

Penerapan Algoritma K-Means Clustering Pada Tingkat Inflasi Kota Di Indonesia

Novia Wulandari

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Nisa Dienwati Nuris

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Saeful Anwar

Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) IKMI Cirebon

Alamat: Cirebon, Jawa Barat, Indonesia 45131

Korespondensi penulis: noviawulan2001@gmail.com

Abstract. *Inflation is a general tendency to increase the prices of goods and services, and it happens all the time. when the prices of domestic goods and services rise, inflation will rise. Depreciation causes the prices of goods and services to rise. Uncontrolled inflation can result in losses for society and the government. Therefore, an appropriate study is needed to map the dynamics of inflation in a region. One way to map the inflation rate is clustering. Clustering is dividing data into groups with the same characteristics. The author took the initiative to analyze the urban inflation rate in Indonesia from 2020 to 2022. The data is sourced from the Central Statistics Agency (BPS) website. This analysis uses the K-Means Clustering method with 5 clusters. the group with the highest inflation is in cluster 0, the high inflation group is in cluster 1, the moderate inflation group is in cluster 2, the low inflation group is in cluster 3, and the lowest inflation is in cluster 4. by categorizing the inflation rate of cities in Indonesia, it can be seen which cities in Indonesia have very high, high, medium, low and very low inflation rates.*

Keywords: *Inflation, Data Mining, K-Means, Clustering*

Abstrak. Inflasi adalah kecenderungan naiknya harga barang dan jasa secara umum, dan itu terjadi sepanjang waktu. ketika harga barang dan jasa dalam negeri naik, inflasi akan naik. Depresiasi menyebabkan naiknya harga barang dan jasa. Inflasi yang tidak terkendali dapat mengakibatkan kerugian bagi masyarakat dan pemerintah. Oleh karena itu, diperlukan studi yang tepat untuk memetakan dinamika inflasi suatu wilayah. Salah satu cara untuk memetakan tingkat inflasi adalah clustering. Clustering adalah pembagian data menjadi kelompok-kelompok dengan karakteristik yang sama. Penulis berinisiatif menganalisis tingkat inflasi kota di Indonesia dari tahun 2020 hingga tahun 2022. Data tersebut bersumber dari website Badan Pusat Statistik (BPS). Analisis ini menggunakan metode K-Means Clustering dengan 5 jumlah cluster. kelompok dengan inflasi paling tinggi berada di cluster 0, kelompok inflasi tinggi ada di cluster 1, kelompok inflasi sedang berada di cluster 2, kelompok inflasi rendah berada di cluster 3, dan inflasi paling rendah berada di cluster 4. dengan mengkategorikan tingkat inflasi kota-kota di Indonesia,

dapat diketahui kota-kota di Indonesia dengan tingkat inflasi sangat tinggi, tinggi, sedang, rendah dan sangat rendah.

Kata kunci: Inflation, Data Mining, K-Means, Clustering

LATAR BELAKANG

Inflasi merupakan proses kenaikan harga yang terjadi terus menerus karena mekanisme pasar secara umum, yang dapat diakibatkan oleh beberapa faktor, antara lain peningkatan konsumsi masyarakat, likuiditas pasar yang berlebihan yang mendorong konsumsi, dan spekulasi, serta distribusi komoditas yang tidak merata. Inflasi merupakan indikator untuk melihat sejauh mana perubahan ketika proses kenaikan harga terus menerus dan saling mempengaruhi..

Salah satu teknik clustering yang sederhana dan populer adalah k-means clustering. Algoritma pengelompokan K-means telah diterapkan di beberapa bidang, beberapa studi yang dilakukan dengan algoritma ini adalah misalnya penelitian yang dilakukan oleh (Pandiangan & Lintang, 2019) “K-Means Clustering Dalam Mengelompokkan Nilai Akhirskripsi Mahasiswa”. Berdasarkan hasil riset klasifikasi Algoritma K-Means Clustering terhadap Nilai Akhir Tugas Akhir mahasiswa dengan pengujian pada berbagai variasi dataset dan cluster yang terdiri dari 100, 200, 300, 400 dan 500 data serta variasi jumlah cluster 2, 3, 4, dan 5. Penerapan klasifikasi menggunakan Algoritma K-Means Clustering berbeda pada setiap dataset, klasifikasi data tidak dipengaruhi oleh jumlah dataset dan jumlah kelompok atau cluster. Oleh karena itu, dapat disimpulkan bahwa algoritma K-Means dapat digunakan untuk mengklasifikasikan nilai akhir tugas akhir mahasiswa.

