

Data Mining

Implementasi Algoritma K-Means untuk Menentukan Angka Harapan Hidup berdasarkan Tingkat Provinsi

Yosinta Pratiwi, Mulyawan

Program Studi Manajemen Informatika, STMIK IKMI Cirebon, Cirebon, Indonesia

INFORMASI ARTIKEL

Diterima Redaksi: 18 Februari 2023
Revisi Akhir: 15 Maret 2023
Diterbitkan Online: 17 Maret 2023

KATA KUNCI

Angka Harapan Hidup; K-Means Clustering; Provinsi

KORESPONDENSI

Phone: 085860568311
E-mail: pratiwiyosinta22@gmail.com

A B S T R A K

Kementerian Kesehatan (Kemenkes) menyatakan usia harapan hidup orang Indonesia semakin meningkat hingga mencapai hampir rata-rata 73,5 tahun 2021. Alasan memilih penelitian ini karena beberapa dari masyarakat Indonesia kurang mengetahui persebaran jumlah angka harapan hidup di beberapa Provinsi di Indonesia selama kurun waktu 3 tahun. Penelitian ini menggunakan metode *K-Means clustering*. *K-Means clustering* merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteristik (*similarity*) antara satu data dengan data yang lain. Tujuan dilakukannya penelitian ini untuk mengetahui tingkat angka harapan hidup di setiap Provinsi di Indonesia menggunakan metode *K-Means clustering*. Teknik penelitian ini yaitu data akan dikelompokkan berdasarkan nama provinsi yang memiliki jumlah Angka Harapan Hidup saat lahir dari tahun 2019 sampai 2021. Setelah dilakukannya pengelompokan maka hasil akan didapatkan jumlah Angka Harapan Hidup saat lahir dengan variabel berupa id, kode provinsi, nama provinsi, jenis kelamin, angka harapan hidup, satuan, tahun, pengelompokan berdasarkan hasil cluster yang didapat, dengan parameter penelitian berupa angka harapan hidup. Diharapkan dengan adanya penelitian ini dapat memberikan masukan kepada pemerintah untuk memberikan solusi dari tingkat Angka Harapan Hidup saat lahir, agar ditahun berikutnya Angka Harapan Hidup bisa lebih baik. Hasil dari pengelompokkan menggunakan tools rapidminer di dapat, bahwa Provinsi persebaran angka harapan hidup tertinggi berada di cluster 0 sebanyak 104 item selama tiga tahun laki-laki dan perempuan, dan Provinsi terendah berada di cluster 1 sebanyak 99 item persebaran laki-laki dan perempuan periode 2019 sampai dengan 2021.

PENDAHULUAN

Kementrian Kesehatan (Kemenkes) menyatakan usia harapan hidup orang Indonesia semakin meningkat hingga mencapai hampir rata-rata 73,5 tahun 2021. Kemenkes menyatakan hal ini berkat semakin meningkatnya kesejahteraan masyarakat. Ukuran umum yang digunakan adalah angka harapan hidup saat lahir yang mencerminkan kondisi kesehatan saat itu. [1]. Angka Harapan Hidup (AHH) adalah rata-rata lamanya seseorang dapat bertahan hidup yang dihitung mulai kelahirannya. Angka harapan hidup berfungsi sebagai salah satu tolak ukur kinerja pemerintah dalam meningkatkan indeks pembangunan masyarakat serta perekonomiannya [2]. Angka Harapan Hidup pada dasarnya merupakan gambaran kondisi suatu wilayah secara garis besar. Semakin tingginya angka harapan menandakan adanya perbaikan status Kesehatan masyarakat [3].

Di Indonesia sendiri angka harapan hidup berkembang secara cepat, ini menunjukkan bahwa masyarakat Indonesia memiliki tingkat kesuburan yang sangat tinggi. Selama periode tahun 2010-2018, Indonesia telah berhasil meningkatkan umur harapan hidup saat lahir sebesar 1,39 tahun atau tumbuh sebesar 0,25 persen pertahun. Semakin tinggi angka harapan hidup akan menyumbang kenaikan Indeks Pembangunan Manusia (IPM).

Untuk itu sebagai salah satu cara memprediksi angka harapan hidup adalah dengan melakukan penerapan angka harapan hidup tangka Provinsi di Indonesia berdasarkan algoritma K-Means Clustering, maka dirumuskan masalah yaitu

bagaimana mengelompokkan angka harapan hidup di setiap Provinsi di Indonesia berdasarkan presentase angka harapan hidup menggunakan algoritma K-Means Clustering.

