi Mobil Berbahan Minyak ke Mobil Listrik Dengan Decision Tree. *Bulletin of*

Information Technology (BIT), 4(4), 414–418.
<https://doi.org/10.47065/bit.v4i4.924>

Adha, R. (2020). Analisis Pembelian Mobil Listrik Menggunakan Metode Analytical Hierarchy Process (AHP) dan Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). *Jist.Publikasiindonesia.Id*, 1(2), 476–485.
<https://doi.org/10.30865/klik.v4i1.1201>

Hasyim, A., Fatchan, M., & Hadikristanto, W. (2022). Penerapan Algoritma Naïve Bayes Dalam Memprediksi Tingkat Penjualan Mobil Tahun 2022. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, 4(02), 207–215.
<https://doi.org/10.46772/intech.v4i02.872>

M. Makbul. (2021). *Metode Pengumpulan Data dan Instrumen Penelitian*.
<https://doi.org/https://doi.org/10.31219/osf.io/svu73>

Syamfithriani, T. S., Mirantika, N., & Trisudarmo, R. (2023). Perbandingan Algoritma *K-Means* dan *K-Medoids* Untuk Pemetaan Daerah Penanganan Diare Pada Balita di Kabupaten Kuningan. *Jurnal Sistem Informasi Bisnis*, 12(2), 132–139. <https://doi.org/10.21456/vol12iss2pp132-139>

Benediktus, Y. K. (2024). *Ikhtiar Dorong Populasi Mobil Listrik*. Kompas.
<https://www.kompas.id/baca/ekonomi/2024/01/04/ikhtiar-dorong-populasi-mobil-listrik>

Dimas, A., & Kartika, D. H. (2024). *Penjualan Mobil Listrik Nasional Capai 17.000 Unit di Tahun 2023*. Kontan. <https://industri.kontan.co.id/news/penjualan-mobil-listrik-nasional-capai-17000-unit-di-tahun-2023>

Prawiro, D. S., Anggela, F. P., & Karawang, K. (2022). *371-Article Text-759-1-10-20220615*. 2(1).

DAFTAR RIWAYAT HIDUP

I. Biodata Mahasiswa

NIM : 19200242
Nama Lengkap : Farhan Jiddan
Tempat/ Tanggal Lahir : Jakarta, 11 Juni 2001
Alamat lengkap : Jl. Rawa Binong Gg. Annur 2B RT 002/010
Kelurahan Lubang Buaya, Kecamatan
Cipayung Jakarta Timur 13810

II. Pendidikan

a. Formal

1. SDS Angkasa 7 Halim P.K, lulus tahun 2013
2. SMPN 267 Jakarta Timur, lulus tahun 2016
3. SMK Budhiwarman 1, lulus tahun 2019

b. Tidak Formal

-

III. Riwayat Pengalaman berorganisasi / pekerjaan

1. Menjadi Video Editor di Kementerian Pertanian Republik Indonesia pada tahun 2017
2. UI Design untuk maidworks pada tahun 2020-2021



Jakarta, 02 Juli 2024

Hormat Saya

Farhan Jiddan

	LEMBAR KONSULTASI SKRIPSI
	UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19200242
 Nama Lengkap : Farhan Jiddan
 Asisten Pembimbing : Agustiena Merdekawati, ST, M.Kom
 Judul Skripsi : Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means

NO	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing
1	01 April 2024	Bimbingan perdana, pengajuan judul dan ACC judul	
2	18 April 2024	Pengajuan bab 1, pengajuan data & olah data yang akan diusulkan	
3	27 April 2024	Revisi bab 1 (Data yang diusulkan dan Roadmap penelitian)	
4	07 Mei 2024	ACC bab 1, dan pengajuan bab 2	
5	18 Mei 2024	Revisi bab 2 (Kerangka Penelitian)	
6	22 Mei 2024	ACC bab 2, dan pengajuan bab 3	
7	08 Juni 2024	Revisi bab 3 (koreksi perhitungan manual)	
8	22 Juni 2024	ACC bab 3, dan pengajuan bab 4	
9	27 Juni 2024	ACC bab 4 dan halaman depan dan belakang skripsi	
10	03 Juli 2024	ACC keseluruhan	