Stabilitas inflasi menjadi syarat penting untuk pertumbuhan ekonomi yang akhirnya membawa keuntungan bagi peningkatan kesejahteraan masyarakat.

Setelah melihat banyaknya informasi mengenai inflasi kota-kota di Indonesia, penulis memutuskan untuk menggunakan algoritma *K-Means Clustering* dalam melakukan klasifikasi inflasi perkotaan di Indonesia. Terdapat lima *cluster* yang dibentuk dengan kelompok nilai yang sangat tinggi berada di *cluster* 0, kelompok nilai tinggi berada di *cluster* 1, kelompok nilai sedang berada di *cluster* 2, kelompok nilai rendah berada di *cluster* 3, dan kelompok nilai sangat rendah berada di *cluster* 4. Dengan

melakukan pengelompokan tingkat inflasi di kota-kota Indonesia, akan diketahui kota-kota mana yang memiliki inflasi sangat tinggi, tinggi, sedang, rendah, dan sangat rendah.

Maka dari itu penulis mengangkat sebuah judul penelitian “Penerapan Algoritma K-Means Clustering Pada Tingkat Inflasi Kota di Indonesia”. Yang mana data yang digunakan dalam data mining untuk mengelompokkan data menjadi variabel atau atribut data adalah data dari tahun 2020-2022.

KAJIAN TEORITIS

A. Data Mining

Data mining adalah serangkaian metode untuk menemukan informasi baru yang sebelumnya tidak diketahui dalam kumpulan data yang besar. Pola yang ditemukan dapat dimanfaatkan untuk mendukung proses pengambilan keputusan (Muttaqin & Defriani, 2020). Penambangan data adalah sebuah proses untuk menemukan dan mengakses informasi penting Menggunakan teknik/metode tertentu berdasarkan pengetahuan atau keterampilan. Hasil penelitian dapat digunakan dalam data mining sebagai alternatif proses pengambilan keputusan (Handayani, 2022).

B. Clustering

Pengelompokan data (*Clustering*) pada dasarnya adalah pengklasifikasian sekumpulan data yang atribut kelasnya tidak ditentukan, berdasarkan konsep tersebut, prinsip pengelompokan adalah memaksimalkan dan meminimalkan kesamaan dalam kelas (Hardiani, 2022). Teknik clustering membagi data ke dalam kelompok-kelompok sehingga data dengan karakteristik yang mirip ditempatkan pada kelompok yang serupa (Priyatman et al., 2019).

C. K-Means

Algoritma k-means adalah algoritma yang mengelompokkan data berdasarkan centroid terdekat. Algoritma K-means dapat digambarkan sebagai clustering non hirarki yang mengelompokkan data ke dalam satu atau lebih kelompok dengan karakteristik serupa (Darmansah, 2021). Metode k-means dapat digunakan untuk membantu melakukan pengelompokan database (Sudibyo, 2020).

D. Rapidminer

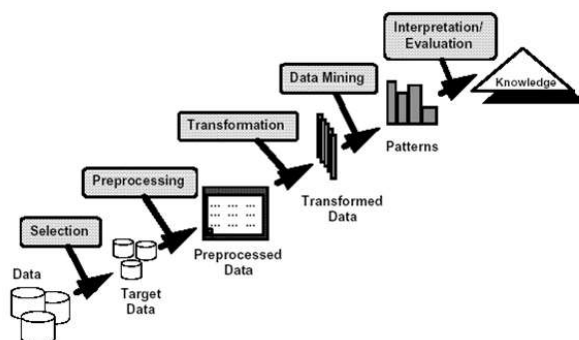
RapidMiner merupakan perangkat lunak sumber terbuka (*open source*) yang menawarkan solusi untuk melakukan analisis data, analisis teks, dan analisis prediktif. *RapidMiner* merupakan perangkat lunak independen untuk analisis data dan mesin penambangan data yang dapat diintegrasikan ke dalam produknya sendiri. Dikembangkan menggunakan bahasa pemrograman Java, *RapidMiner* dapat dijalankan pada semua sistem operasi. *RapidMiner* mengaplikasikan beragam teknik deskriptif dan prediktif untuk memberikan wawasan kepada pengguna agar dapat mengambil keputusan yang terbaik (Nabila et al., 2021)

E. Inflasi

Inflasi adalah peristiwa moneter yang menurunkan harga mata uang terhadap suatu barang (Mulyani, 2020). Inflasi merupakan masalah yang selalu dan pasti dialami oleh setiap negara (Fuady & Nugraha, 2017). Inflasi adalah salah satu masalah ekonomi klasik yang dapat mempengaruhi pendapatan riil, yang memiliki efek negatif permanen pada ekonomi makro. Dalam hal ini, inflasi menjadi indikator penting untuk menjaga stabilitas ekonomi (Santosa, 2017).

