Penerapan K-Means Clustering Untuk Menentukan Jumlah Pengangguran Berdasarkan Umur (Studi Kasus Di Badan Statistik Provinsi DKI Jakarta 2020-2022)

by Andi Diah Kuswanto

Submission date: 04-Jul-2024 01:47PM (UTC+0700)
Submission ID: 2412397696
File name: REPEATER_-_VOL._2_NO._3_JULI_2024_Hal_135-146.pdf (1.56M)
Word count: 2998
Character count: 18597

Repeater : Publikasi Teknik Informatika dan Jaringan
Vol. 2 No. 3 Juli 2024
e-ISSN : 3046-7276, p-ISSN : 3046-7284, Hal.135-146
DOI: https://doi.org/10.62951/repeater.v2i3.116

Penerapan K-Means Clustering Untuk Menentukan Jumlah Pengangguran Berdasarkan Umur (Studi Kasus Di Badan Statistik Provinsi DKI Jakarta 2020-2022)

Andi Diah Kuswanto [1], Azumardi Nabil Fadhila [2], Paulus Tri Setiawan [3], Muhammad Kevin Setiawan [4], Dody Renal Syahputra[5]
[1-5] Universitas Bina Sarana Informatika

Korespondensi penulis : Andi.ahk@bsi.ac.id[1] , azumardinabilfadila0292@gmail.com[2] , Relegiosetiawan@gmail.com[3] , Kevinsetiawan322@gmail.com[4] , Syahputrarenal111@gmail.com[5]

***ABSTRACT**: Unemployment is a persistent problem in the labor market, thus hampering economic development and national prosperity. Indonesia, including its capital Jakarta, continues to face significant levels of unemployment compared to neighboring countries. This research focuses on analyzing the structure of unemployment in Jakarta using K-Means Clustering to categorize unemployment data based on age groups (2020-2022) sourced from the Central Statistics Agency. Analysis carried out via RapidMiner revealed three clusters:-Cluster 0: Age 30-60 years and above, Cluster 1: Age 20-24 years, Cluster 2: Age 15-19 and 25-29 years. The findings show that the 20-24 year age group has the highest unemployment rate (399,167 people), while the 30-60 year and above age group shows the lowest unemployment rate (75,560 people). This clustering approach provides insight into the distribution of unemployment by different age demographics in Jakarta, highlighting areas where targeted interventions may be needed to effectively address this socio-economic challenge*

***Keywords:** K-Means, Unemployment, Statistics*

ABSTRAK: Pengangguran merupakan masalah yang terus-menerus terjadi di pasar tenaga kerja, sehingga menghambat pembangunan ekonomi dan kesejahteraan nasional. Indonesia, termasuk ibu kotanya Jakarta, terus menghadapi tingkat pengangguran yang signifikan dibandingkan negara tetangga. Penelitian ini fokus menganalisis struktur pengangguran di Jakarta dengan menggunakan K-Means Clustering untuk mengkategorikan data pengangguran berdasarkan kelompok umur (2020-2022) yang bersumber dari Badan Pusat Statistik. Analisis yang dilakukan melalui RapidMiner mengungkapkan tiga cluster:-Cluster 0: Usia 30-60 tahun ke atas, Klaster 1: Usia 20-24 tahun, Klaster 2: Usia 15-19 dan 25-29 tahun. Temuan menunjukkan bahwa kelompok usia 20-24 tahun mempunyai tingkat pengangguran tertinggi (399.167 orang), sedangkan kelompok usia 30-60 tahun ke atas menunjukkan tingkat pengangguran terendah (75.560 orang). Pendekatan pengelompokan ini memberikan wawasan mengenai distribusi pengangguran berdasarkan demografi usia yang berbeda di Jakarta, menyoroti bidang-bidang di mana intervensi yang ditargetkan mungkin diperlukan untuk mengatasi tantangan sosio-ekonomi ini secara efektif

Kata kunci : K-Means, Pengangguran, Statistik

LATAR BELAKANG

Pengangguran merupakan masalah ketenagakerjaan yang selalu menjadi penghambat pembangunan ekonomi dan kesejahteraan suatu bangsa. Sampai saat ini masalah pengangguran terus melanda berbagai negara termasuk Indonesia. Berdasarkan fakta yang ada tingkat pengangguran di Indonesia masih cukup tinggi dibandingkan dengan beberapa negara tetangga. Pengangguran termasuk dalam permasalahan yang akan selalu diperhatikan dengan terus menerus terutama pada negara-negara berkembang, seperti Indonesia, terutama ibu kota Provinsi DKI Jakarta (Arianto Zega et al., 2022).

