

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

ANALISIS DATA KUALITAS UDARA DI INDONESIA DENGAN K-MEANS CLUSTERING PADA BULAN AGUSTUS 2023

Chandra Alim¹, Herwis Gultom², Jazuli Mukhtar³

^{1,2,3} Prodi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang

e-mail: *1dosen02551@unpam.ac.id

ABSTRAK: Pencemaran udara merupakan salah satu permasalahan yang selalu mengganggu kesehatan lingkungan, seperti kabut asap yang terjadi setiap tahunnya di wilayah Indonesia yang mengganggu kesehatan manusia. Indonesia merupakan negara terpadat keempat di dunia, sehingga lingkungan yang sehat dibutuhkan bagi masyarakat Indonesia. Lingkungan yang sehat merupakan tempat dengan kualitas udara yang bersih dan nyaman untuk dihirup. Alat pembakaran rumah tangga, kendaraan bermotor, fasilitas industri dan kebakaran hutan merupakan sumber pencemaran udara yang umum. Polutan yang menjadi perhatian utama kesehatan masyarakat meliputi materi partikulat, karbon monoksida, ozon, nitrogen dioksida, dan sulfur dioksida. Polusi udara luar dan dalam ruangan menyebabkan penyakit pernapasan dan penyakit lainnya serta merupakan sumber penting kesakitan dan kematian. Untuk itu diterapkan ilmu dari data mining dengan metode K- means clustering yang dapat mengelompokkan polutan pencemar udara. Pengolahan Dataset dengan metode K- Means dengan tahapan yang menentukan atau mendeskripsikan nilai kuantitatif dari tingkat kemiripan atau ketidakmiripan data(propinquity measure) menggunakan bantuan dari aplikasi Rapidminer untuk menentukan jumlah cluster optimal dalam proses Cluster yang akan dilakukan perbandingan metode perhitungan jarak untuk menghasilkan cluster- cluster dari dataset yang ada. Pada tahapan awal dilakukan drawing data dengan mencari missing value, dan dilanjutkan dengan transformasi data dengan menormalisasikan data, proses normalisasi data juga merupakan proses mereduksi perhitungan komputasi yang terlalu besar maka dilakukan proses normalisasi data kedalam range yang lebih kecil yaitu skala 0,0 sampai 1,0 dengan menggunakan aplikasi Rapidminer untuk mempermudah proses menormalisasikan data.

Kata kunci: Pencemaran udara, K-Means Clustering, Rapidminer.

ABSTRACT: Air pollution is one of the problems that always disrupts environmental health, such as the smog that occurs every year in Indonesia which disrupts human health. Indonesia is the fourth most populous country in the world, so a healthy environment is needed for the Indonesian people. A healthy environment is a place with air quality that is clean and comfortable to breathe. Household combustion equipment, motor vehicles, industrial facilities and forest fires are common sources of air pollution. Pollutants of major public health concern include particulate matter, carbon monoxide, ozone, nitrogen dioxide, and sulfur dioxide. Outdoor and indoor air pollution causes respiratory and other illnesses and is an important source of morbidity and death. For this reason, knowledge from data mining is applied with the K-means clustering method which can group air pollutants. Dataset processing using the K-Means method with stages that determine or describe the quantitative value of the level of similarity or dissimilarity of the data (propinquity measure) using the help of the Rapidminer application to determine the optimal number of clusters in the Cluster process which will carry out a comparison of distance calculation methods to produce clusters from existing datasets. In the initial stage, data drawing is carried out by looking for missing values, and continued with data transformation by normalizing the data. The data normalization process is also a process of reducing computational calculations that are too large, so the data normalization process is carried out into a smaller range,

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

namely a scale of 0.0 to 1, 0 by using the Rapidminer application to simplify the process of normalizing data.

Keywords: Google Slides training, skills, education, community service...