Catatan untuk Dosen Pembimbing.
Bimbingan Skripsi

- Dimulai pada tanggal : 01 April 2024
- Diakhiri pada tanggal : 03 Juli 2024
- Jumlah pertemuan bimbingan : 10

Disetujui oleh,
Dosen Pembimbing



(Agustiena Merdekawati, ST, M.Kom)



LEMBAR KONSULTASI SKRIPSI

UNIVERSITAS BINA SARANA INFORMATIKA

NIM : 19200242
Nama Lengkap : Farhan Jiddan
Asisten Pembimbing : Yuri Yuliani, M.Kom
Judul Skripsi : Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means

NO	Tanggal Bimbingan	Pokok Bahasan	Paraf Asisten Pembimbing
1	23 April 2024	Pengajuan bab 1, data yang digunakan untuk penelitian	
2	29 April 2024	Revisi bab 1, teknik penulisan	
3	13 Mei 2024	ACC bab 1, dan pengajuan bab 2	
4	20 Mei 2024	ACC bab 2	
5	28 Mei 2024	Pengajuan bab 3, revisi struktur penulisan	
6	10 Juni 2024	ACC bab 3, dan pengajuan bab 4	
7	28 Juni 2024	ACC bab 4	
8	04 Juli 2024	ACC keseluruhan	

Catatan untuk Asisten Pembimbing.

Bimbingan Skripsi

- Dimulai pada tanggal : 23 April 2024
- Diakhiri pada tanggal : 04 Juli 2024
- Jumlah pertemuan bimbingan : 8

Disetujui oleh,
Asisten Pembimbing

(Yuri Yuliani, M.Kom)

**SURAT PERNYATAAN KEBENARAN/KEABSAHAN DATA
HASIL RISET UNTUK KARYA ILMIAH**

Yang bertanda tangan di bawah ini, saya:

Nama : Farhan Jiddan
NIM : 19200242
Jenjang : Sarjana (S1)
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika
Perguruan Tinggi : Universitas Bina Sarana Informatika

Dengan ini menyatakan bahwa data dan atau informasi yang saya gunakan dalam penulisan karya ilmiah dengan judul "**Klasifikasi Tingkat Penjualan Mobil Listrik Di Indonesia Menggunakan Algoritma K-Means**" merupakan data dan atau informasi yang saya peroleh melalui hasil penelitian sendiri dan tidak didasarkan pada data atau informasi hasil riset dari perusahaan/instansi/lembaga manapun.

Saya bersedia untuk bertanggung jawab secara pribadi, tanpa melibatkan pihak **Universitas Bina Sarana Informatika**, atas materi/isi karya ilmiah tersebut, termasuk bertanggung jawab atas dampak atau kerugian yang timbul dalam bentuk akibat tindakan yang berkaitan dengan data dan atau informasi yang terdapat pada karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 08 Agustus 2024

Mengetahui, Yang menyatakan,

Dosen Pembimbing



Agustienna Merdekawati, ST, M.Kom



Farhan Jiddan

BUKTI HASIL PENGECEKAN PLAGIARISME

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Di era modern yang dipenuhi dengan kemajuan teknologi, dampak kemajuan ini sangat terasa di berbagai aspek kehidupan manusia, termasuk dalam industri otomotif. Mobil listrik adalah salah satu perkembangan terkini yang menonjol. Mobil ini menggunakan motor listrik sebagai penggerak utama dan menggunakan baterai sebagai sumber energi. Mobil listrik memiliki banyak keuntungan, termasuk efisiensi energi dan efek positif terhadap lingkungan. Karena setiap merek memiliki karakteristik tertentu yang membedakan produknya dari yang lain, merek tersebut dapat mempengaruhi minat pembeli. Permintaan akan barang berkualitas tinggi mendorong perusahaan untuk meningkatkan kualitas produk dan mempertahankan reputasi merek mereka. Dengan begitu Perusahaan dapat mengidentifikasi minat pelanggan dalam membeli mobil listrik dengan melihat statistik penjualannya.