Received: Juni 30,2024 , Accepted: Juli 04,2024 , Published: Juli 31,2024
* Andi Diah Kuswanto, Andi.ahk@bsi.ac.id

Provinsi DKI Jakarta menghadapi permasalahan tingkat pengangguran yang tinggi seperti kota-kota di negara berkembang lainnya. Pengangguran terbuka menjadi masalah bagi perekonomian karena pengangguran menunjukkan ketidakseimbangan dalam perekonomian, yaitu kelebihan penawaran tenaga kerja. Penelitian ini bertujuan untuk menganalisis struktur pengangguran di Jakarta.

Untuk mengetahui jumlah penduduk yang terbanyak di wilayah Provinsi DKI Jakarta. Teknik clustering dapat membantu untuk mengelompokkan data secara otomatis tanpa perlu diberitahu label kelasnya. Para ahli banyak mengusulkan metode clustering, salah satunya adalah K-Means. Metode K-Means digunakan dalam berbagai aplikasi kecil hingga menengah dan merupakan algoritma klasterisasi yang paling banyak karena kemudahan dalam mengaplikasikannya. Algoritma K-Means, cluster yang dihasilkan cukup baik, sehingga metode ini bisa direkomendasikan sebuah clustering yang baik. Metode K-Means metode *not hierarchical clustering* yang berusaha membantu penyertaan variabel kelompok dalam kelas hasil perhitungan akhir (Di et al., 2024).

LANDASAN TEORI

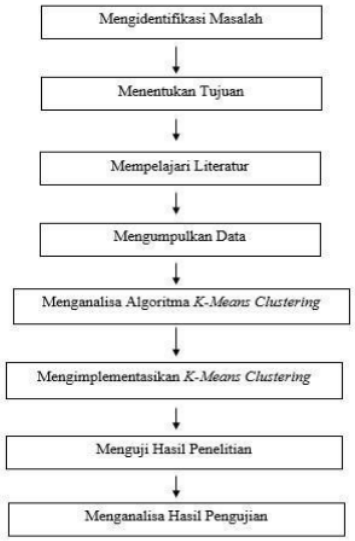
Data Mining

Data mining mulai ada sejak 1990-an sebagai cara yang benar dan tepat untuk mengambil pola dan informasi yang digunakan untuk menemukan hubungan antara data untuk melakukan pengelompokkan ke dalam satu atau lebih cluster sehingga objek-objek yang berada dalam satu cluster akan mempunyai kesamaan yang tinggi antara satu dengan lainnya (Dongga et al., 2023).

Data mining didefinisikan sebagai proses menemukan pola-pola dalam data. Proses ini otomatis atau seringnya semiotomatis. Pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi. Data yang dibutuhkan dalam jumlah besar. Data mining atau yang sering disebut juga *Knowledge Discovery in Database* (KDD) merupakan proses pengumpulan dan pengolahan data yang bertujuan untuk mengekstrak informasi penting pada kumpulan data. Proses pengumpulan dan ekstraksi informasi tersebut dapat dilakukan menggunakan perangkat lunak dengan bantuan perhitungan statistika, matematika, ataupun teknologi Artificial Intelligence (AI) (Canda Naya, 2023).

Tahap ini merupakan proses utama ketika menerapkan metode untuk menemukan pengetahuan yang berguna dari data. Jenis data dalam penelitian ini adalah clustering yang digunakan untuk mengelompokan data pengangguran berdasarkan tingkat pengangguran Provinsi DKI Jakarta(Pastia & Dikananda, 2023).

Adapun kerangka kerja yang digunakan dalam penulisan ini adalah seperti terlihat pada gambar 1.



Gambar 1 kerangka kerja 1

Data WareHouse

Data warehouse merupakan kumpulan dari struktur data yang saling berhubungan, setiap struktur memiliki Key Performance Indicator untuk proses bisnis tertentu, dan setiap Key Performance Indicator berhubungan dengan entiti tertentu. Hubungan ini akan menciptakan dimensional model. Data warehouse diperlukan bagi para pengambil keputusan manajemen dari suatu organisasi/perusahaan. Dengan adanya data warehouse, akan mempermudah pembuatan aplikasi DSS dan EIS karena memang kegunaan dari data warehouse adalah khusus untuk membuat suatu database yang dapat di gunakan. Data warehouse sangat penting karena semua informasi yang dibutuhkan sebuah perusahaan dapat ditemukan disini, oleh karena itu perusahaan yang ingin mengimplementasikan data warehouse harus dapat membuatnya dengan baik.

