

Implementasi Data Mining Untuk Jumlah Penduduk Berdasarkan Usia Menurut Kabupaten/Kota Menggunakan Algoritma K-Means

Siti Aminah¹, Mulyawan²

¹e-mail: sumi09siti@gmail.com

²e-mail : wm7488748@gmail.com

¹ STMIK IKMI Cirebon

Abstract

Residents are all people who live in specific areas and regions at specific times, with a relationship between their number, age, marital status, religion, gender, birth, death, quality, mobility, and resilience in relation to the economy, social, cultural, and political rights and obligations. But over time, the distribution of the population can increase and decrease. An increase in the population of a city or district if it is not evenly distributed will have negative impacts on such things as economic growth and population density. So in this study, the population in districts and cities will be grouped according to age groups, namely: children, early adolescents, late adolescents, early adults, and late adults. Using data mining is one of the ideas to classify the number of residents in each region or city so that there is no overcrowding. The method of data collection that will be used for each resident of the regency is based on age, namely productive age (15–64 years), non-productive age (0–14 years), and senior age (> 65 years). By using the Knowledge Discovery Data data mining technique, the results obtained from manual calculations using Microsoft Excel and RapidMiner have the same results, namely that for cluster 0 (C0), there are 106 age category data, cluster 1 (C1), 24 age category data, and cluster 2 (C2), as many as 94 age category data, with Davies-Bouldin results of 0.157.

Keywords: Data mining, K-means, Clustering, Knowledge Discovery Data

Abstrak

Penduduk adalah semua orang yang bertempat tinggal di wilayah dan perwilayahan tertentu pada waktu tertentu dengan keterkaitan hubungan antar jumlah,usia,status perkawinan, agama, jenis kelamin, kelahiran, kematian, kualitas, mobilitas, dan juga ketahanan yang berkaitan dengan ekonomi, social, budaya, dan perpolitik yang hak dan kewajiban. Namun seiring berjalannya waktu, persebaran penduduk bisa bertambah dan berkurang. Peningkatan jumlah penduduk suatu kota ataupun kabupaten jika tidak merata akan menimbulkan dampak negatif seperti pertumbuhan ekonomi dan kepadatan penduduk. Maka dalam penelitian ini akan dilakukan pengelompokkan jumlah penduduk di kabupaten/kota menurut kelompok umur yaitu: Kanak-kanak, Remaja Awal, Remaja Akhir, Dewasa Awal, dan Dewasa Akhir. Menggunakan data mining merupakan salah satu ide data untuk mengklasifikasikan jumlah penduduk yang ada di setiap Kabupaten/Kota agar tidak terjadi kepadatan penduduk. Metode pengumpulan data tiap penduduk Kabupaten yang akan dikumpulkan berdasarkan usia yaitu usia produktif (15-64 tahun), usia nonproduktif (0-14 tahun) dan usia manula (>65 tahun). Dengan menggunakan data

mining teknik Knowledge Discovery Data didapatkan hasil dari perhitungan manual menggunakan Microsoft excel dan rapidminer hasilnya sama, yaitu untuk cluster 0 (C0) sebanyak 106 data kategori usia, cluster 1 (C1) 24 data kategori usia, dan cluster 2 (C2) sebanyak 94 data kategori usia, dengan hasil davies bouldin 0.157.

Keywords: Data mining, K-means, Clustering, Knowledge Discovery Data

1. Pendahuluan

Badan Pusat Statistik Indonesia menjelaskan bahwa penduduk adalah jumlah orang yang hidup di daerah setiap saat dan ada konsekuensinya proses demografi seperti fertilitas, mortalitas, dan migrasi. Populasi terdiri dari orang dan dalam relasi individu atau kelompok yang menduduki wilayah tersebut di negara-negara tertentu selama setidaknya satu tahun pada saat pendataan atau pencacahan. (Saraswati 2018)

Ketika populasi suatu kota atau lingkungan bertambah, jika tidak merata maka akan menimbulkan dampak negatif seperti tingkat pengangguran, meningkatnya polusi, pertumbuhan ekonomi, dan juga pertumbuhan ekonomi. Pengelompokan populasi dalam penelitian ini berdasarkan umur, yaitu: Anak-anak, remaja awal, remaja akhir, dewasa awal dan akhir, tua, lebih tua, dan manula mendapatkan informasi kependudukan berdasarkan wilayah pada tahun 2021. Dengan menggunakan informasi tersebut, peneliti mengklasifikasikan populasi ke dalam kelompok umur anak-anak, remaja, dan dewasa.