21. PENDAHULUAN

Pencemaran udara merupakan salah satu permasalahan yang selalu mengganggu kesehatan lingkungan, seperti halnya kabut asap yang setiap tahun menyebar di Indonesia akhir-akhir ini yang mengganggu kesehatan masyarakat. Indonesia merupakan negara dengan jumlah penduduk terbesar keempat di dunia, sehingga masyarakat Indonesia membutuhkan lingkungan yang sehat. Lingkungan yang sehat adalah tempat yang kualitas udaranya bersih dan nyaman untuk dihirup. Udara bersih pada lingkungan yang sehat dihasilkan dari minimalnya polusi atau pencemaran udara pada lingkungan tersebut. Oleh karena itu, lingkungan yang sehat merupakan kebutuhan penting bagi semua orang. Peralatan pembakaran rumah tangga, kendaraan bermotor, pabrik industri dan kebakaran hutan merupakan sumber polusi udara yang umum. Polutan kesehatan masyarakat yang utama adalah partikel, karbon monoksida, ozon, nitrogen dioksida, dan sulfur dioksida. Polusi udara luar dan dalam ruangan menyebabkan penyakit pernapasan dan penyakit lainnya serta merupakan sumber kesakitan dan kematian yang signifikan.

Oleh karena itu, informasi yang diperoleh dari data mining diterapkan dengan metode K-means clustering yang dapat dikluster tanpa pengotor. Pengolahan data K-Means dengan langkah-langkah yang menentukan atau menggambarkan suatu nilai kuantitatif (proximity size) tingkat kemiripan atau ketidaksamaan antar data menggunakan aplikasi Rapidminer untuk menentukan jumlah cluster yang optimal dalam proses clustering. dibandingkan dengan metode perhitungan jarak untuk menghasilkan cluster dari dataset yang ada. Pada tahap awal pembersihan data dilakukan dengan mencari nilai yang hilang dan dilanjutkan dengan transformasi data melalui normalisasi data. Proses normalisasi data juga merupakan proses untuk mengurangi perhitungan komputer yang berlebihan, sehingga data dinormalisasi pada rentang yang lebih kecil yaitu skala 0.0-1.0, dengan menggunakan Rapidminer untuk mempermudah proses normalisasi data. Jakarta merupakan ibu kota Indonesia yang menjadi salah satu pusat kegiatan perekonomian di Indonesia. Menurut Survei Penduduk Antar Sensus (SUPAS), 10,61 juta orang akan tinggal di salah satu kota terpadat di Indonesia pada tahun 2021. Kepadatan penduduk dan pertumbuhan yang pesat menimbulkan permasalahan pencemaran, salah satunya pencemaran udara. Lingkungan yang padat menyebabkan degradasi udara akibat berbagai aktivitas antropogenik seperti aktivitas industri, peningkatan jumlah kendaraan pribadi, pembakaran lahan, asap rokok, dan lain-lain.

Menurut Organisasi Kesehatan Dunia (WHO), polusi udara adalah kontaminasi lingkungan dalam atau luar ruangan oleh zat kimia, fisik, atau biologis apa pun yang mengubah karakteristik alami atmosfer.

Data WHO menunjukkan bahwa hampir seluruh penduduk dunia (99%) menghirup udara yang melebihi batas pedoman WHO dan mengandung polutan tingkat tinggi, dengan negara-negara berpendapatan rendah dan menengah yang

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

paling banyak terpapar.

Kualitas udara berkaitan erat dengan iklim bumi dan ekosistem di seluruh dunia. Banyak penyebab polusi udara (misalnya pembakaran bahan bakar fosil) juga merupakan sumber gas rumah kaca. Oleh karena itu, kebijakan untuk mengurangi polusi udara memberikan strategi yang bermanfaat bagi iklim dan kesehatan dengan mengurangi penyakit terkait polusi udara dan berkontribusi terhadap mitigasi perubahan iklim dalam jangka pendek dan panjang.

Menurut National Institute for Occupational Safety and Health (NIOSH), terdapat lima sumber pencemaran dalam ruangan, yaitu (1) pencemaran dari dalam gedung berupa produk pembersih, pestisida, asap rokok, dan lain-lain, (2) pencemaran dari luar gedung, seperti tumpukan asap atau knalpot kendaraan bermotor, (3) kontaminasi dari bahan konstruksi seperti lem, fiberglass, asbes, formaldehida, dan bahan konstruksi lainnya, (4) kontaminasi dari mikroba seperti jamur, bakteri, protozoa, dan mikroba lain yang terdapat pada saluran udara, dan (5) udara segar akibat gangguan dan kurangnya pemeliharaan ventilasi. Selain itu, Tabel 1 mencantumkan penyebab polusi udara dalam ruangan (Adler, 2000).