Langkah-langkah dalam membuat data warehouse:

1. *Determine business objective*
2. *Collect and analyze information*
3. *Identify core business processes*
4. *Construct a conceptual data model*
5. *Locate data source and plan data transformations*
6. *Set tracking duration*

7. *Implement the plan* (Rifky et al., 2021).

Metode Clustering

Clustering adalah suatu metode pengelompokan berdasarkan ukuran kedekatan. Perbedaan Clustering dengan grup, kalau grup berarti kelompok yang sama kondisinya kalau tidak ya pasti bukan kelompoknya. Tetapi kalau cluster tidak harus sama akan tetapi pengelompokannya berdasarkan kedekatan dari suatu karaktteristik sample yang ada, salah satunya dengan menggunakan rumuh jarak *euclidean*. Aplikasi cluster ini sangat banyak, karena hampir banyak dalam mengidentifikasi permasalahan atau pengambilan keputusan selalu tidak sama persis akan tetapi cenderung memiliki kemiripan saja (Al et al., 2022).

Algoritma K-Means

K-Means adalah sebuah metode klastering yang tergolong dalam pembelajaran tanpa pengawasan, yang aplikasikan untuk mengelompokkan data ke dalam beberapa kelompok melalui proses partisi. Dengan algoritma ini, data dapat diolah tanpa memerlukan label kelas. Inilah perbedaan *K-Means* dengan metode *KNearest Neighbor* (KNN), *Hierarchical Clustering* serta algoritma pembelajaran unsupervised learning lainnya yang membutuhkan vektor sebagai masukan. Algoritma ini akan mengklasifikasikan data atau objek ke dalam kelompok yang telah ditentukan. Setiap kelompok akan memiliki titik pusat yang disebut centroid yang merepresentasikan kelompok tersebut. Secara sederhana, algoritma *K-Means* adalah algoritma data mining yang digunakan untuk memecahkan masalah pengelompokan (*clustering*). Dalam proses pengolahan data menggunakan algoritma *K-Means Clustering*, langkah awalnya adalah menentukan centroid awal secara acak untuk setiap kelompok, kemudian dilakukan perhitungan berulang agar posisi centroid optimal. *K-Means* akan diproses dengan RapidMiner dengan perhitungan *Euclidean Distance* (Basalamah & Setyadi, 2023).

Rapid Minner

RapidMiner adalah aplikasi atau perangkat lunak yang berfungsi sebagai alat pembelajaran dalam ilmu data mining. Platfrom dikembangkan oleh perusahaan yang didedikasikan untuk semua langkah yang melibatkan sejumlah besar data dalam bisnis komersial, penelitian, pendidikan, pelatihan, dan pembelajaran. RapidMiner memiliki sekitar 100 solusi pembelajaran untuk pengelompokan, klasifikasi dan analisis regresi. RapidMiner juga mendukung sekitar 22 format file, seperti .xls, .csv, dan sebagainya. RapidMiner membawa kecerdasan buatan kepada perusahaan melalui platform ilmu data yang terbuka dan dapat diskalakan. RapidMiner dibangun untuk tim analisis, mengintegrasikan seluruh siklus ilmu

data, dari persiapan data hingga pembelajaran mesin hingga penyebaran model prediksi. Lebih dari 625.000 profesional analitis menggunakan produk RapidMiner untuk meningkatkan pendapatan, mengurangi biaya, dan menghindari risiko. RapidMiner adalah perangkat lunak independen yang digunakan untuk menganalisa data dan mesin penambangan data, yang dapat diintegrasikan dengan berbagai bahasa pemrograman secara mudah. RapidMiner ditulis dengan bahasa pemrograman Java, sehingga dapat berkerja pada banyak sistem operasi. RapidMiner menyediakan UI untuk mendesain pipa analisis, di mana akan menghasilkan file XML yang dapat menjelaskan proses analisis yang ingin diterapkan oleh pengguna ke data. RapidMiner akan membaca file ini untuk menjalankan analisa secara otomatis. RapidMiner menyediakan tampilan (UI) yang ramah pengguna, sehingga memudahkan pengguna saat menggunakannya. Tampilan yang terdapat pada RapidMiner disebut *Perspective. Terdapat 3 Perspective*, yaitu *Welcome Perspective, Design Perspective dan Result Perspective* (Prasetyo et al., 2021).