Data mining adalah teknik umum untuk mengklasifikasikan penduduk setiap negara bagian/kota administratif dengan cara yang tidak menunjukkan kepadatan penduduk, dengan menggunakan data populasi negara bagian yang dikumpulkan berdasarkan usia berisiko, yaitu usia kerja (15-64 tahun), usia kerja (0-14 tahun) dan lebih tua (>65 tahun). Dilihat dari gambarannya, informasi dasar dikelompokkan menjadi beberapa teknik yaitu deskripsi, kemudian evaluasi, prediksi, klasifikasi, pengelompokan dan asosiasi. Manajemen data adalah inti dari proses KDD (Knowledge Discovery in Database), dimana KDD adalah proses terorganisir untuk menemukan informasi dan model yang sesuai untuk data besar.

Penelitian terdahulu tentang pengelompokan jumlah penduduk dapat dilihat di jurnal Tekinkom, volume 2, edisi 2 Desember 2019, e-issn: 2621-3079 doi 10.37600/tekinkom.v2i2.115 yang berjudul mengelompokkan jumlah penduduk menurut umur menggunakan k-method dan menjelaskan bahwa pengendalian penduduk merupakan salah satu tugas pemerintah Indonesia. Pertambahan dan pergerakan penduduk di setiap daerah memungkinkan daerah tertentu untuk mengatasi perubahan jumlah penduduk yang sangat besar yang dapat mempengaruhi tingkat perekonomian daerah tersebut. Data diolah dengan metode k-means clustering yang sesuai dengan jumlah penduduk kota pematangsiantar per kecamatan. Dengan pengelompokan ini terlihat bahwa jumlah penduduk tiap kecamatan dipilah berdasarkan umur, sehingga memungkinkan program

kepegawaian yang lebih tepat diterapkan di pusat-pusat klaster. (Sonang, Purba, and Pardede 2019)

Data untuk penelitian ini yaitu berisi data jumlah penduduk berdasarkan kelompok umur dan kabupaten/kota di Provinsi Jawa Barat periode tahun 2021. Data ini terkait topik Kependudukan yang dihasilkan oleh Dinas Kependudukan Dan Pencatatan Sipil yang dikeluarkan dalam periode 1 tahun sekali. (Jabar 2021)

Tabel 1 Dataset Jumlah Penduduk

id	kode provinsi	nama provinsi	kode kabupaten kota	nama kabupaten kota	kelompok umur	jenis kelamin	jumlah penduduk	satuan	tahun
1	32	Jawa Barat	3201	Kabupaten Bogor	00-04	Laki-laki	212809	Jiwa	2021
2	32	Jawa Barat	3201	Kabupaten Bogor	00-04	Perempuan	198774	Jiwa	2021
3	32	Jawa Barat	3201	Kabupaten Bogor	05-09	Laki-laki	265642	Jiwa	2021
4	32	Jawa Barat	3201	Kabupaten Bogor	05-09	Perempuan	248026	Jiwa	2021
5	32	Jawa Barat	3201	Kabupaten Bogor	10-14	Laki-laki	258327	Jiwa	2021
....
860	32	Jawa Barat	3279	Kota Banjar	65-69	Perempuan	3524	Jiwa	2021
861	32	Jawa Barat	3279	Kota Banjar	70-74	Laki-laki	1966	Jiwa	2021
862	32	Jawa Barat	3279	Kota Banjar	70-74	Perempuan	2221	Jiwa	2021
863	32	Jawa Barat	3279	Kota Banjar	> 75	Laki-laki	2114	Jiwa	2021
864	32	Jawa Barat	3279	Kota Banjar	> 75	Perempuan	2903	Jiwa	2021

Masalah dalam penelitian ini adalah pertumbuhan populasi yang tidak merata dan ketidakakuratan data yang mengarah pada hal-hal buruk seperti kelebihan populasi dan polusi yang mempengaruhi ekonomi dan sumber daya manusia. Dalam hal data kependudukan, teknik clustering data mining sangat berguna untuk mengelompokkan data kependudukan berdasarkan informasi yang ada sehingga dapat digunakan sebagai dasar untuk memprediksi jumlah penduduk yang akan datang. .