Pembakaran bahan bakar di dalam rumah merupakan salah satu sumber pencemaran udara dalam ruangan. Hal ini dikarenakan bahan bakar yang digunakan untuk pemanasan dan memasak dapat menghasilkan nitrogen oksida (NO_x), karbon monoksida (CO), sulfur dioksida (SO₂), hidrokarbon dan partikulat (Kusnoputranto, 2002). NO_x terdiri dari nitrogen oksida (NO) dan nitrogen dioksida (NO₂). Menurut Fardiaz (1992), kedua bentuk NO_x tersebut mempunyai efek yang berbahaya bagi manusia. Kedua komponen tersebut menunjukkan bahwa NO₂ empat kali lebih toksik dibandingkan NO, padahal NO tidak berbahaya dan tidak menyebabkan iritasi. Pada konsentrasi udara normal, senyawa NO dapat teroksidasi dan menjadi lebih beracun. Sementara itu, CO merupakan polutan udara yang paling umum. Menurut Soedomo (2001), CO dihasilkan oleh pembakaran tidak sempurna bahan-bahan berbasis karbon seperti batu bara, kayu, bahan bakar minyak dan zat organik lainnya yang digunakan untuk pemanasan, pembakaran limbah dan proses industri. Menghirup karbon monoksida dapat menyebabkan sakit kepala yang cukup parah, kerusakan sel otak, pusing, koma bahkan kehilangan kesadaran (Kusnoputranto, 2002). SO₂ diproduksi dengan membakar bahan bakar fosil belerang dalam bahan mentah seperti batu bara, tembaga, aluminium, minyak bumi, besi dan timah. Sumber belerang di perkotaan adalah pembangkit listrik (Wahyuddin et al., 2016). Pada saat yang sama, partikel merupakan salah satu parameter pencemaran udara yang dapat mempengaruhi gangguan pernafasan manusia (Zannaria et al., 2009).

Dampak kesehatan dari polusi udara dalam ruangan termasuk masalah pernafasan, kanker dan bahkan masalah mata (Von Schirnding et al., 2002). Selain itu, polusi udara dalam ruangan juga dapat menyebabkan infeksi saluran pernafasan akut (ISPA), penyakit paru obstruktif kronik, kanker paru, asma, kanker nasofaring dan laring, TBC, berat badan lahir rendah (BBLR), dan penyakit mata (Smith et al., 2004) Dalam Peraturan Kementerian Kesehatan Republik Indonesia Nomor

1077/Menkes/Per/V/2011 tentang Pedoman Kesehatan Udara Dalam Ruangan disebutkan bahwa ISPA selalu masuk dalam sepuluh besar penyakit terbanyak di Indonesia. Hal ini menunjukkan bahwa dampak pencemaran udara berkisar dari dampak jangka pendek (iritasi mata dan tenggorokan) hingga dampak

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

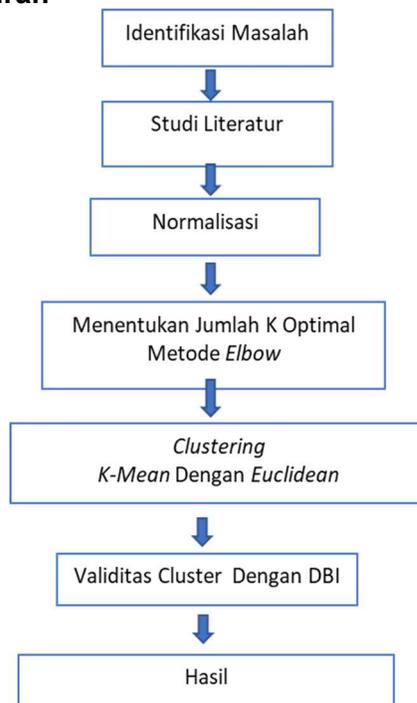
jangka panjang (penyakit pernafasan dan kanker).