PEMBAHASAN

Data Pengangguran DKI Jakarta

Perhitungan di excel ini merupakan hasil data pengangguran provinsi DKI Jakarta dari tahun 2020-2022 berdasarkan kelompok umur yang terdiri dari umur 15-19, 20-24, 25-29. 30-34, 35-39, 40-44, 45-49, 50-54, 55-59 dan 60. Data ini diambil dari website badan pusat statistik provinsi DKI Jakarta yang di format dalam bentuk mirocsoft excel seperti contoh gambar 1.

Data Pengagguran DKI Jakarta Dari Tahun 2020 - 2022 Berdasarkan Kelompok Umur			
Kelompok Umur 15-60	pengangguran 2020	pengangguran 2021	pengangguran 2022
15 - 19	82870.00	59372.00	88056.00
20 - 24	148978.00	129362.00	120827.00
25 - 29	99279.00	70159.00	62144.00
30 - 34	56680.00	59991.00	24504.00
35 - 39	52818.00	32008.00	17116.00
40 - 44	30897.00	35385.00	16620.00
45 - 49	28041.00	26047.00	13142.00
50 - 54	29784.00	13701.00	13371.00
55 - 59	18645.00	6105.00	5477.00
60 +	24788.00	7769.00	16037.00

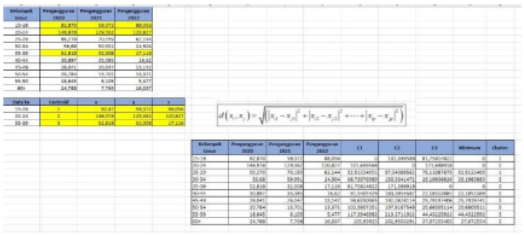
Gambar 1 Data pengganguran DKI Jakarta

Perhitungan Data Excel

Perhitungan data di excel merujuk pada penggunan metode clustering k-means. Yang pertama kita lakukan adalah menetukan data centroid secara random disini kita mengambil data umur 15-19, 20-24, 25-29. Yang kedua menghitung jarak antara objek dengan masing-masing centroid, bisa menggunakan rumus *Euclidean Distance*.

$$d(x_i, x_j) = \sqrt{\left(|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \cdots + |x_{ip} - x_{jp}|^2\right)}$$

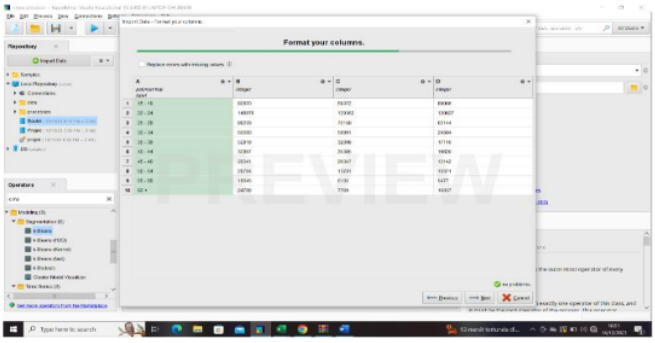
Yang ketiga menetukan nilai minimum dari masing-masing centroid. Dan yang keempat menentukan cluster dengan cara mengklasifikasikan nilai minimum berdasarkan jarak terdekat dengan centroid. Contoh nya seperti pada gambar 3.1.2.



Gambar 3.1.2. Perhitungan Data Excel

Format Data Rapid Minner

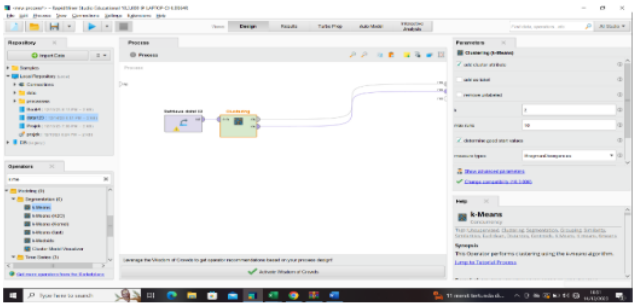
Pada format data rapid minner merupakan data pengangguran DKI Jakarta yang di format dari excel ke rapid miner. Yang pertama kali kita lakukan adalah melakukan import data melalui rapid miner setelah itu kita cari data excel yang ingin kita format setelah mengklik tombol next di bagian table a kita rubah role nya menjadi label sedangkan di table b, c dan d kita rubah type nya menjadi integer agar data nya menjadi bilangan bulat contoh nya seperti pada gambar 3.2.1.