Berdasarkan Penelitian sebelumnya yang telah dilakukan oleh Riska dkk. (2020) yang berjudul “Penerapan metode algoritma K-Means Clustering dalam pengelompokkan angka harapan hidup saat lahir menurut Provinsi” yang melakukan pengelompokkan angka harapan hidup berdasarkan Provinsi di Indonesia dari tahun 2015-2018 dengan menggunakan algoritma K-Means Clustering”. [4]. Kemudian penelitian yang telah dilakukan oleh Riyan Muda dkk (2019) yang berjudul “Pengaruh Angka Harapan Hidup, Tingkat Pendidikan dan Pengeluaran perkapita Terhadap Pertumbuhan Ekonomi di Sulawesi Utara Pada Tahun 2003-2017”. Penelitian ini menjadi salah satu referensi untuk melaksanakan penelitian ini. Perbedaan dari penelitian sebelumnya dan penelitian sekarang yaitu data yang digunakan berbeda kemudian pada penelitian sekarang data di olah di rapidminer [5]

Penelitian ini bertujuan untuk mengetahui tingkat harapan hidup seseorang berdasarkan hasil cluster yang di dapat. Sehingga pada penelitian ini dapat memberikan solusi terhadap pemerintah dalam memperbaiki tingkat angka harapan hidup saat lahir, agar ditahun berikutnya angka harapan hidup bisa lebih baik lagi.

Oleh karena itu diusulkan penelitian dengan judul “Penerapan algoritma K-Means pada angka harapan hidup berdasarkan tingkat Provinsi”. Adapun yang menjadi alasan dilakukannya penelitian ini dengan judul tersebut karena kurangnya pengetahuan masyarakat akan angka harapan hidup di Indonesia, dan diharapkan agar tingkat kesejahteraan masyarakat pun semakin baik.

TINJAUAN PUSTAKA

Data Mining

Data Mining adalah pencarian atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar. *Data Mining* juga disebut sebagai serangkaian proses untuk mencari lebih dalam nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data. *Data mining* sering juga disebut *knowledge discovery in database* (KDD). KDD adalah kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar [6].

Data mining diartikan sebagai suatu proses ekstraksi informasi berguna dan berpotensi dari sekumpulan data yang terdapat secara implisit dalam suatu basis data. Data mining meliputi tugas-tugas sebagai ekstraksi pengetahuan, arkeolog data, eksplorasi dalam pemrosesan pola data dan memanen informasi. Semua aktifitas ini dilakukan secara otomatis dan memungkinkan adanya penemuan cepat bahkan oleh programmer [7]

Secara garis besar, data mining dikelompokkan menjadi dua yaitu:

1. Deskripsi Mining, suatu proses menemukan sebuah karakter yang penting di dalam basis data, Teknik ini termasuk *Descriptive Mining* seperti *Clustering Association*, dan *Sequential Mining*
2. *Predictive*, suatu proses menemukan sebuah pola dari data sehingga dapat digunakan untuk variabel lain pada masa depan, salah satu tekniknya terdapat di dalam *predictive mining* adalah klasifikasi, karena itulah fungsi satu ini dikatakan sebagai fungsi prediksi sama halnya dengan melakukan *predictive analisis* [8]

Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) merupakan proses Analisa terstruktur untuk memperoleh informasi yang benar, baru, bermanfaat dan menemukan pola dari data yang besar dan kompleks. *Data Mining* menjadi inti dari proses KDD, yakni dengan menggunakan algoritma tertentu untuk mengeksplorasi data, membangun model dan menemukan pola yang belum diketahui. Model digunakan untuk memahami fenomena data, Analisa maupun prediksi [9].

Clustering

Clustering merupakan suatu metode pengelompokkan berdasarkan ukuran kedekatan (kemiripan). Clustering adalah proses pengelompokkan objek berdasarkan informasi yang diperoleh dari data yang menjelaskan hubungan dari setiap

objek, yang memaksimalkan kemiripan dalam cluster yang berbeda. Tujuannya menemukan kluster yang berkualitas dalam waktu yang layak. Klastering dalam *data mining* berguna untuk proses Analisa data [10].

Clustering adalah proses pengelompokan benda serupa ke dalam kelompok yang berbeda, atau lebih tepatnya partisi dari sebuah dataset kedalam subset, sehingga data dalam setiap subset memiliki arti yang bermanfaat. Dimana dalam cluster terdiri dari kumpulan benda-benda yang mirip antara satu dengan yang lainnya dan berbeda dengan yang lainnya dan berbeda dengan benda yang terdapat pada cluster lainnya [11].