Akibat dampak dan tingginya volume data pencemaran udara di Indonesia yang belakangan ini memburuk, hal ini menyebabkan peningkatan jumlah kasus ISPA yang dilaporkan di Indonesia. Analisis data kluster kualitas udara di Indonesia diperoleh dari data laporan pemantauan kualitas udara di website Kementerian Lingkungan Hidup

Permasalahan yang dihadapi adalah tingginya polusi udara di Indonesia yang menyebabkan peningkatan laporan masyarakat yang terkena ISPA. Penelitian ini bertujuan untuk mengidentifikasi masalah tersebut dengan menggunakan metode k-means untuk mengklusterisasi data kualitas udara di Indonesia. Ruang lingkup penelitian mencakup penggunaan algoritma k-means untuk mengkluster pola perburukan kualitas udara dengan data sekunder dari portal Kementerian Lingkungan Hidup pada bulan Agustus 2023. Rumusan masalah yang diangkat meliputi cara yang tepat untuk mengidentifikasi daerah dengan udara yang paling tercemar di Indonesia serta optimasi pencarian data pencemaran udara di Jakarta dalam bentuk kluster menggunakan metode k-means. Tujuan penelitian ini adalah untuk mengetahui data referensi kualitas udara di Indonesia pada bulan Agustus 2023 dan mencari hasil validasi nilai Davies-Bouldin Index terbaik.

22. METODA

Kerangka Pemikiran



Gambar 1. Kerangka Pemikiran

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

Keterangan :

1. Identifikasi Masalah

Identifikasi masalah dilakukan pada tahap awal penelitian untuk mencari serta mengidentifikasi masalah apa yang ditemukan. Masalah yang telah ditemukan yakni kualitas udara di Indonesia yang terus mengalami perburukan setiap harinya,

2. Studi literatur

Studi literatur digunakan sebagai bahan acuan dalam penelitian ini, yang bersumber dari jurnal-jurnal maupun ebook.

3. Normalisasi

Data yang berbentuk belum bisa diolah karena terdapat beberapa kesenjangan nilai data yang terlalu tinggi oleh karena itu data harus melalui proses normalisasi agar data tersebut bisa menghasilkan hasil yang valid dan sesuai.

4. Metode *Elbow*

Penentuan jumlah *cluster* paling optimal dilakukan dengan menggunakan metode *elbow*.

5. Clustering

Langkah selanjutnya dilakukan proses klusterisasi dengan metode *K-Means*, dengan menggunakan *distance measure* yaitu *euclidean distance*.

6. Validitas *Cluster* Dengan DBI

Proses selanjutnya dilakukan dengan perhitungan nilai validitas *cluster* menggunakan *davies bouldin index*

7. Hasil

Selanjutnya akan didapatkan hasil cluster

Analisis Kebutuhan

1. Metode Panellation Sampel

Sampel adalah sejumlah subjek yang merupakan bagian dari populasi yang mempunyai sifat yang sama (Siregar, 2019). Populasi merupakan semua nilai hasil perhitungan maupun pengukuran baik kuantitatif maupun kualitatif dari karakteristik tertentu mengenai sekelompok objek yang lengkap dan jelas. Adapun populasi dalam penelitian ini adalah seluruh data warga di Indonesia pada bulan Agustus 2023. Dimana terdapat 60 pemantau udara pada tiap titik di Indonesia

2. Metoda Pengumpulan Data

Peneliti menggunakan dataset German Credit Data yang bersifat publik (terbuka) dimana data-data yang diperoleh bersumber dari website lingkungan hidup. Selain itu informasi yang terdapat pada media elektronik menjadi salah satu referensi peneliti dalam memahami pergerakan data setiap harinya. Dataset yang diperoleh terdiri dari 12 variabel data sebagai berikut:

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

Tabel 1. Data yang akan dijadikan sampel Persebaran *Pencemaran udara*

Nama Field	Tipe	Keterangan
Tanggal	Numerik	Tanggal 1 Agustus sampai 30 Agustus 2023
Nama Lokasi	String	Nama lokasi stasiun pemantau udara di Indonesia
Index	Numerik	
Parameter	String	

Perancangan Penelitian

Penelitian ini menggunakan pendekatan deskriptif yang bertujuan untuk menggambarkan objek yang diteliti melalui data atau sampel yang telah dikumpulkan, diolah, dianalisis, dan disajikan secara sistematis. Metode deskriptif yang digunakan adalah studi kasus fenomena memburuknya kualitas udara di Indonesia, dengan tujuan memberikan gambaran detail mengenai latar belakang, sifat, dan karakteristik khas dari dataset yang diperoleh.