Pengelompokan penduduk menurut umur bertujuan untuk memudahkan pencarian data kependudukan dan memudahkan pemerintah atau pemangku kepentingan dalam merencanakan bentuk dan arah pembangunan serta perencanaan sesuai dengan kemampuan penduduk. Melalui pengelompokan ini, pemerintah juga dapat menganalisis jumlah penduduk dan mengetahui berapa banyak anak yang perlu bersekolah, berapa tenaga pendidik dan guru yang mendukung kegiatan ini, berapa sarana dan prasarana yang dibutuhkan.

2. Rumusan Masalah

Berdasarkan pemaparan latar belakang diatas maka tugas akhir merumuskan beberapa masalah sebagai berikut :

- 1) Bagaimana cara pengelompokan jumlah penduduk dengan menggunakan algoritma k-means?

- 2) Bagaimana hasil dari pengelompokan jumlah penduduk dengan menggunakan algoritma k-means tersebut ?

3. Batasan Masalah

Batasan permasalahan yang diangkat dan dibahas pada perancangan ini yaitu :

- 1) Mengelompokkan jumlah penduduk berdasarkan usia dalam kategori anak-anak, remaja, dan dewasa.
- 2) Data yang diambil hanya 224 data dari 864, informasi yang dihasilkan mengenai jumlah penduduk, kelompok usia, dan kategori usia.

4. Tujuan

Dilihat dari rumusan masalah diatas, maka tujuan yang dapat diambil yaitu :

- 1) Untuk mengetahui kelompok pada jumlah penduduk untuk kategori anak-anak, remaja, dan dewasa.
- 2) Untuk mengetahui hasil pengelompokan dari data jumlah penduduk

5. Tinjauan Pustaka

a) Teori Data Mining

Data mining adalah inti dari proses KDD dan melibatkan algoritme tebakan yang memeriksa data, membuat model, dan menemukan pola yang tidak diketahui. KDD otomatis, dapat didefinisikan untuk mengatur proses untuk pengenalan yang benar, berguna dan pencocokan pola dalam kumpulan data yang besar dan kompleks. Data mining memecahkan masalah dengan menganalisis data dalam basis data, di mana data disimpan secara elektronik dan pencarian dilakukan secara otomatis, misalnya di komputer (Arumawadu et al., 2015). Kemudian Harryanto dan Hansun (2017) menggunakan aplikasi algoritma C4.5 untuk mengidentifikasi karyawan potensial suatu perusahaan dengan memprediksi dan mengklasifikasikan kandidat potensial suatu perusahaan dengan membangun pohon keputusan berdasarkan informasi yang ada. Knowledge Discovery in Database (KDD) adalah seluruh proses non-trivial untuk menemukan dan mengenali pola dalam data, di mana pola yang ditemukan valid, baru setelah itu dapat berguna dan dapat dipahami. (Putra and Wadisman 2018)

b) Teori Clustering

Clustering pada suatu data adalah suatu tahapan untuk menggolongkan himpunan data yang atribut kelasnya belum dideskripsikan, secara konsep clustering adalah untuk memaksimalkan meminimalkan kemiripan intra antar kelas. Sebagai contoh ada suatu himpunan obyek, proses pertama dapat di klasterisasi menjadi beberapa himpunan kelas selanjutnya menjadi sebuah himpunan beraturan sehingga dapat diturunkan berdasarkan kelompok klasifikasi tertentu. Cluster juga dapat diartikan sebagai kelompok. Maka analisis clustering pada dasarnya akan menghasilkan sejumlah cluster (kelompok). Sebelum kita melakukan analisa perlu diterapkan pemahaman bahwa suatu himpunan dari data tertentu sebenarnya sudah memiliki kemiripan