K-Means Clustering

K-Means Clustering adalah satu metode data clustering *non hierarki* yang mengelompokkan data dalam bentuk satu atau lebih cluster. Data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu cluster yang lain sehingga data yang berada dalam satu kluster kelompok memiliki tingkat variasi yang kecil [11].

Algoritma *K-Means Clustering* adalah salah satu algoritma dengan *partitional*, karena *K-Means Clustering* didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya. Algoritma *K-Means Clustering* didasarkan pada penentuan jumlah awal kelompok dengan mendefinisikan nilai *centroid* awalnya. Algoritma *K-Means Clustering* menggunakan proses secara berulang-ulang untuk mendapatkan basis *clustering* menggunakan proses secara berulang-ulang untuk mendapatkan basis data cluster [12].

Angka Harapan Hidup

Angka Harapan Hidup (AHH) saat lahir merupakan perkiraan rata-rata tambahan umur seseorang dari kandungan seorang ibu saat proses kelahiran yang diharapkan bisa dapat hidup secara normal dan sehat. Menurut statistic Indonesia , angka harapan hidup pada saat lahir (*life expentancy at birth*) ialah rata-rata tahun hidup yang akan dijalani oleh bayi yang baru lahir pada suatu tahun tertentu. Angka Harapan Hidup di suatu wilayah berbeda dengan wilayah lainnya tergantung dari kualitas hidup yang mampu dicapai oleh penduduk [4]

Menurut Anggraini dan Listiyaningsih (2013), angka harapan hidup (AHH) berkaitan erat dengan pembangunan social ekonomi suatu wilayah. Semakin tinggi angka harapan hidup (AHH) di suatu wilayah mengindikasikan pembangunan sosial ekonomi suatu wilayah tersebut semakin maju. Keberhasilan program Kesehatan dan program sosial ekonomi pada umumnya dapat dilihat dari peningkatan usia harapan hidup penduduk suatu wilayah. [13].

Rapidminer

Rapidminer merupakan perangkat lunak terbuka (*open source*). *Rapidminer* merupakan solusi untuk menganalisis penambangan data, penambangan teks, dan analitis prediktif. *Rapidminer* menggunakan berbagai teknik deskriptif dan prediktif untuk memberikan wawasan kepada pengguna dan membantu mereka membuat keputusan terbaik. *Rapidminer* memiliki sekitar 500 operator *data mining*, termasuk operator *input*, *output*, *preprocessing* data, dan visualisasi. *Rapdminner* adalah perangkat lunak mandiri untuk analisis data dan mesin penambangan data yang diintegrasikan kedalam produk anda sendiri. *Rapidminer* ditulis dalam java, sehingga bekerja pada semua sistem operasi [14]

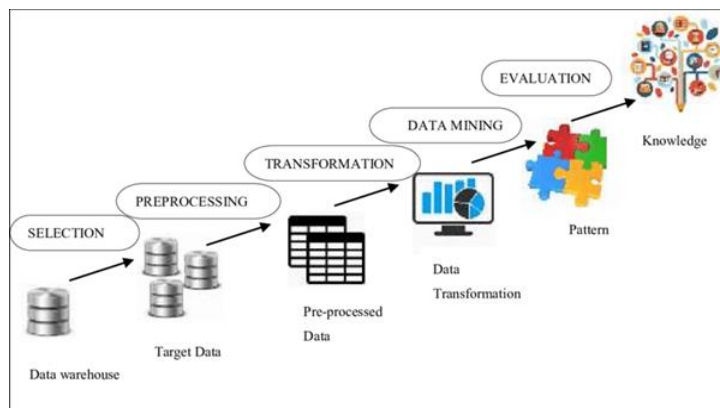
Indeks Pembangunan Manusia

Mengutip isi *Human Development Report* (HDR) pertama tahun 1990. Indeks pembangunan manusia adalah suatu proses untuk memperbanyak pilihan-pilihan yang dimiliki manusia. Diantara banyak pilihan tersebut, pilihan yang terpenting adalah untuk berumur Panjang dan sehat, untuk berilmu pengetahuan, dan untuk mempunyai akses terhadap sumber daya yang dibutuhkan agar dapat hidup secara layak. Indeks pembangunan manusia (IPM) mengukur capaian pembangunan manusia berbasis sejumlah komponen dasar kualitas hidup, IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut mencakup umur Panjang dan sehat, pengetahuan, dan kehidupan yang layak [10].