Gambar 3.2.1 format data rapid minner

Tampilan Desain Rapid Minner

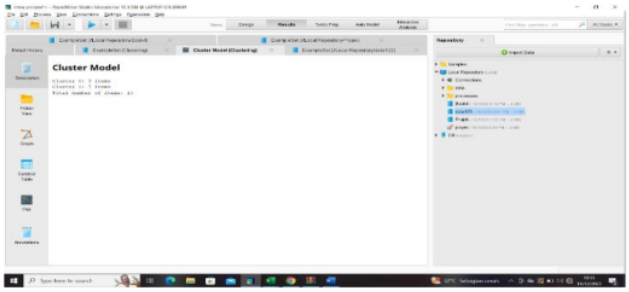
Pada tampilan desain rapid miner yaitu setelah sudah menginput data kita pergi ke tampilan desain untuk menghubungkan dengan metode k-means yang pertama kita lakukan kita masukan data pengangguran DKI Jakarta selanjutnya kita pilih operator yang kita gunakan yaitu k-means setelah itu kita hubungkan data pengangguran DKI Jakarta dengan k-means dan kita pilih menjadi 2 kelompok selanjutnya kita klik tombol run. Contoh hasil nya seperti pada gambar 3.2.2.



Gambar 3.2.2. Tampilan Desain Rapid Minner

Tampilan Cluster Model

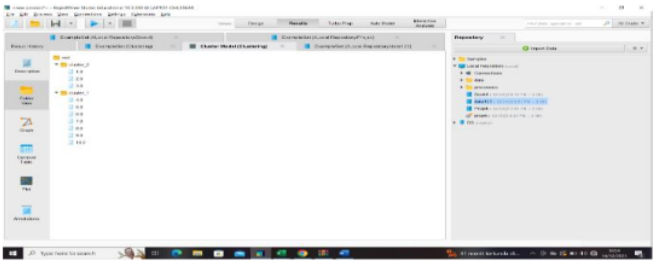
Pada tampilan cluster model yaitu setelah kita klik run kita bisa melihat data nya dengan mengklik cluster model dan data tersebut dibagi menjadi 2 kelompok yang pertama cluster0 terdiri dari 3 items dan kelompok yang kedua cluster1 terdiri dari 10 items dengan total items yaitu 10. Cotoh hasil nya seperti pada gambar 3.2.3.



Gambar 3.2.3. Tampilan Cluster Model

Tampilan Folder View

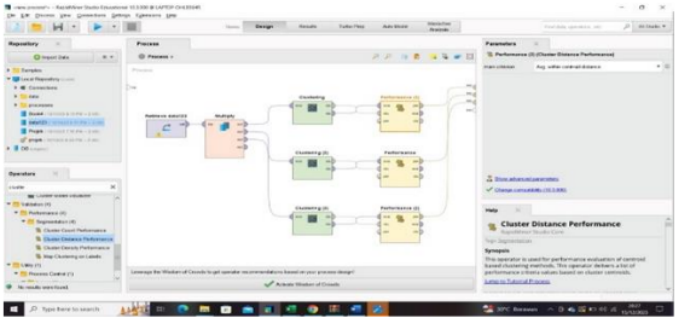
Pada tampilan folder view pada tampilan tersebut kita bisa melihat kelompok 1 yaitu cluster0 terdiri data 1, 2 dan 3 sedangkan kelompok 2 yaitu cluster1 terdiri dari data 4, 5, 6, 7, 8, 9 dan 10. Contoh hasil nya seperti pada gambar 3.2.4.