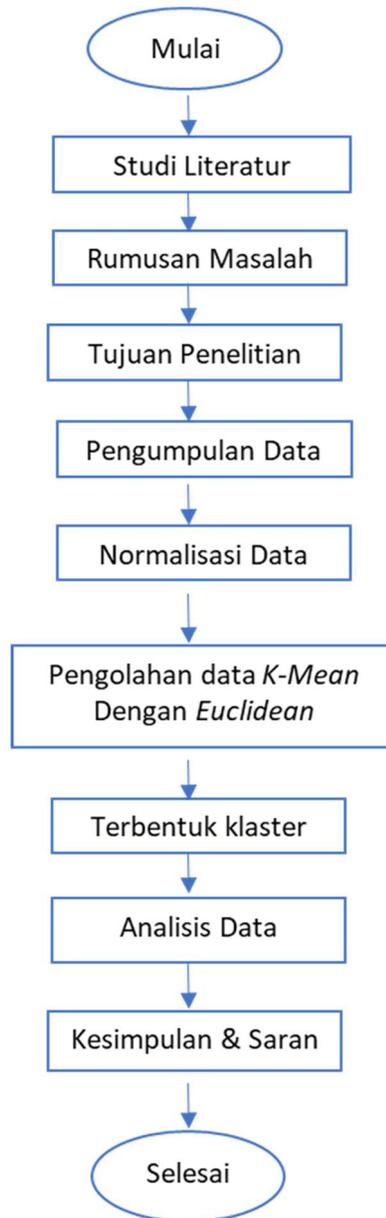
Penelitian ini juga memanfaatkan metode kuantitatif dengan pendekatan Unsupervised Learning, yang memungkinkan pengelompokan data ke dalam kluster tanpa memerlukan data latih. Data yang digunakan adalah data kuantitatif dari website lingkungan hidup, dan setelah melalui proses normalisasi, data diolah menggunakan metode clustering K-Means dengan bantuan aplikasi RapidMiner. RapidMiner membantu dalam analisis data mining, text mining, dan analisis prediksi untuk membentuk kluster kualitas udara di Indonesia.

Setelah kluster terbentuk, proses analisis data dilakukan untuk menentukan kesimpulan dan saran. Validasi kluster dilakukan menggunakan Davies Bouldin Index (DBI), yang mengukur validitas kluster dengan memaksimalkan jarak intra-kluster dan meminimalkan jarak antar kluster. Proses ini memastikan bahwa setiap objek dalam kluster memiliki karakteristik yang tinggi dan perbedaan antar kluster lebih jelas.

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024



Gambar 2. Perancangan Penelitian

23. HASIL DAN PEMBAHASAN

Pengolahan Data Awal

Dalam pengolahan dataset menggunakan metode K-Means, peneliti memanfaatkan aplikasi RapidMiner untuk menentukan jumlah cluster optimal dan melakukan perbandingan metode perhitungan jarak, seperti Euclidean distance. Pengukuran jarak ini penting untuk menentukan kemiripan atau ketidakmiripan data, yang berperan besar dalam proses clustering. Penentuan proximity measure atau

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

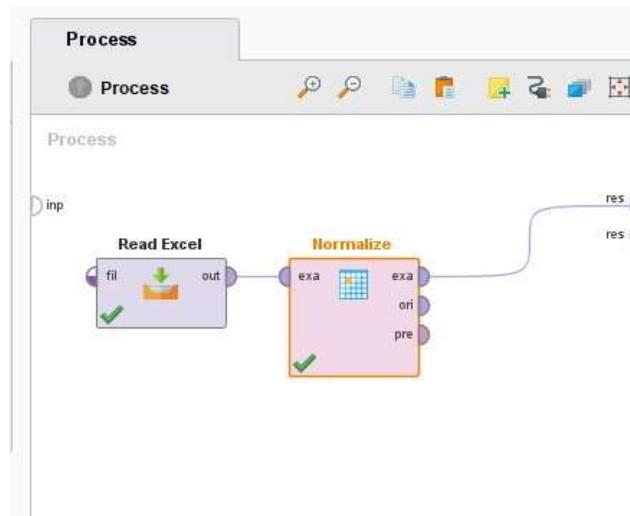
tingkat kemiripan data menjadi kunci dalam menghasilkan kluster yang akurat.