Kesejahteraan adalah kondisi agregat dari kepuasan individu. Kesejahteraan biasanya diukur dengan Indeks Pembangunan Manusia (IPM) untuk melihat kondisi sumber daya manusia pada suatu negara. Pembangunan manusia berarti pertumbuhan yang positif dan perubahan dalam tingkat kesejahteraan, yang meliputi aspek sosial, ekonomi,

budaya, politik dan lingkungan. Oleh sebab itu fokus utama pembangunan manusia adalah pada manusia dan kesejahteraannya [15].

METODOLOGI



Gambar 1. Proses *Knowledge Discovery in Database* (KDD)

Pada proses tahapan perancangan data ini berdasarkan Knowledge Discovery in Database (KDD) adalah proses analisa terstruktur untuk memperoleh informasi yang benar, baru, bermanfaat dan menemukan pola dari data yang besar dan kompleks. Berikut ini ada beberapa tahapan Knowledge Discovery in Database antara lain:

Data Selection

Data yang digunakan pada penelitian ini menggunakan data asli angka harapan hidup pada tahun 2019-2021. Atribut yang ada pada data ini berjumlah 4 atribut yang terdiri dari, Nama Provinsi, Jenis Kelamin, Presentase Angka Harapan Hidup, dan Tahun, dan data awal angka harapan hidup sebanyak 203. Operator yang digunakan pada tahap *data selection*.

Tabel 1. Data Angka Harapan Hidup

id	Kode Provinsi	Nama Provinsi	Jenis Kelamin	Angka Harapan Hidup	Satuan	Tahun
1	1100	ACEH	LAKI-LAKI	67,98	TAHUN	2019
2	1100	ACEH	PEREMPUAN	71,85	TAHUN	2019
3	1200	SUMATRA UTARA	LAKI-LAKI	67,07	TAHUN	2019
4	1200	SUMATRA UTARA	PEREMPUAN	70,92	TAHUN	2019
5	1300	SUMATRA BARAT	LAKI-LAKI	67,42	TAHUN	2019

Pada table angka harapan hidup terdiri dari 7 variabel yaitu id, Kode Provinsi, Nama Provinsi, Jenis Kelamin, Angka Harapan Hidup, Satuan, dan Tahun.

Data Preprocessing

Dataset yang di dapat dari open data Jawa Barat yang berasal dari Badan Pusat Statistik Indonesia berupa data Angka Harapan Hidup dari tahun 2019-2021. Operator yang digunakan pada tahap *preprocessing data* adalah operator *replace missing value* dikarenakan pada saat import data menggunakan operator *read excel* terdapat *missing value* yang di lihat dari tampilan statistic data.

Name	Type	Missing	Statistics	Filter (5 / 5 attributes):	Search for Attributes
Id ▼ Nama_Provinsi	Polynomial	0	Least PAPUA (5)	Most ACEH (6)	Values ACEH (6), BALI (6), ... [32 more]
Cluster ▼ cluster	Nominal	0	Least cluster_1 (99)	Most cluster_0 (104)	Values cluster_0 (104), cluster_1 (99)
▼ Jenis_Kelamin	Integer	0	Min 1	Max 2	Average 1.498
▼ Presentase_Angka_Harapan_...	Real	0	Min 62.960	Max 76.890	Average 70.079
▼ Tahun	Integer	0	Min 2019	Max 2021	Average 2019.995

Gambar 2. Tampilan Statistic Terdapat Missing Value

Data Transformation

Hasil dari transformasi data yang diolah di rapidminer.

Tabel 2. Data Angka Harapan Hidup setelah Transformasi

Nama Provinsi	Jenis kelamin	Presentase Angka Harapan Hidup	Tahun
ACEH	1	67,98	2019
ACEH	2	71,85	2019
SUMATERA UTARA	1	67,07	2019
SUMATERA UTARA	2	70,92	2019
SUMATERA BARAT	1	67,42	2019
SUMATERA BARAT	2	71,29	2019
RIAU	1	69,62	2019
RIAU	2	73,43	2019
JAMBI	1	69,11	2019
JAMBI	2	72,97	2019
SUMATERA SELATAN	1	67,78	2019
SUMATERA SELATAN	2	71,63	2019
BENGKULU	1	67,28	2019
BENGKULU	2	71,13	2019

Data Mining

Pada pemodelan algoritma K-Means berguna untuk melakukan clustering data, dimana terlebih dahulu melakukan pemilihan jumlah cluster yang telah ditentukan. Jumlah cluster yang dipakai yaitu berjumlah 2 cluster. Selanjutnya dari model yang telah diterapkan, maka menghasilkan hasil clustering. Hasil proses ini merupakan data yang dikelompokkan berdasarkan kemiripan karakteristik dari setiap data, sehingga dapat hasil pola atau informasi yang tersembunyi dari data-data tersebut.