METODE PENELITIAN

Pendekatan desain dalam penelitian ini berbasis *Knowledge Discovery in Databases (KDD)*. KDD berkaitan dengan seni sintesis dan penemuan ilmiah, interpretasi dan visualisasi dalam berbagai jenis data. Penemuan pengetahuan di gudang data (KDD) adalah cara penting untuk menyelidiki pola dalam data dan mendapatkan hasil yang akurat, inovatif, dapat ditindaklanjuti, dan berwawasan.



Gambar 1 Tahapan KDD

1. *Data Selection*

Data Selection adalah proses pemilihan (sampling) data dari suatu dataset yang harus dilakukan sebelum *data mining* di KDD. Data pemrosesan pesan disimpan dalam file terpisah dari database aplikasi.

2. *Data Preprocessing*

Data preprocessing adalah teknik penambangan data. Sebelum tahap konstruksi. Pertama, data mentah diproses. *Preprocessing* data atau data *preprocessing* dilakukan dengan menghilangkan pesan *error*.

3. *Transformation*

Transformasi data digunakan dalam pemrosesan data untuk mengubah data menjadi format yang sesuai. Banyak teknik transformasi data termasuk normalisasi, pemilihan atribut, dan diskritisasi.

4. *Data Mining*

Sebuah proses di mana berbagai teknik digunakan untuk mengekstraksi pola potensial dari data yang berguna.

5. *Interpretation/Evaluation*

Langkah ini melibatkan pemeriksaan apakah model atau data yang dihasilkan konsisten dengan informasi atau asumsi sebelumnya.

Teknik Pengumpulan Data

Teknik Pengumpulan data yaitu dengan melakukan teknik *online research*, adapun teknik *online research* adalah melakukan penelusuran data melalui internet dengan mengakses situs web Badan Pusat Statistik (BPS). Dataset yang diambil yaitu data Inflasi kota di Indonesia pada tahun 2020 sampai 2022. Jumlah record data dari dataset ini sebanyak 270 data.

Dalam penelitian ini menggunakan data sekunder, dimana data tersebut diperoleh dari *Website* resmi Badan Pusat Statistik (BPS). Data tersebut diambil pada tanggal 04 Januari 2023, data yang diambil adalah data dari tahun 2020 sampai 2022, yang dapat diakses melalui:

<https://www.bps.go.id/indicator/3/1708/1/inflasi-90-kota-umum-.html>

HASIL DAN PEMBAHASAN

A. Mengelompokkan Kota di Indonesia Berdasarkan Tingkat Inflasi Menggunakan Algoritma K-Means

1. Data

Data yang digunakan dalam penelitian ini adalah data inflasi 90 Kota di Indonesia tahun 2020, 2021 dan 2022.

Tabel 1 Data Inflasi Tahun 2020

No	Kota	Januari	Februari	Maret	Desember
1.	Kota Meulaboh	1.44	-0.10	0.52	1.02
2.	Kota Banda Aceh	0.77	0.54	0.61	0.85
3.	Kota Lhokseumawe	0.08	0.49	0.64	1.27
4.	Kota Sibolga	0.20	0.69	-0.79	1.32
5.	Kota Pematang Siantar	0.62	0.12	-0.12	1.47
6.	Kota Medan	0.58	0.14	-0.19	0.65
7.	Kota Padangsidempuan	0.32	-0.01	0.53	1.00

90	Kota Jayapura	0.17	0.40	-0.29	1.02

Tabel 2 Data Inflasi Tahun 2021

No	Kota	Januari	Februari	Maret	Desember
1.	Kota Meulaboh	0.83	-0.95	-0.57	0.81
2.	Kota Banda Aceh	0.77	-0.56	-0.45	0.74
3.	Kota Lhokseumawe	0.80	-0.70	-0.09	0.59
4.	Kota Sibolga	0.74	-0.68	-0.29	0.17
5.	Kota Pematang Siantar	1.13	-0.46	-0.23	0.85
6.	Kota Medan	0.38	-0.33	-0.03	0.44
7.	Kota Padangsidempuan	0.38	-0.28	-0.57	0.35