di antara anggotanya tersebut. Oleh karena itu, setiap anggota-anggota yang memiliki kemiripan karakteristik dikelompokkan ke dalam satu atau lebih dari suatu kelompok. Tujuan dari clustering data adalah untuk meminimalisasikan fungsi tujuan yang ditetapkan didalam proses clustering, dan umumnya selalu meminimalisasi variasi suatu cluster dan memaksimalkan variasi antar cluster. (Muliono and Sembiring 2019)

c) Teori Algoritma K-Means

Algoritma K-Means dimulai dengan pemilihan K secara acak, dimana K adalah jumlah cluster yang akan terbentuk. Kemudian secara acak menetapkan nilai K, menetapkan nilai sementara ke pusat cluster atau biasa dikenal dengan centroid, mean atau "rata-rata". Hitung jarak dari setiap data yang diberikan ke setiap centroid menggunakan rumus Euclidean sampai Anda menemukan jarak terdekat dari setiap data ke centroid. Klasifikasikan setiap titik data berdasarkan seberapa dekat dengan pusat massa. Lakukan langkah tersebut hingga nilai titik berat tidak berubah (stabil). (nurul rohmawati, sofi defiyanti 2015)

Rumus dan tahapan pada K-Means Clustering yaitu sebagai berikut :

- 1) Tentukan k sebagai banyaknya himpunan yang dapat dibentuk.
- 2) Hasilkan nilai acak untuk pusat cluster awal (titik tengah) hingga k.
- 3) Menghitung jarak semua data input ke setiap centroid menggunakan rumus jarak Euclidean sampai ditemukan jarak terdekat dari semua data ke centroid. Berikut persamaan jarak Euclid:

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2}$$

Dimana :

x_i : data kriteria,

μ_j : centroid pada cluster ke-j

Memperbaharui nilai centroid. Menurut Rahman dkk (2017) nilai centroid baru di peroleh dari rata-rata cluster yang bersangkutan dengan menggunakan rumus:

$$C_k = \frac{1}{nk} \sum d_i$$

Keterangan:

nk = jumlah data dalam cluster k

d_i = jumlah dari nilai jarak yang masuk dalam masing-masing cluster.

Melakukan perulangan dari langkah 2 hingga 5 sampai anggota tiap cluster tidak ada yang berubah.

Hitung jarak setiap data ke masing-masing centroid menggunakan persamaan Euclidean Distance yaitu:

$$d(P, Q) = \sqrt{\sum (x_j(P) - x_j(Q))^2}$$

Bersumber : (Ilmiah and Pendidikan 2022)

6. Kategori Usia

Kelompok umur mengandung banyak informasi penting tentang seseorang, seperti: Jenis Kelamin, Ras, dan Usia. Wajah seseorang dapat mengalami banyak perubahan sementara atau permanen. Contoh dari perubahan tersebut adalah kerutan. Menurut Hardiantara (2011), kerutan merupakan proses penuaan alami karena kulit menjadi lebih tipis, kering dan kurang elastis seiring bertambahnya usia. Usia manusia dapat dibagi menjadi interval atau kelompok, masing-masing mewakili tahap perkembangan manusia. Salah satu kelompok umur atau kelompok umur yang dipublikasikan oleh Departemen Kesehatan RI (2009) di website resmi depkes.go.id: anak ke-1 = 0-5 tahun, anak ke-2 = 6-11 tahun, pubertas dini = 12-16 tahun, pubertas terlambat = 17-25 tahun, dewasa awal = 26-35 tahun, dewasa akhir = 36-45 tahun, muda = 46-55 tahun, tua = 56-65 tahun, lebih tua = 65 tahun ke atas. (Matematika and Amin 2017)

7. Sumber data

Sumber data penelitian ini disediakan oleh Disdukcapil Provinsi Jawa Barat Kecamatan Coblong, Kota Bandung, Jawa Barat 40131. Dataset yang diekstrak adalah jumlah penduduk menurut kelompok umur dan subwilayah pada tahun 2021. Total data dari dataset ini adalah 864 data dan hanya diambil 224 data. Untuk mempelajari diakses pada tanggal 18 Januari 2023 sebagai bahan pendukung penelitian ini.