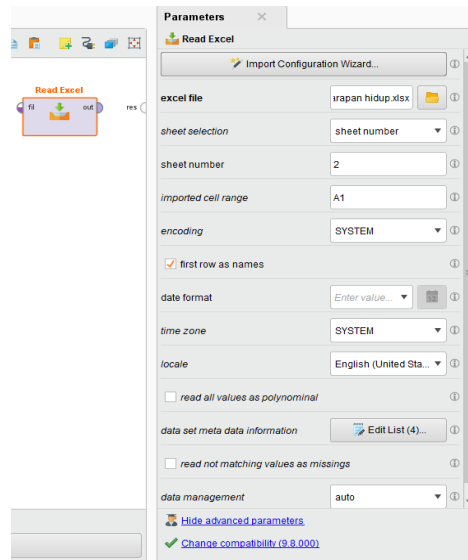
Evaluasi

Interpretation/Evaluation

Pada tahap Evaluation yaitu tahap pengujian yang menggunakan dataset angka harapan hidup dengan menerapkan pengolahan di aplikasi rapidminer. Berikut adalah proses klasterisasi pada rapidminer:

Read Excel

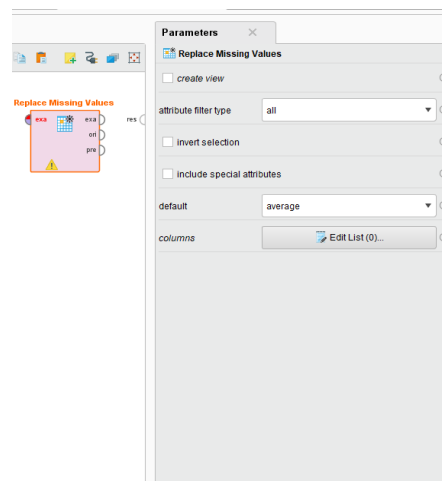
Operator *read excel* digunakan untuk memuat data langsung dari *Microsoft excel*, dan mempermudah untuk membaca file dari *Microsoft excel*.



Gambar 3. Setting Operator Read Excel

Replace Missing Value

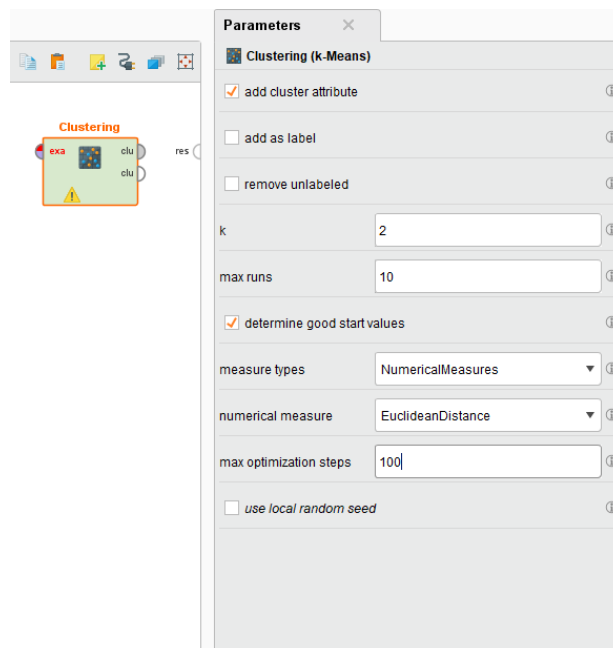
Operator *Replace Missing Value* digunakan karena dataset ini terjadi missing value, sehingga perlu ditambahkan operator ini.



Gambar 4. Tampilan Operatos Setting *Replace Missing Value*

Clustering

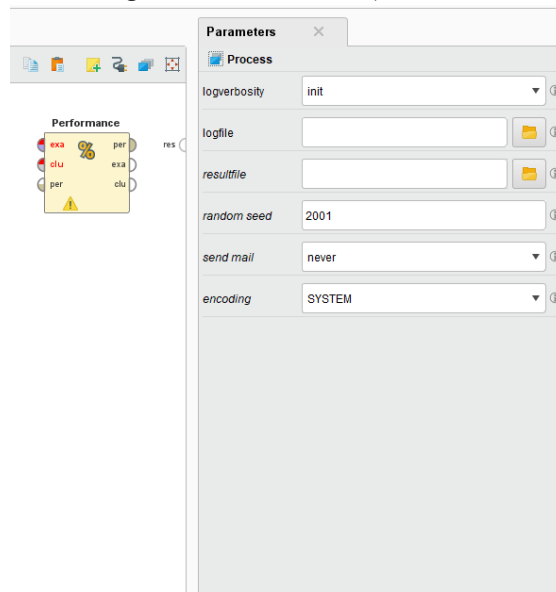
Operator clustering digunakan untuk mengelompokkan beberapa cluster berdasarkan hasil DBI (*Devies Boulden Indeks*) yang terkecil.