90	Kota Jayapura	0.10	0.08	0.13	0.57

Tabel 3 Data Inflasi Tahun 2022

No	Kota	Januari	Februari	Maret	Desember
1.	Kota Meulaboh	0.95	-0.87	0.69	0.98
2.	Kota Banda Aceh	1.01	-0.34	0.59	0.64
3.	Kota Lhokseumawe	1.12	-0.99	0.85	1.51
4.	Kota Sibolga	1.53	-0.21	0.93	1.44
5.	Kota Pematang Siantar	0.96	0.18	0.77	1.61
6.	Kota Medan	1.04	-0.28	0.68	1.54
7.	Kota Padangsidimpuan	0.90	0.37	1.11	0.64

90	Kota Jayapura	0.56	-0.02	0.66	0.66

2. *Data Selection*

Pada tahap selection penulis menyeleksi atau memilih data yang diperlukan data tersebut kemudian disesuaikan dengan kebutuhan. Data dalam penelitian ini terdiri dari 12 atribut dengan jumlah data 270 *record*.

3. *Pre-Processing/Cleaning*

Tujuan dilakukannya *Pre-Processing* adalah untuk pembersihan data dari noise dan memastikan tidak ada duplikasi data. Pada tahap ini dilakukan *case folding* yang digunakan untuk mengubah huruf besar menjadi huruf kecil.

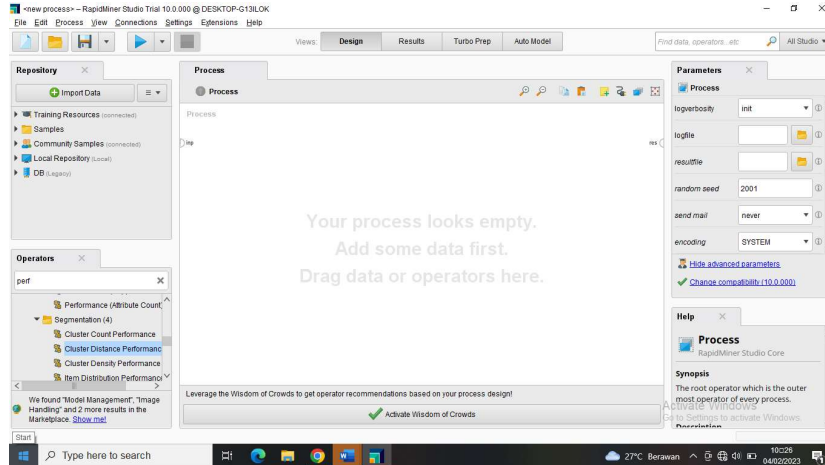
4. *Transformation*

Transformasi data merupakan proses untuk mengubah jenis atribut data menjadi format yang tepat untuk penambangan data. Karena dalam penelitian ini akan diuji menggunakan alat perangkat lunak Rapidminer, maka data yang telah diproses sebelumnya akan diubah agar sesuai dengan algoritma K-Means yang digunakan, yang terdiri dari 12 atribut dan 270 catatan.

5. *Proses Clustering*

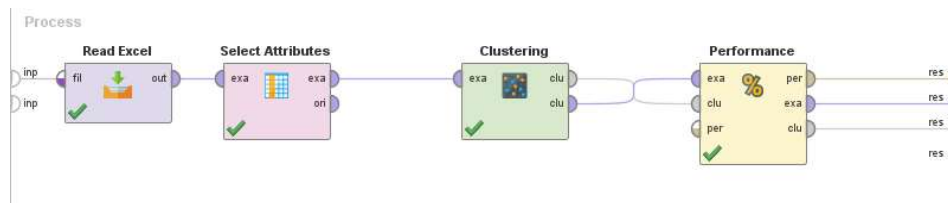
Proses *clustering* dengan menggunakan algoritma K-Means akan dilakukan terhadap 270 data Inflasi di Indonesia. Hasil *clustering* akan diperoleh kelompok data inflasi yang dapat digunakan untuk mengetahui daerah mana yang mengalami tingkat inflasi rendah, sedang dan tinggi.

Algoritma yang digunakan dalam proses clustering pada tingkat inflasi di Indonesia adalah algoritma K-Means dengan menggunakan Rapidminer.