Gambar 3.2.4. Tampilan Folder View

Tampilan Performance

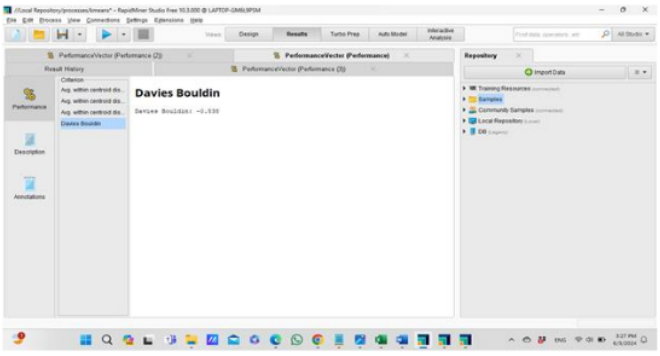
Pada tampilan ini kita ingin melihat *performance* pada data pengagguran DKI Jakarta mungkin saja menjadikan data ini menjadi 2 kelompok bukan pilihan yang terbaik mungkin saja 3 dan 4. Caranya adalah kita tambah operator multiply selanjutnya kita hubungkan data pengagguran DKI Jakarta dengan data cluster yang kita buat menjadi 2, 3 dan 4 kelompok selanjutnya untuk melihat performance kita tambahkan operatornya yaitu *cluster distance performance* selanjutnya kita hubungkan dengan cara mengsilangkan clustering dengan performancenya kemudian kita hubungkan permormancenya ke resort. Contoh hasil nya seperti pada gambar 3.2.5.



Gambar 3.2.5. Tampilan Performance

Tampilan Davies Bouldin Pada Kelompok 2

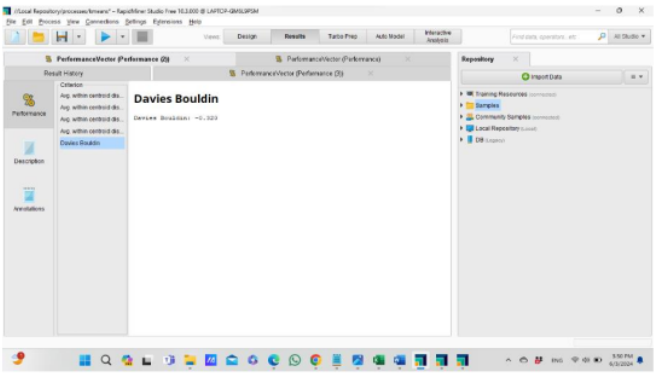
Setelah melakukan *running* kita lihat pada tampilan davies bouldin pada kelompok 2 yaitu 0.538 disini kita abaikan nilai minus nya karena ini bentuknya adalah absolut. Contoh hasil nya seperti pada gambar 3.2.6.



Gambar 3.2.6. Tampilan Davies Bouldin Kelompok 2

Tampilan Davies Bouldin Kelompok 3

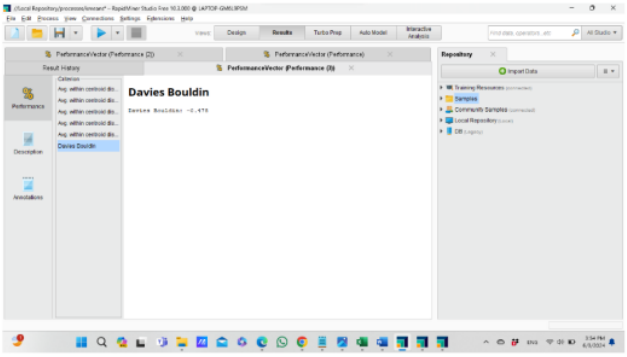
Setelah melakukan *running* kita lihat pada tampilan *davies bouldin* pada kelompok 2 yaitu 0.320 disini kita abaikan nilai minus nya karena ini bentuknya adalah *absolut*. Contoh hasil nya seperti pada gambar 3.2.7.



Gambar 3.2.7. Tampilan Davies Bouldin Kelompok 3

Tampilan Davies Bouldin Kelompok 4

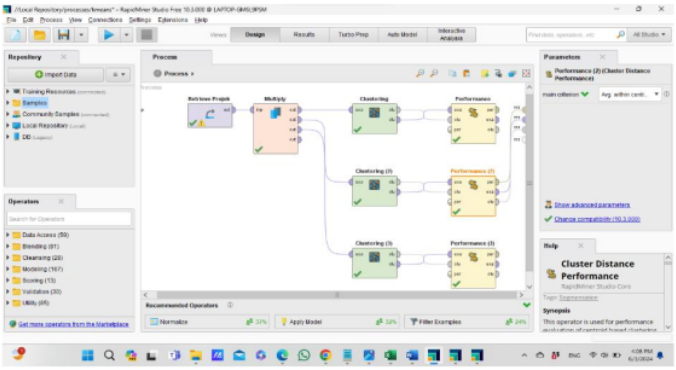
Setelah melakukan running kita lihat pada tampilan davies bouldin pada kelompok 4 yaitu 0.470 disini kita abaikan nilai minus nya karena ini bentuknya adalah absolut. Contoh hasil nya seperti pada gambar 3.2.8.