Proses pengolahan data dimulai dengan cleaning data untuk mencari missing value, yang dilanjutkan dengan normalisasi data menggunakan fitur Clustering pada RapidMiner. Normalisasi ini mereduksi perhitungan komputasi dengan mengubah data ke dalam rentang skala 0,0 hingga 1,0, yang mempermudah proses pengolahan data selanjutnya. Proses ini dilakukan melalui dua operator utama, yaitu Read Excel dan Normalize, untuk memastikan data siap digunakan dalam klusterisasi.



Name	Type	Missing	Statistics	Filter (2 / 2 attributes):
nama	Polynomial	0	Sum: yogyakarta [...] suman (1)	Min: DKI 1 Bundaran HI (1) Max: DKI 1 Bundaran HI (1), DKI
nilai ispu	Integer	0	Min: 16 Max: 258 Average: 59.667	

Gambar 3. Tidak Terdapat Missing Value



Gambar 4. Normalisasi Data

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024



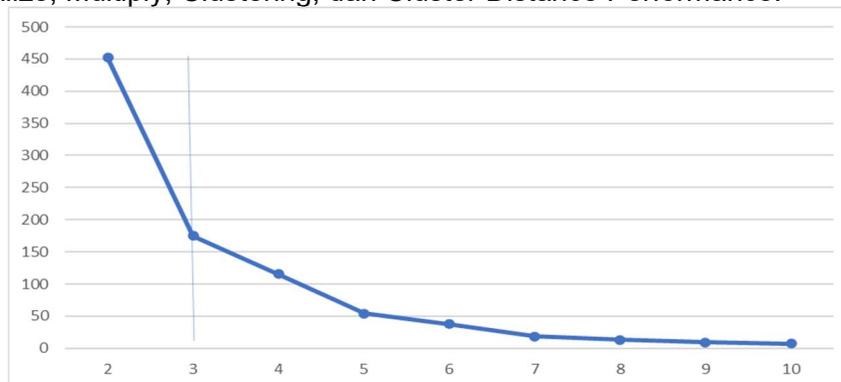
Row No.	nama	nilai ispu
46	ternate marik...	0.132
47	yogyakarta go...	0.140
48	balikpapan b...	0.029
49	balikpapan pl...	0.041
50	bantar gebang	0.281
51	banyuasin	0.116
52	cirebon 01	0.054
53	DKI 1 Bundar...	0.207
54	DKI 2 Kelapa...	0.174
55	DKI 3 Jagaka...	0.178
56	DKI 4 Lubang...	0.273
57	DKI 5 Kebun ...	0.120
58	kotawaringin ...	0.186
59	ogan ilir	0.182
60	sumur batu	0.277

ExampleSet (60 examples, 1 special attribute, 1 regular attribute)

Gambar 5. Hasil Normalisasi Data

Penentuan Jumlah Cluster Terbaik

Setelah menentukan jumlah sampel data dan melakukan normalisasi, langkah berikutnya adalah menentukan jumlah cluster yang tepat menggunakan aplikasi RapidMiner. Penulis menggunakan metode elbow untuk menentukan jumlah K terbaik, dengan melihat nilai Sum of Square Error (SSE) pada grafik. Uji coba dilakukan dengan 60 sampel data, menguji jumlah cluster dari K=2 hingga K=10. Sudut paling tajam pada kurva elbow digunakan sebagai acuan dalam pemilihan jumlah cluster, dengan bantuan enam operator seperti Read Excel/Retrieve, Normalize, Multiply, Clustering, dan Cluster Distance Performance.