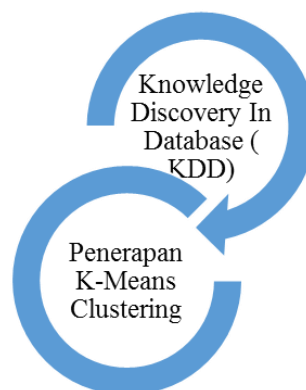
Data sekunder yaitu data yang diambil peneliti dari sumber-sumber yang ada. Sumber data sekunder yang digunakan dalam penelitian ini adalah data dari (Jabar, 2021). Informasi yang akan digali adalah data dan atribut penduduk menurut kelompok umur dan wilayah perkotaan tahun 2021 dalam format .xlsx. data yang ada hingga 864 data.

8. Teknik Pengumpulan data

Dalam penelitian ini teknik pengumpulan data adalah studi dokumentasi yaitu melakukan teknik penelitian online, sedangkan teknik penelitian melakukan pencarian data melalui internet melalui website Open Data Jabar.

9. Tahap Perancangan

Tahapan pada perancangan pada penelitian ini meliputi :



Gambar 1 Tahapan Perancangan

1) Proses KDD (Knowledge Discovery in Database)

Tahap desain yang digunakan dalam perancangan ini menggunakan desain Knowledge Discovery In Databases (KDD). Knowledge Discovery In Database (KDD) merupakan metode untuk mengumpulkan informasi dari database yang ada. Secara umum, KDD adalah proses yang bertujuan untuk memeriksa dan menganalisis kumpulan data dalam jumlah besar dan mengekstraksi informasi dan pengetahuan yang berguna. Proses KDD terdiri dari beberapa langkah, yaitu:

a. Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang akan digunakan untuk proses Data mining disimpan dalam suatu berkas, terpisah dari basis data operasional. (Yuli Mardi 2019)

Seleksi ini dilakukan untuk mendapatkan kelompok data berdasarkan atribut yang diperlukan seperti nama Kabupaten dan Kelompok Usia. Dalam perancangan tugas akhir ini terjadi seleksi data dari data jumlah penduduk berdasarkan usia. Total atribut yang ada sebanyak 9 atribut, sedangkan yang diambil untuk perancangan tugas akhir ini hanya 4 atribut data.

b. Pre-processing / Cleaning

Dalam proses data cleaning, data-data yang tidak relevan atau tidak sesuai akan dibersihkan agar data dapat digunakan untuk proses selanjutnya. Dalam perancangan tugas akhir ini terjadi tahapan cleaning data. Dimana atribut data sebanyak 5 dihapuskan, sehingga atribut yang dipakai untuk proses selanjutnya sebanyak 4 atribut.

c. Transformation Data

Pada fase ini, hasil dari data mining dijelaskan. Data atribut yang diperoleh dari proses pembersihan dan reduksi digunakan sebagai kumpulan data untuk uji desain dalam pekerjaan ini. Data penduduk berdasarkan perubahan kabupaten/kota diproses dengan Rapidminer.

d. Data Mining

Pada proses data mining, data akan diproses dan menghasilkan pattern.

e. Interpretation / Evaluation

Pada tahap ini, proses KDD dilakukan, data mining yang sudah dihasilkan adalah alur informasi sehingga dapat ditampilkan dengan tampilan yang sederhana agar bisa dipahami dan dimengerti.

2) Penerapan K-Means Clustering Pada Rapidminer

Pada tahap kedua yaitu penerapan K-Means Clustering pada Rapidminer. Tahap ini menerapkan algoritma K-Means Clustering untuk mengelompokkan jumlah penduduk kategori usia yaitu kanak-kanak, remaja, dan dewasa. Dalam penerapan ini menggunakan perhitungan Euclidean Distance yaitu perhitungan yang menghasilkan selisih jarak antara data satu dengan titik uji centroid. Jarak disini yang akan digunakan untuk dasar penentuan pengelompokan kategori usia.

10. Hasil Dan Pembahasan

Dalam tahap pengumpulan data, data yang digunakan yaitu data asli dari website opendata.jabarprov.go.id mengenai Jumlah Penduduk Berdasarkan Usia Dan Kabupaten Kota pada tahun 2021. Atribut yang ada pada data ini berjumlah 9 atribut terdiri dari Kode Provinsi, Nama Provinsi, Kode Kabupaten Kota, Nama Kabupaten Kota, Kelompok Umur, Jenis Kelamin, Jumlah Penduduk, Satuan, dan Tahun. Jumlah data keseluruhan sebanyak 864 baris data, tetapi data dan atribut yang digunakan untuk penelitian ini hanya 224 baris data dan 4 atribut yaitu Kategori Usia, Inisiasi Usia, Jenis Kelamin, dan Jumlah Penduduk.