Penjualan mobil listrik, juga dikenal sebagai mobil bertenaga listrik (BEV), di Indonesia mencatatkan angka sebanyak 17.062 unit pada periode Januari hingga Desember 2023, melonjak sebesar 65,22% pertahun dari 10.327 unit pada tahun sebelumnya. Data ini dikumpulkan oleh Gabungan Industri Kendaraan Bermotor

BAB III

HASIL DAN PEMBAHASAN

3.1 Perumusan Masalah dan Studi Litelatur

Untuk mengarahkan fokus penelitian dan menetapkan tujuan yang jelas, perumusan masalah adalah langkah awal dalam proses penelitian atau penyelesaian proyek. Langkah ini melibatkan identifikasi, definisi, dan penyempurnaan masalah yang akan dipecahkan.

Pada penelitian ini perumusan masalahnya sebagai berikut :

- Bagaimana pengelompokan tingkat pembelian mobil listrik di Indonesia dengan Metode K-means?
- Bagaimana hasil kualitas pengelompokan atau clustering yang didapatkan?
- Kapan musim atau periode tertentu dalam setahun di mana penjualan mobil dari brand tertentu cenderung tinggi atau rendah?

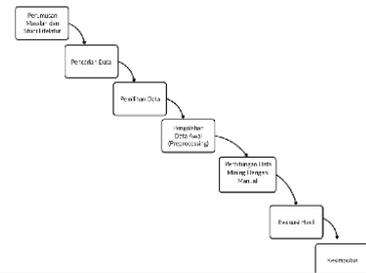
Pengumpulan, peninjauan, dan analisis berbagai sumber literatur yang relevan dengan topik penelitian dikenal sebagai studi literatur. Tujuan dari studi ini adalah untuk mendapatkan pemahaman yang lebih baik tentang apa yang sudah diketahui

BAB II

METODOLOGI PENELITIAN

2.1 Kerangka Penelitian atau Tahap Penelitian

Kerangka penelitian merupakan konseptualisasi yang terdiri dari sejumlah konsep, teori, metode, dan pendekatan yang digunakan, kerangka penelitian berfungsi sebagai struktur komprehensif yang membahas landasan teoritis dan metodologis penelitian serta mengarahkan langkah-langkah praktis yang diperlukan untuk menyusun, melaksanakan, dan menganalisis data untuk mencapai tujuan penelitian.



BAB IV

PENUTUP

4.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan oleh penulis, maka dapat ditarik kesimpulan sebagai berikut :

- Penelitian ini bertujuan untuk mengelompokkan tingkat pembelian mobil listrik di Indonesia pada tahun 2019-2023.
- Hasil Davies-Bouldin Index (DBI) dalam mengukur kualitas clustering menghasilkan nilai sebesar 0,099. Yang artinya menunjukkan bahwa hasil clustering yang dilakukan sangat baik.
- Hasil akurasi yang didapat dalam pengolahan penelitian menggunakan *Microsoft Excel* ini adalah:

Cluster 1 yang berjumlah 12 data, maka nilai akurasi 80%

Cluster 2 yang berjumlah 1 data, maka nilai akurasi 6,7%

Cluster 3 yang berjumlah 1 data, maka nilai akurasi 6,7%

Cluster 4 yang berjumlah 1 data, maka nilai akurasi 6,7%

- Hasil rata-rata *performance* yang didapat dalam pengolahan penelitian

BUKTI HASIL PENGECEKAN PLAGIARISME