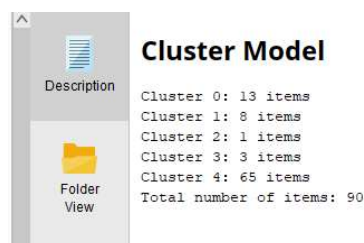
Gambar 2 Halaman Muka Rapidminer

Atribut yang digunakan untuk menghitung tingkat inflasi di Indonesia adalah atribut *Read Excel* yang berfungsi untuk membaca data utama, atribut *Select Attributes* yang berfungsi untuk memilih subset dari *ExampleSet* dan menghapus atribut lainnya, atribut *Clustering Operator* yang mengurutkan data menjadi yang telah ditentukan, dan fungsi *performance* yang bekerja untuk mengevaluasi kinerja model memberikan daftar tujuan kinerja sesuai dengan tugas yang diberikan. Contoh model *clustering K-Means* dapat dilihat pada Gambar 3 berikut.

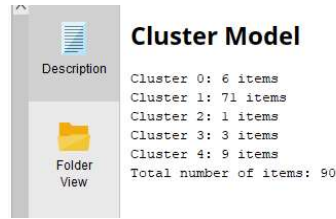


Gambar 3 Model K-Means Clustering

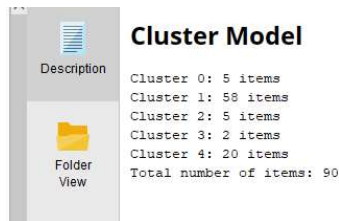
Setelah proses dijalankan, dapat dilihat hasil pada *cluster* model (*Clustering*), Untuk pilihan *Description* dapat dilihat pada Gambar 4 untuk tahun 2020, gambar 5 untuk tahun 2021 dan gambar 6 untuk tahun 2022 berikut.



Gambar 4 Hasil pilihan description tahun 2020

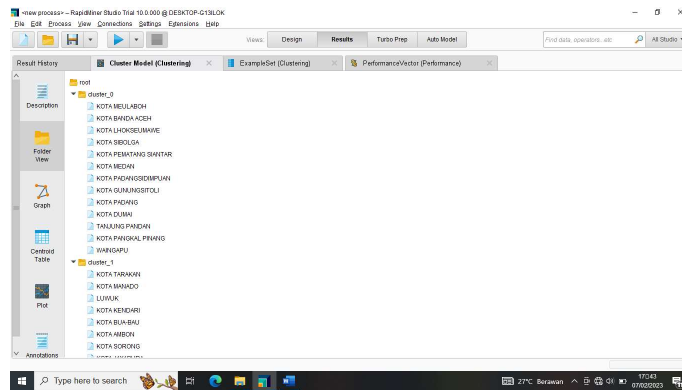


Gambar 5 Hasil pilihan description tahun 2021

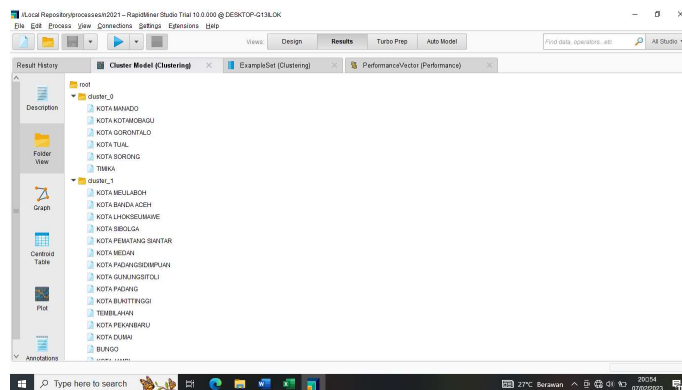


Gambar 6 Hasil pilihan description tahun 2022

Terdapat lima *cluster* dari setiap tahunnya, pada gambar Gambar 7 tampilan pilihan *folder view* pada tahun 2020, pada Gambar 8 tampilan pilihan *folder view* pada tahun 2021 dan pada Gambar 9 tampilan pilihan *folder view* pada tahun 2022.