1) Penerapan Algoritma K-Means Clustering Menggunakan Microsoft Excel

Implementasi algoritma K-means clustering dengan microsoft excel 2010 membutuhkan beberapa langkah dalam proses ini sebagai berikut:

a. Menentukan Pusat Cluster

Pusat cluster, atau bisa disebut centroid digunakan untuk menghitung jarak antara data di setiap cluster. Proses ini menentukan nilai titik cluster yang dapat ditentukan dengan metode acak sesuai dengan peneliti selama nilai centroid masih terdapat dalam nilai data di setiap atribut. Selain itu, nilai pusat cluster dapat ditentukan dengan menggunakan rata-rata dari setiap atribut.

Untuk perhitungan jarak ini menentukan kelompok usia yaitu kategori kanak-kanak, remaja, dan dewasa. Dari perhitungan jarak atau centroid diatas didapatkan 3 cluster dengan 2 hasil iterasi, yaitu :

1. Cluster 0 (C0), sebagai cluster tertinggi dengan jumlah data sebanyak 106 data.
2. Cluster 1 (C1), sebagai cluster terendah dengan jumlah data sebanyak 24 data.
3. Cluster 2 (C2), sebagai cluster sedang dengan jumlah data sebanyak 94 data.

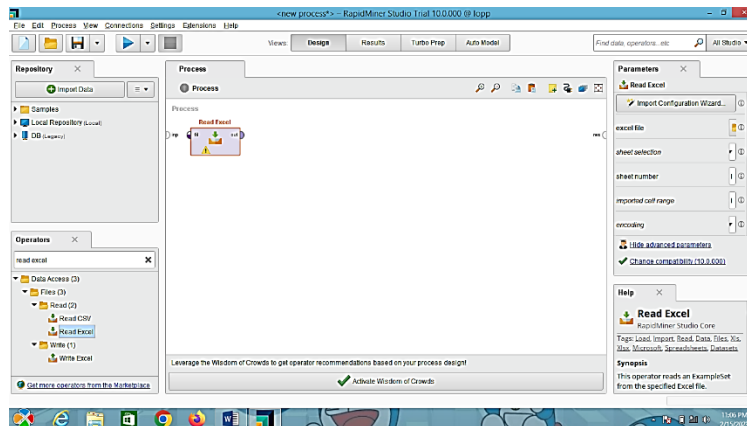
b. Menentukan Centroid Baru

Untuk menentukan centroid baru ini ditentukan dengan menghitung rata-rata cluster pada hasil iterasi 1.

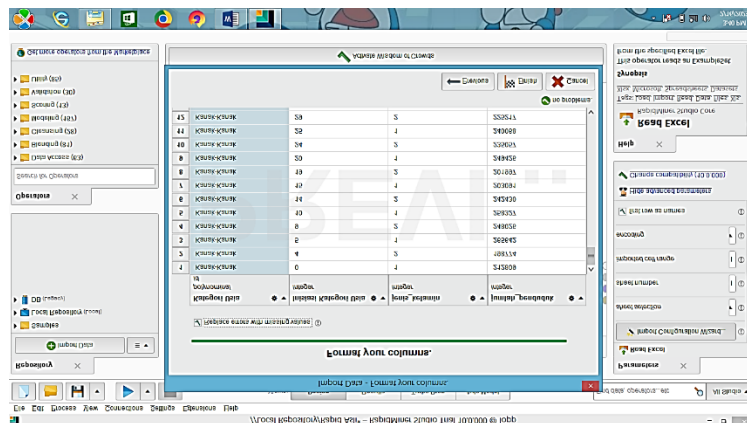
2) Hasil Pengelompokan Jumlah Penduduk Kategori Usia Menggunakan Rapidminer

a. Tampilan Read Excel Rapidminer

Pada gambar dibawah ini menjelaskan cara memasukan data baru yang akan diproses lebih lanjut, pada saat ini data yang akan diproses berupa data Excel. Klik pada bacaan read excel lalu pilih "Import Read Excel". Kemudian akan muncul seperti gambar dibawah ini.