Gambar 5. Operator Clustering dan Parameter

Performance

Operator *Performance* digunakan untuk mengetahui hasil dari DBI (*Devies Boulden Indeks*).



Gambar 6. Operator Performance dan parameter

Pada tahap evaluasi KDD penggunaan operator performance yaitu untuk menentukan nilai K optimal dengan parameter Max run/iterasi sebanyak 10 kali percobaan.

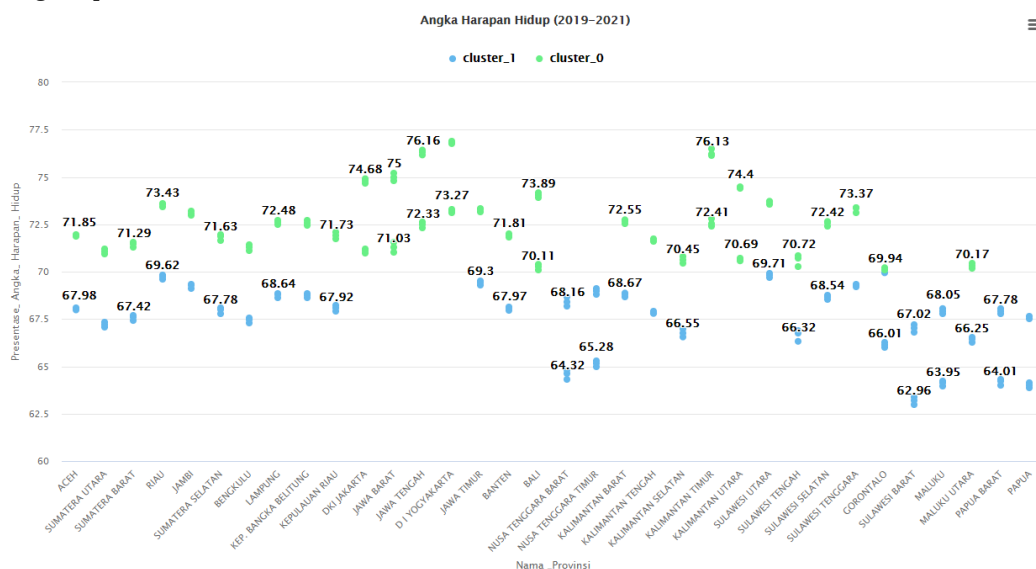
HASIL DAN PEMBAHASAN

Hasil Pengolahan RapidMiner

Tabel 3. Nilai (K) dan DBI

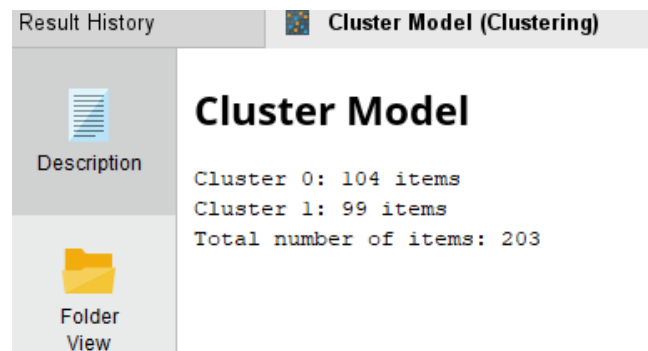
CLUSTER	DBI
2	0.222
3	0.250
4	0.266
5	0.279
6	0.332
7	0.332
8	0.318
9	0.318
10	0.317

Dari hasil pengolahan diketahui, untuk optimasi jumlah cluster adalah 2 dengan nilai DBI0,222. Maka untuk data ini, pengelompokan yang dilakukan dengan optimasi jumlah cluster adalah 2. Maka untuk data ini, pengelompokan yang dilakukan dengan optimasi cluster adalah 2.