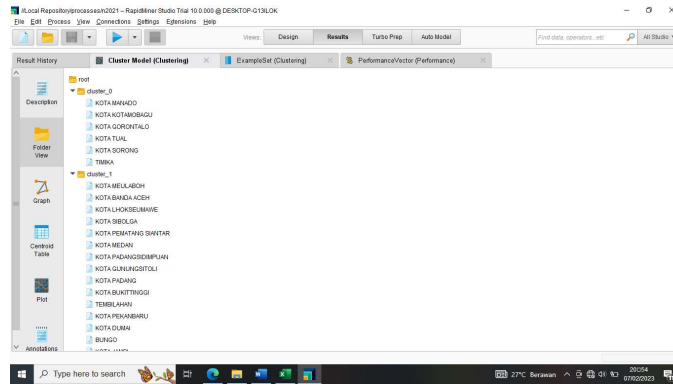


Gambar 7 Hasil pilihan Folder View Tahun 2020



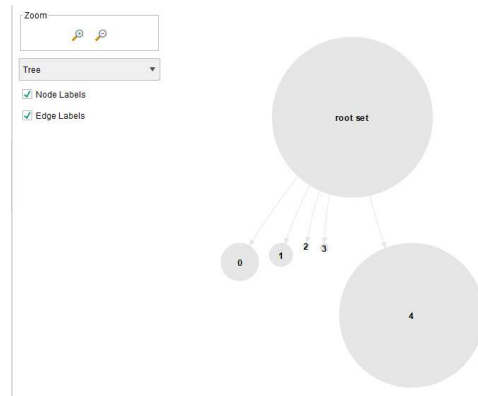
Gambar 8 Hasil pilihan Folder View Tahun 2021

Penerapan Algoritma K-Means Clustering Pada Tingkat Inflasi Kota Di Indonesia

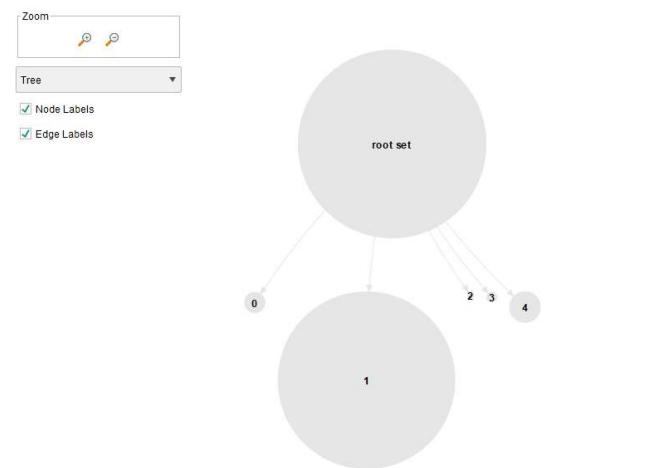


Gambar 9 Hasil pilihan Folder View Tahun 2022

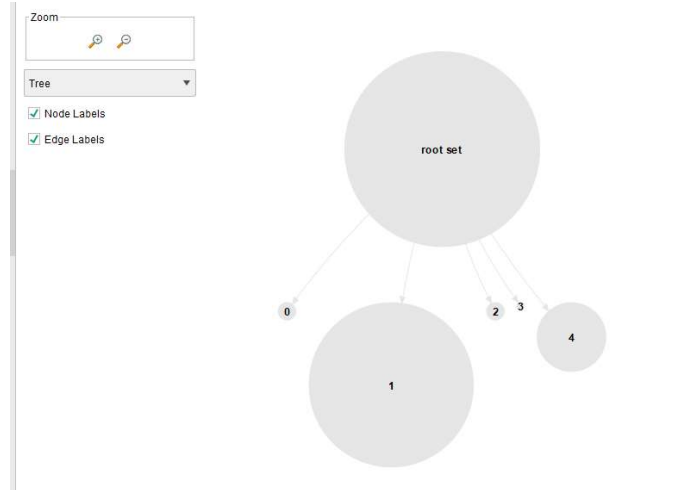
Pada gambar Gambar 10 terdapat hasil pilihan *Graph* pada tahun 2020, pada gambar Gambar 11 terdapat hasil pilihan *Graph* pada tahun 2021 dan pada Gambar 12 terdapat hasil pilihan *Graph* pada tahun 2022.



Gambar 10 Hasil pilihan Graph tahun 2020



Gambar 11 Hasil pilihan Graph tahun 2021



Gambar 12 Hasil pilihan Graph tahun 2022

Pada gambar Gambar 13 terdapat hasil pilihan *Centroid Table* pada tahun 2020, pada gambar Gambar 14 terdapat hasil pilihan *Centroid Table* pada tahun 2021 dan pada Gambar 15 terdapat hasil pilihan *Centroid Table* pada tahun 2022

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Januari	0.695	-0.145	1	-0.003	0.431
Februari	-0.066	0.015	0.810	1.000	0.354
Maret	0.135	-0.365	-1.910	-1.103	-0.040
April	-0.398	0.226	0.720	0.277	-0.027
Mei	0.483	-0.024	0.900	0.583	0.127
Juni	-0.016	0.533	0.920	0.113	0.279
Juli	-0.220	0.274	1.450	-0.840	-0.092
Agustus	0.062	0.224	0.410	0.127	-0.093
September	0.145	-0.051	-0.830	-0.387	-0.072
Oktober	0.486	-0.520	0.240	-0.713	0.135
November	0.353	0.176	0.030	0.850	0.339
Desember	1.080	0.474	0.350	0.887	0.459