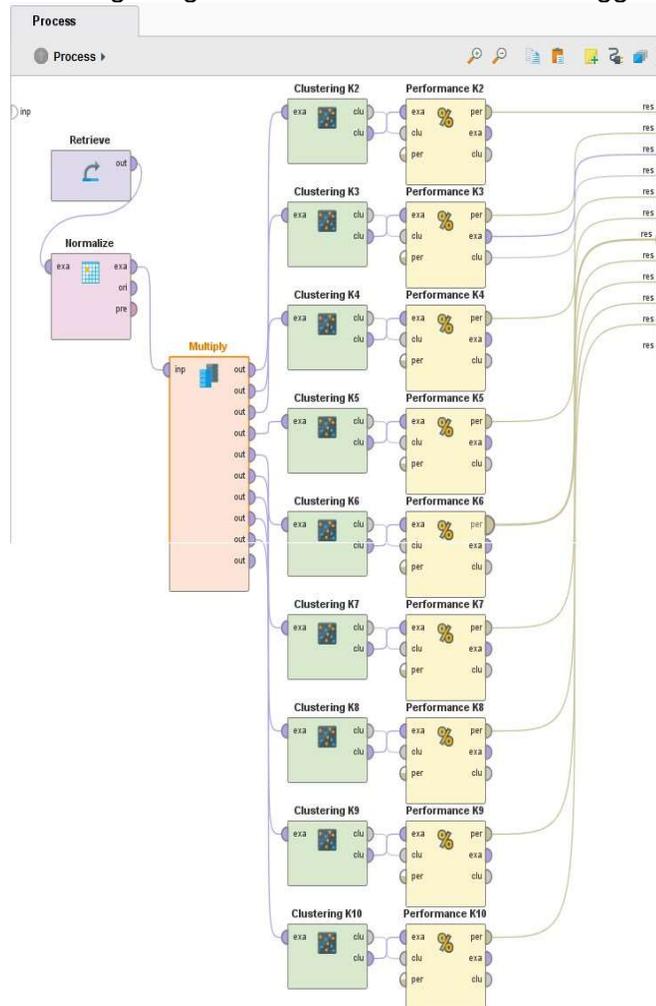
Gambar 6. Data proses pencarian cluster

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

Setelah normalisasi data, langkah selanjutnya adalah menambahkan multiply untuk menguji beberapa cluster, dalam hal ini 9 clustering dan 9 performance. Dari hasil simulasi, K=3 menunjukkan nilai selisih tertinggi dan sudut siku tertajam pada grafik, sehingga disimpulkan bahwa jumlah cluster terbaik adalah K=3. Proses pengolahan data clustering dengan metode K-Means akan menggunakan tiga cluster



Gambar 7. Model Proses Cluster

24. SIMPULAN DAN SARAN

Dari hasil penelitian ini dapat disimpulkan beberapa hal sesuai tujuan dari penelitian yaitu Setelah melakukan perhitungan dan pengolahan data dengan menggunakan metode K-Means dalam pengklasteran data kualitas udara di wilayah indonesia dapat disimpulkan sebagai berikut:

1. Pola cluster data kualitas udara dikelompokkan menjadi 3 klaster terbaik, yang dibagi menjadi tidak sehat, sedang dan baik.
2. Didapatkan Cluster_0 dengan jumlah anggota sebanyak 26 dengan

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

- persentase 43% dengan kategori BAIK
3. Didapatkan Cluster_1 dengan jumlah anggota sebanyak 1 dengan persentase 2% dengan kategori TIDAK SEHAT
 4. Didapatkan Cluster_2 dengan jumlah anggota sebanyak 33 dengan persentase 55% dengan kategori SEDANG
 5. Hasil validasi nilai Davies Bouldin Index dengan menggunakan metode K-Means dengan distance measure menggunakan Euclidean dengan nilai 0,428

Hasil dari penelitian ini tentunya masih banyak hal yang bisa dikembangkan, Saran bagi peneliti berikutnya untuk mencapai hasil yang lebih baik diantaranya:

Dari analisa cluster perlu dilakukan lagi penelitian lanjutan agar clusterisasi data dapat dilakukan secara lebih valid dan cepat, dengan menggunakan metode lainnya maupun distance measure yang lain. Bisa juga dengan membangun aplikasi pengelolaan data pengklusteran data kualitas udara yang lebih rinci.