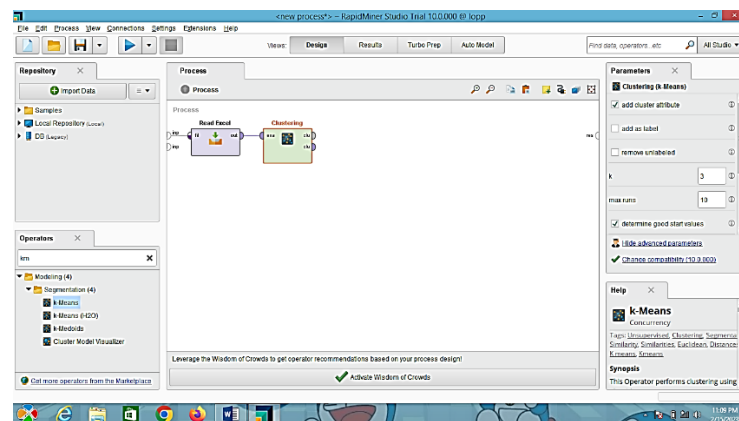
Gambar 2 Tampilan Read Excel



Gambar 3 Data Transformasi

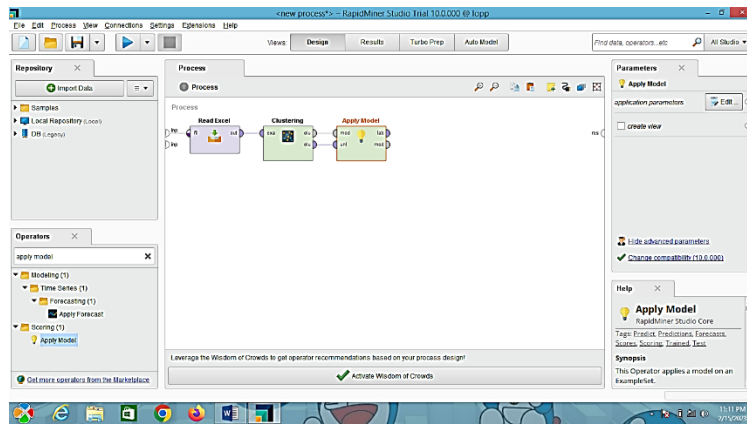
b. Pemrosesan Sistem Rapidminer

Tahapan pertama dengan mengklik k-means, setelah itu mengganti jumlah K sesuai dengan penelitian yang kita ambil. Pada penelitian ini jumlah K yang diambil yaitu sebanyak 3 cluster (K). Setelah selesai mengganti jumlah cluster (k), lalu sambungkan read excel ke clustering sehingga muncul tampilan seperti gambar di bawah ini.



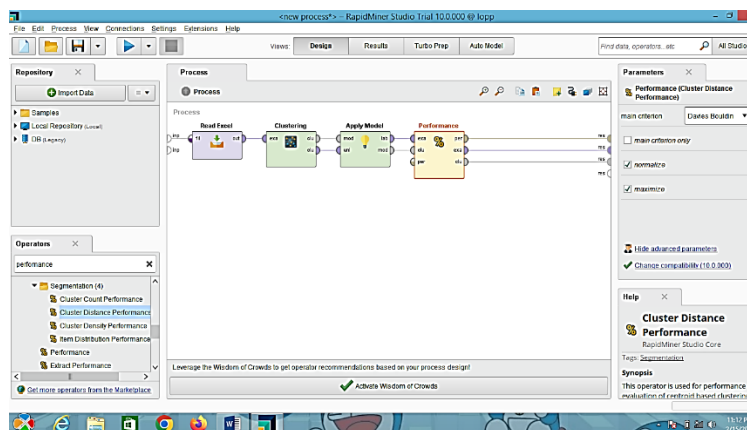
Gambar 4 Operator K-Means

Tahap Kedua yaitu dengan mengklik operator apply model.