Gambar 3.2.8. Tampilan Davies Bouldin Kelompok 4

Tampilan Desain Kelompok 3

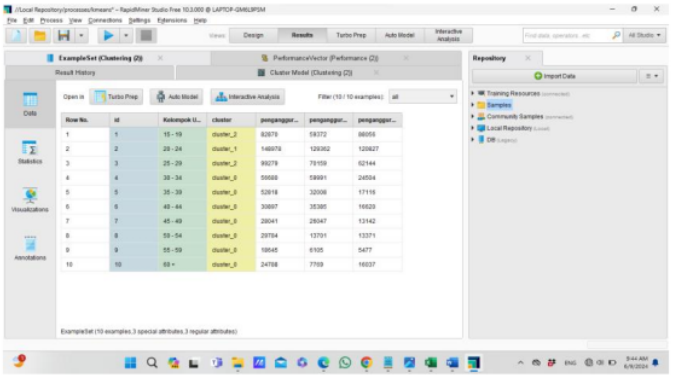
Karena pada kelompok 3 adalah nilai yang paling kecil yaitu 0.320 maka kelompok 3 adalah yang terbaik untuk dijadikan pengelompokkan. Disini kita tampilkan hasilnya dengan cara menghubungkan exa dan cluster nya ke res dan selanjutnya kita klik run. Contoh hasil nya seperti pada gambar 3.2.9.



Gambar 3.2.9. Tampilan Desain Kelompok 3

Tampilan Hasil Data Pada Kelompok 3

Setelah kita klik run muncul hasil cluster pada kelompok 3 yang terdiri dari cluster 0 yaittu kelompok umur 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60 ke atas. Sedangkan cluster 1 terdiri dari kelompok umur 20-24. Sedangkan cluster 2 terdiri dari kelompok umur 15-19, 25-29. Contoh nya seperti pada gambar 3.2.10.



Gambar 3.2.10. Tampilan Hasil Data

Hasil Hipotesa

Hasil hipotesa dari data pengganguran Dki Jakarta 2020-2022 berdasarkan rapid miner dan perhiyungan excel di atas adalah pada data ini cocok di buat dalam 3 kelompok yang terdiri dari cluster 0 yaitu umur 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60 ke atas. Sedangkan cluster 1 yaitu umur 20-24. Sedangkan cluster 2 yaitu umur 15-19, 25-29. Pada temuan data ini juga tingkat angka pengagguran tertinggi adalah cluster 1 dengan total 399.167 penduduk di tahun 2020-2022 pada umur 20-24 tahun. Sedangkan tingkat angka pengagguran terendah adalah cluster 0 dengan total 75.560 penduduk di tahun 2020-2022 pada umur 30-60 tahun ke

atas. Pada tingkat angka pengagguran sedang adalah cluster 2 dengan total 230.940 di tahun 2020-2022 pada umur 15-19 dan 25-29 tahun.

KESIMPULAN DAN SARAN

Kesimpulan

Penelitian ini menggunakan metode *K-Means Clustering* untuk menganalisis pengangguran di DKI Jakarta berdasarkan umur (2020-2022). Data dari Badan Pusat Statistik dianalisis dengan RapidMiner dan dikelompokkan menjadi tiga cluster:

- **Cluster 0**: Umur 30-60 ke atas.
- **Cluster 1**: Umur 20-24.
- **Cluster 2**: Umur 15-19 dan 25-29.

Hasilnya menunjukkan bahwa kelompok umur 20-24 memiliki tingkat pengangguran tertinggi (399.167), sedangkan kelompok umur 30-60 ke atas memiliki tingkat pengangguran terendah (75.560).

Saran

Untuk Mengurangi tingkat pengangguran di DKI Jakarta dapat dicapai dengan beberapa langkah strategis. Pertama, fokuskan pelatihan keterampilan untuk kelompok usia 20-24 tahun sesuai dengan kebutuhan pasar kerja. Kedua, tingkatkan akses pendidikan dan pelatihan vokasional untuk usia 15-19 dan 25-29 tahun. Ketiga, kembangkan program kewirausahaan bagi kaum muda. Keempat, perkuat kerjasama dengan industri untuk program magang dan pekerjaan paruh waktu. Kelima, manfaatkan teknologi untuk menciptakan lapangan kerja baru di sektor digital. Implementasi langkah-langkah ini diharapkan dapat secara signifikan mengurangi tingkat pengangguran di DKI Jakarta.