Gambar 7. Tampilan hasil chart scatter pengelompokan angka harapan hidup

Dari Tampilan hasil scatter di atas dapat dilihat bahwa nama Provinsi berada dibawah dan presentase angka harapan hidup ada di sisi kiri, dari hasil pengelompokan bahwa cluster terbagi menjadi dua yaitu cluster 0 dan cluster 1 yaitu cluster tertinggi dan terendah, cluster tertinggi berada di cluster 0 dan cluster terendah, berada di cluster 1, cluster 0 berwarna hijau dan cluster 1 berwarna biru.



Gambar 8. Tampilan hasil cluster model

Text View merupakan sheet untuk menampilkan database yang telah diolah secara keseluruhan lengkap dengan clusternya. Tampilan dari cluster model (*clustering*). Pada cluster model terbagi menjadi dua cluster, cluster 0 dan cluster 1, dimana cluster 0 terdapat 104 items, cluster 1 99 item, jadi jumlah keduanya sebanyak 203 item.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 1.313
Avg. within centroid distance_cluster_0: 1.306
Avg. within centroid distance_cluster_1: 1.320
Davies Bouldin: 0.225
```

Gambar 9. Tampilan hasil performance vector

Hasil uji performance dengan evaluasi Davies-Bouldin Index (DBI dalam kinerja pengelompokkan cluster. Dimana seberapa baik kinerja dari centroid yang dihasilkan juga bisa diukur dengan penggunaan operator Cluster Distance Performance. Dari hasil implementasi menggunakan RapidMiner dalam menguji data jenis penilaian kinerja, nilai Davies Bouldin Index (DBI). Berdasarkan hasil Davies-Bouldin Index yang diperoleh yaitu cluster_2 dengan hasil DBI 0,222, cluster 3 hasil DBI 0,250, cluster 2 dengan hasil DBI 0,222 dan cluster 3 hasil DBI 0,250. Maka dapat disimpulkan hasil evaluasi dari hasil clustering bahwa hasil yang terbaik berada pada cluster 2 karena nilai DBI yang diperoleh sudah mendekati nilai 0.

Row No.	Nama Provi...	cluster ↑	Jenis kelamin	Presentase ...	Tahun
2	ACEH	cluster_0	2	71.850	2019
4	SUMATERA ...	cluster_0	2	70.920	2019
6	SUMATERA B...	cluster_0	2	71.290	2019
8	RIAU	cluster_0	2	73.430	2019
10	JAMBI	cluster_0	2	72.970	2019
12	SUMATERA S...	cluster_0	2	71.630	2019
14	BENGKULU	cluster_0	2	71.130	2019
16	LAMPUNG	cluster_0	2	72.480	2019
18	KEP. BANGK...	cluster_0	2	72.460	2019
20	KEPULAUAN ...	cluster_0	2	71.730	2019
21	DKI JAKARTA	cluster_0	1	70.960	2019
22	DKI JAKARTA	cluster_0	2	74.680	2019
23	JAWA BARAT	cluster_0	1	71.030	2019
24	JAWA BARAT	cluster_0	2	74.810	2019
25	JAWA TENGAH	cluster_0	1	72.330	2019
26	JAWA TENGAH	cluster_0	2	76.160	2019
27	D I YOGYAKA...	cluster_0	1	73.130	2019
28	D I YOGYAKA...	cluster_0	2	76.760	2019
30	JAWA TIMUR	cluster_0	2	73.150	2019

ExampleSet (205 examples, 2 special attributes, 3 regular attributes)

Gambar 10. Tampilan dataset setelah di olah di rapidminner

Tampilan dataset setelah diolah, dimana cluster terbagi menjadi dua, dan jenis kelamin setelah di inisialisasi dan Provinsi dijadikan id di rapidminner. Folder view: merupakan sheet untuk menampilkan database yang telah diolah secara keseluruhan lengkap dengan clusternya.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan Penelitian yang di dapat pada proyek penelitian tugas akhir ini sebagai berikut: Hasil dari penelitian ini didapatkan dua cluster yaitu cluster 0 dan cluster 1. Presentase angka harapan hidup tertinggi berada di cluster 0, dengan jumlah 104 item dari 34 Provinsi selama 3 tahun dari 2019 sampai dengan 2021 laki-laki dan perempuan dan persentase angka harapan hidup terendah berada di cluster 1 dengan jumlah 99 item dari 34 Provinsi selama 3 tahun dari 2019 sampai dengan 2021 laki-laki dan perempuan. Berdasarkan penelitian yang telah dilakukan didapatkan Nilai DBI yang terkecil (teroptimum) berada di kluster 2 yaitu dengan nilai 0,222.