Gambar 13 Hasil pilihan Centroid Table tahun 2020

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Januari	0.288	0.388	-0.580	0.143	-0.382
Februari	-0.072	-0.003	-0.210	0.467	0.241
Maret	0.162	0.062	0.930	0.910	-0.203
April	0.765	0.190	-0.280	-0.483	0.274
Mei	0.072	0.315	1.820	-0.250	0.508
Juni	0.360	-0.091	0.990	-0.207	0.244
Juli	0.768	0.084	-0.600	0.107	0.211
Agustus	-0.248	-0.052	-0.720	-0.493	-0.234
September	-0.498	0.000	-0.820	-0.517	0.032
Oktober	0.135	0.114	0.960	0.480	0.400
November	-0.162	0.383	0.680	0.080	0.751
Desember	1.212	0.651	1.020	1.713	0.628

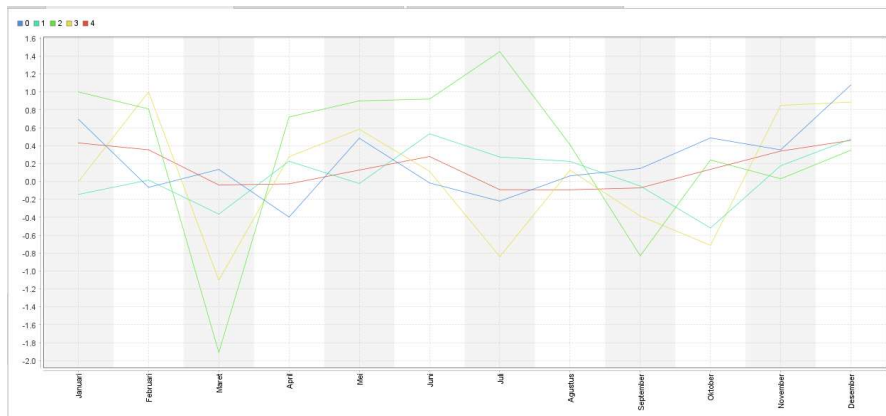
Gambar 14 Hasil pilihan Centroid Table tahun 2021

Penerapan Algoritma K-Means Clustering Pada Tingkat Inflasi Kota Di Indonesia

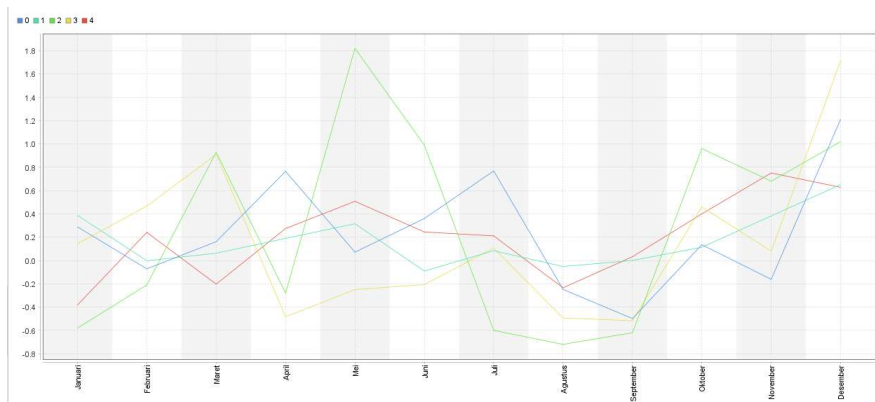
Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Januari	0.278	0.646	0.038	0.280	1.014
Februari	-0.038	-0.040	-0.900	-1.755	-0.294
Maret	0.942	0.771	0.222	1.315	0.907
April	0.866	1.116	1.200	1.825	0.930
Mei	0.304	0.520	0.832	1.890	0.859
Juni	2.008	0.686	0.298	0.340	0.886
Juli	0.884	0.620	1.806	0.845	0.827
Agustus	-1.046	-0.332	0.102	-1.220	-0.855
September	0.848	1.144	0.334	-0.120	0.944
Oktober	-0.512	-0.104	-0.192	-0.120	-0.347
November	0.012	0.130	0.576	0.150	-0.024
Desember	1.072	0.521	0.744	1.640	1.145