DAFTAR PUSTAKA

- Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan Pada RSUD Pekanbaru," *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- R. K. Dinata, Safwandi, N. Hasdyna, and N. Azizah, "Analisis K-Means Clustering pada Data Sepeda Motor," *INFORMAL Informatics J.*, vol. 5, no. 1, p. 10, 2020, doi: 10.19184/isj.v5i1.17071.
- C. Astria, A. P. Windarto, A. Wanto, and E. Irawan, "Metode K-Means pada Pengelompokan Wilayah Pendistribusian Listrik," *Semin. Nas. Sains Teknol. Inf.*, pp. 306–312, 2019.
- R. Adhitama, A. Burhanuddin, and R. Ananda, "Penentuan Jumlah Cluster Ideal Smk Di Jawa Tengah Dengan Metode X-Means Clustering Dan K-Means Clusterin," *JIKO (Jurnal Inform. dan Komputer)*, vol. 3, no. 1, pp. 1–5, 2020, doi: 10.33387/jiko.v3i1.1635.
- R. I. Fajriah, H. Sutisna, and B. K. Simpony, "Perbandingan Distance Space Manhattan Dengan Euclidean Pada K-Means Clustering Dalam Menentukan Promosi," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 4, no. 1, pp. 36–49, 2019.
- B. Yusuf, R. Mahara, H. Ahmadian, S. Wahyuni, and K. AR, "Analisis Clustering Penduduk Miskin Di Provinsi Aceh Menggunakan Algoritma K-Means Dan X-Means," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 1, pp. 26–35, 2022, doi: 10.32672/jnkti.v5i1.3961.
- M. Mughnyanti, S. Efendi, and M. Zarlis, "Analysis of determining centroid clustering x-means algorithm with davies-bouldin index evaluation," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 725, no. 1, 2020, doi: 10.1088/1757-899X/725/1/012128.
- A. Fira, C. Rozikin, and Garno, "Komparasi Algoritma K-Means dan K-Medoids Untuk Pengelompokan Penyebaran Covid-19 di Indonesia," *J. Appl. Informatics Comput.*, vol. 5, no. 2, 2021.

Prosiding Seminar Nasional Informatika Dan Sistem Informasi

P-ISSN 2549-4805

Volume 8, No. 1, Juni-Agus 2024

- N. P. E. Merliana, Ernawati, and A. J. Santoso, "Analisa Penentuan Jumlah Cluster Terbaik pada Metode K-Means," *Pros. Semin. Nas. MULTI DISIPLIN ILMU&CALL Pap. UNISBANK*, pp. 978–979.
- R. Nainggolan and G. Lumbantoruan, "Optimasi Performa Cluster K-Means Menggunakan Sum Of Squared Error (SSE)," vol. 2, no. 2, pp. 103–108, 2018.
- D. A. I. C. Dewi and D. A. K. Pramita, "Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali," *Matrix J. Manaj. Teknol. dan Inform.*, vol. 9, no. 3, pp. 102–109, 2019, doi: 10.31940/matrix.v9i3.1662.
- N. Nurhaliza and Mustakim, "Clustering of Data Covid-19 Cases in the World Using DBSCAN Algorithms," *Indones. J. Inform. Res. Softw. Eng.*, vol. 1, no. 1, pp. 1–8, 2021.
- V. Fitriyani, "Analisis Clustering Provinsi Indonesia Berdasarkan Persebaran Virus Corona (Covid-19) Menggunakan Algoritma K-Means," *Pros. Semin. Pendidik. Mat. ...*, vol. 3, no. 2721, 2021, [Online]. Available: <http://prosiding.himatikauny.org/index.php/prosidingism/article/view/143>
- Darmansah and N. W. Wardani, "Analisis Pesebaran Penularan Virus Corona di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 105–117, 2021, doi: 10.35957/jatisi.v8i1.590.
- E. Muningsih and S. Kiswati, "Sistem Aplikasi Berbasis Optimasi Metode Elbow Untuk Penentuan Clustering Pelanggan," *Joutica*, vol. 3, no. 1, p. 117, 2018, doi: 10.30736/jti.v3i1.196.
- M. Nishom, "Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square," *J. Inform. J. Pengemb. IT*, vol. 4, no. 1, pp. 20–24, 2019, doi: 10.30591/jpit.v4i1.