DAFTAR PUSTAKA

Al, U., Indonesia, A., Aisyah, O., Arsyad, T., Nurlatifah, M. B. A. H., Sunarmo, M. M., Program, M. S., Manajemen, S., Ekonomi, F., & Bisnis, D. (2022). Laporan akhir penelitian: Penerapan K-Means clustering dalam menentukan strategi promosi. November.

Arianto Zega, Y., Leona, M., Putra, S., Angelina, N., Phang, S., & Loo, E. (2022). Analisa kebijakan pemerintah terkait ancaman pengangguran pasca kenaikan inflasi. Jurnal Multidisiplin Indonesia, 1(4), 1121–1126. https://doi.org/10.58344/jmi.v1i4.108

Basalamah, A. T., & Setyadi, R. (2023). Penerapan algoritma K-Means clustering pada tingkat penyelesaian pendidikan di Provinsi Indonesia. Jurnal Informatika Dan Teknologi Komputer, 4(2), 114–121. https://ejurnalunsam.id/index.php/jicom/

Canda Naya, A. S. (2023). Implementasi data mining untuk pengelompokan pengangguran terbuka di Indonesia dengan metode clustering. Jurnal Teknologi Pelita Bangsa, 14(2), 99–104.

Di, P., Dki, P., & Handayanna, F. (2024). Penerapan algoritma K-Means untuk mengelompokkan kepadatan. Journal of Applied Computer Science and Technology (JACOST), 5(1), 50–55.

Dongga, J., Sarungallo, A., Koru, N., & Lante, G. (2023). Implementasi data mining menggunakan algoritma Apriori dalam menentukan persediaan barang (Studi kasus: Toko Swapen Jaya Manokwari). G-Tech: Jurnal Teknologi Terapan, 7(1), 119–126. https://doi.org/10.33379/gtech.v7i1.1938

Pastia, N. I., & Dikananda, F. N. (2023). Pengelompokan data pengangguran terbuka menggunakan algoritma K-Means berdasarkan Provinsi Jawa Barat. Jurnal Dinamika Informatika, 12(1), 59–69.

Prasetyo, V. R., Lazuardi, H., Mulyono, A. A., & Lauw, C. (2021). Penerapan aplikasi RapidMiner untuk prediksi nilai tukar rupiah terhadap US Dollar dengan metode linear regression. Jurnal Nasional Teknologi Dan Sistem Informasi, 7(1), 8–17. https://doi.org/10.25077/teknosi.v7i1.2021.8-17

Rifky, R. A., Musrini B, M., & Fitrianti F, N. (2021). Membangun data warehouse untuk menganalisa pola siswa yang mendaftar di ITENAS (Studi kasus Institut Teknologi Nasional Bandung). Jurnal Ilmiah Teknologi Infomasi Terapan, 8(1), 45–56. https://doi.org/10.33197/jitter.vol8.iss1.2

Penerapan K-Means Clustering Untuk Menentukan Jumlah Pengangguran Berdasarkan Umur (Studi Kasus Di Badan Statistik Provinsi DKI Jakarta 2020-2022)

ORIGINALITY REPORT

25%	23%	8%	7%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	journal.widyatama.ac.id Internet Source	4%
2	journal.isas.or.id Internet Source	3%
3	repository.unmuhjember.ac.id Internet Source	3%
4	journal.aptii.or.id Internet Source	2%
5	jurusan.tik.pnj.ac.id Internet Source	2%
6	repository.uin-suska.ac.id Internet Source	1%
7	lontar.ui.ac.id Internet Source	1%
8	repository.fe.unj.ac.id Internet Source	1%
	journal.unilak.ac.id	

9	Internet Source	1 %
10	e-journal.hamzanwadi.ac.id Internet Source	1 %
11	ejournal.seaninstitute.or.id Internet Source	1 %
12	Submitted to Universitas Sebelas Maret Student Paper	1 %
13	ejournal.itn.ac.id Internet Source	1 %
14	journal.stieamkop.ac.id Internet Source	1 %
15	Tri Bayu Pamungkas, Siti Maesaroh, Pebri Ardiansyah. "IMPLEMENTASI DATA MINING PADA STOK PENGGUNAAN BARANG DI GMF AEROASIA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING", Jurnal Ilmiah Sains dan Teknologi, 2023 Publication	1 %
16	www.neliti.com Internet Source	1 %

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		