Saran

Adapun saran yang dikemukakan dalam penelitian ini adalah sebagai berikut: Setelah mengetahui tingkat angka harapan hidup, diharapkan untuk mengetahui tingkat harapan hidup seseorang berdasarkan hasil cluster yang di dapat. Sehingga pada proyek tugas akhir ini dapat memberikan solusi terhadap pemerintah dalam memperbaiki tingkat angka harapan hidup saat lahir, agar ditahun berikutnya tingkat angka harapan hidup bisa lebih baik lagi. Untuk peneliti selanjutnya diharapkan dapat menggunakan analisis statistik yang lain dalam pengolahan data.

DAFTAR PUSTAKA

- [1] faturahman, "ANGKA HARAPAN HIDUP SEBAGAI INTERVENING ANGGARAN KESEHATAN MASYARAKAT DAN SUMBER DAYA MANUSIA KESEHATAN TERHADAP KEMISKINAN DI KOTA JAMBI," 2019.
- [2] diana wulan septyananda, "IMPLEMENTASI PRINCIPAL COMPONENT ANALYSIS DALAM MENANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI ANGKA HARAPAN HIDUP DI INDONESIA," 2022.
- [3] L. Ginting Andi, "Dampak Angka Harapan Hidup dan Kesempatan Kerja Terhadap Kemiskinan," *EcceS (Economics, Social, and Development Studies)*, vol. 7, no. 1, p. 42, Jun. 2020, doi: 10.24252/ecc.v7i1.13197.
- [4] T. J. H. I. Oktavia Riska, "Penerapan Metode Algoritma K-means Dalam Pengelompokan Angka Harapan Hidup Saat Lahir Menurut Provinsi," 2020. [Online]. Available: www.bps.go.id/
- [5] K. R. B. J. K. Muda Riyan, "PENGARUH ANGKA HARAPAN HIDUP, TINGKAT PENDIDIKAN DAN PENGELUARAN PERKAPITA TERHADAP PERTUMBUHAN EKONOMI DI SULAWESI UTARA PADA TAHUN 2003-2017," 2019.
- [6] A. Fauzi *et al.*, "METHOMIKA: Jurnal Manajemen Informatika & Komputerisasi Akuntansi Halaman 152," *sistem pendukung pemilihan pekerjaan menggunakan metode apriori berdasarkan korelasi jurusan dengan ipk untuk mengetahui pekerjaan yang tepat*, vol. 2, no. 2, 2018, doi: 10.46880/jmika.Vol2No2.pp152-159.
- [7] S. Takalapeta, "Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5," *JIMP-Jurnal Informatika Merdeka Pasuruan*, vol. 3, 2018.
- [8] M. Benri, H. Metisen, and S. Latipa, "ANALISIS CLUSTERING MENGGUNAKAN METODE K-MEANS DALAM PENGELOMPOKAN PENJUALAN PRODUK PADA SWALAYAN FADHILA," 2015.
- [9] A. H. Zanuardi, "Knowledge Discovery in Database Analysis of Traffic Accident Characteristic on Ahmad Yani Road Surabaya through Knowledge Discovery in Database Approach," 2018.
- [10] A. Charolina, D. Ruswanti, P. Studi Informatika, F. Sains, T. dan Kesehatan, and U. Sahid Surakarta Jl Adi Sucipto, "KLAUSTERING KOTA DAN KABUPATEN DI INDONESIA BERDASARKAN UMUR HARAPAN HIDUP SAAT LAHIR DENGAN K-MEDOIDS," 2019.
- [11] R. A. Indraputra and R. Fitriana, "K-Means Clustering Data COVID-19," 2020.
- [12] N. Wayan Wardani, "Analisis Pesebaran Penularan Virus Corona Di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering," 2021, [Online]. Available: <http://jurnal.mdp.ac.id>
- [13] R. Herawaty Bangun, "ANALISIS DETERMINAN ANGKA HARAPAN HIDUP KABUPATEN MANDAILING NATAL (Life Expectations Determinants Analysis In Mandailing Natal Regency)," vol. 4, no. 3, 2019.
- [14] Imami Zulia A, "ANALISIS SENTIMEN REVIEW KOSMETIK PADA WEBSITE FEMALE DAILY MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT VECTOR MACHINE BERBASIS PARTICLE SWARM OPTIMIZATION," 2019.

- [15] I. K. G. B. I Made Cahyadi Putra Kusuma, “ANALISIS PENGARUH ANGKA HARAPAN HIDUP, RATA-RATA LAMA SEKOLAH, TINGKAT KEMISKINAN TERHADAP KESEJAHTERAAN PADA KABUPATEN/KOTA DI PROVINSI BALI,” pp. 4059–4081, 2023.