Gambar 5 Operator Apply Model

Tahap ketiga yaitu klik operator performance (cluster distance performance). Setelah operator performance ditampilkan, lalu sambungkan dari operator apply model ke operator performance dan operator performance disambungkan seperti gambar berikut.



Gambar 6 Operator Performance

11. Kesimpulan

Berdasarkan pembahasan, mengenai pengelompokan jumlah penduduk menggunakan metode algoritma k-means clustering dapat disimpulkan bahwa :

1. Hasil dari penelitian yang diperoleh dapat disimpulkan bahwa pada perhitungan manual pada *Microsoft excel* dan dibantu oleh aplikasi rapidminer, bahwa metode ini mampu menyelesaikan pengelompokan berdasarkan kategori usia secara detail bagaimana cara menghitung jarak dan centroid dari cluster 1 dan berikutnya.
2. Hasil kedua yaitu diperoleh dari metode K-Means clustering, yang diterapkan ke rapidminer memiliki hasil yang sama. Akan tetapi pada rapidminer k-means hanya berlaku jika nilai min dapat dihitung jika atributnya bernilai integer

numerical saja, dan tidak bisa menentukan jarak centroid dari cluster 1 dan yang lainnya.

3. Hasil dari perhitungan manual menggunakan Microsoft excel dan rapidminer hasilnya sama, yaitu untuk cluster 0 (C0) sebanyak 106 data kategori usia sebagai cluster tertinggi, cluster 1 (C1) 24 data kategori usia sebagai cluster terendah, dan cluster 2 (C2) sebanyak 94 data kategori usia sebagai cluster sedang, dengan hasil davies bouldin 0.157.

12. Saran

Penulis berharap agar informasi yang telah disampaikan dapat bermanfaat bagi Dinas Kependudukan dan pemerintah yang bersangkutan dalam mengembangkan pengelompokan jumlah penduduk berdasarkan usia, dan bagi peneliti lainnya dapat lebih lanjut menggunakan beberapa algoritma pengelompokan lain untuk mencapai hasil yang optimal.

Daftar Pustaka

- Ilmiah, Jurnal, and Wahana Pendidikan. 2022. "Analisis K-Means Clustering Pada Pengiriman Produk Bearing Danendra Bima Adhi Pramana 1 , Tifani Amalina 2 , Riza Ibnu Adam 3 Teknik Informatika, Universitas Singaperbangsa Karawang." 8(September):128–37.
- Jabar, Open Data. 2021. "Jumlah Penduduk Berdasarkan Kelompok Umur Dan Kabupaten/Kota Di Jawa Barat." *Dinas Kependudukan Dan Pencatatan Sipil*. Retrieved February 18, 2023 (<https://opendata.jabarprov.go.id/id/dataset/jumlah-penduduk-berdasarkan-kelompok-umur-dan-kabupatenkota-di-jawa-barat>).
- Matematika, Jurnal Ilmiah, and Muchammad Al Amin. 2017. "MATH Unesa." 2(6).
- Muliono, Rizki, and Zulfikar Sembiring. 2019. "Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen." *CESS (Journal of Computer Engineering, System and Science)* 4(2):2502–2714.
- nurul rohmawati, sofi defiyanti, mohamad jajuli. 2015. "Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa." *Jitter 2015* I(2):62–68.
- Putra, Randi Rian, and Cendra Wadisman. 2018. "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means Implementation of Data Mining for Potential Customer Selection Using K-Means Algorithm." *Journal of Information Technology and Computer Science (INTECOMS)* 1(1):72–77.
- Saraswati, Mila. 2018. "Be Smart Ilmu Pengetahuan Sosial." *Grafindo Media Pratama* 15.

- Sonang, Sahat, Arifin Tua Purba, and Ferri Ojak Imanuel Pardede. 2019. “Pengelompokan Jumlah Penduduk Berdasarkan Kategori Usia Dengan Metode K-Means.” *Jurnal Teknik Informasi Dan Komputer (Tekinkom)* 2(2):166. doi: 10.37600/tekinkom.v2i2.115.
- Yuli Mardi. 2019. “Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data Mining Merupakan Bagian Dari Tahapan Proses Knowledge Discovery in Database (KDD) . Jurnal Edik Informatika.” *Jurnal Edik Informatika* 2(2):213–19.