Gambar 15 Hasil pilihan Centroid Table tahun 2022

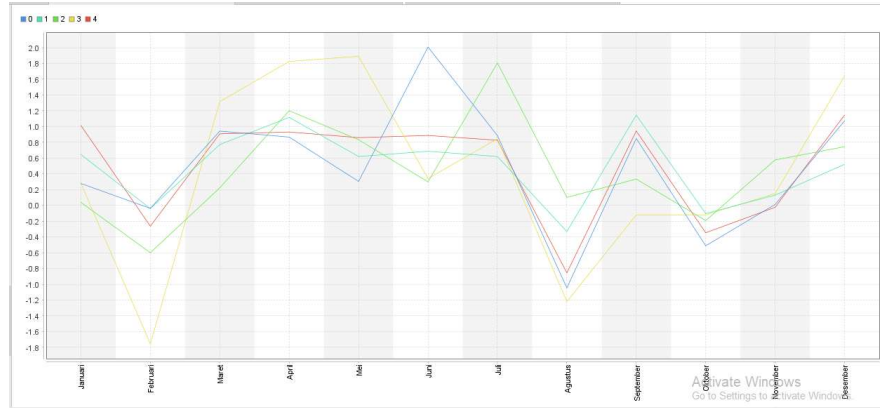
Pada gambar Gambar 16 terdapat hasil pilihan *Plot* pada tahun 2020, pada gambar Gambar 17 terdapat hasil pilihan *Plot* pada tahun 2021 dan pada Gambar 18 terdapat hasil pilihan *Plot* pada tahun 2022.



Gambar 16 Hasil pilihan Plot tahun 2020



Gambar 17 Hasil pilihan Plot tahun 2021



Gambar 18 Hasil pilihan Plot tahun 2022

Pada Tabel 4 terdapat hasil analisis *clustering* laju inflasi kota-kota di Indonesia dengan menggunakan algoritma k-means pada tahun 2020, pada Tabel 5 terdapat hasil analisis *clustering* laju inflasi kota-kota di Indonesia dengan menggunakan algoritma k-means pada tahun 2021 dan Pada Tabel 6 terdapat hasil analisis *clustering* laju inflasi kota-kota di Indonesia dengan menggunakan algoritma k-means pada tahun 2022.

Tabel 4 Hasil Clustering tahun 2020

Cluster 0	1. Kota Meulaboh	8. Kota Gunungsitoli
	2. Kota Banda Aceh	9. Kota Padang
	3. Kota Lhokseumawe	10. Kota Dumai
	4. Kota Sibolga	11. Tanjung Pandan
	5. Kota Pematang Siantar	12. Kota Pangkal Pinang
	6. Kota Medan	13. Waingapu
	7. Kota Padangsidimpuan	
Cluster 1	1. Kota Tarakan	5. Kota Bua-Bau
	2. Kota Manado	6. Kota Ambon
	3. Luwuk	7. Kota Sorong
	4. Kota Kendari	8. Kota Jayapura
Cluster 2	1. Timika	
Cluster 3	1. Kota Ternate	3. Merauke
	2. Manokwari	
Cluster 4	1. Kota Bukittinggi	34. Kota Madiun
	2. Tembilahan	35. Kota Surabaya
	3. Kota Pekanbaru	36. Kota Tangerang
	4. Bungo	37. Kota Cilegon
	5. Kota Jambi	38. Kota Serang
	6. Kota Palembang	39. Singaraja
	7. Kota Lubuklinggau	40. Kota Denpasar
	8. Kota Bengkulu	41. Kota Mataram
	9. Kota Bandar Lampung	42. Kota Bima
	10. Kota Metro	43. Maumere
	11. Kota Batam	44. Kota Kupang
	12. Kota Tanjung Pinang	45. Sintang
	13. Dki Jakarta	46. Kota Pontianak

14. Kota Bogor	47. Kota Singkawang
15. Kota Sukabumi	48. Sampit
16. Kota Bandung	49. Kota Palangka Raya
17. Kota Cirebon	50. Kota Baru
18. Kota Bekasi	51. Tanjung
19. Kota Depok	52. Kota Banjarmasin
20. Kota Tasikmalaya	53. Kota Balikpapan
21. Cilacap	54. Kota Samarinda
22. Purwokerto	55. Tanjung Selor
23. Kudus	56. Kota Kotamobagu
24. Kota Surakarta	57. Kota Palu
25. Kota Semarang	58. Bulukumba
26. Kota Tegal	59. Watampone
27. Kota Yogyakarta	60. Kota Makassar
28. Jember	61. Kota Pare-Pare
29. Banyuwangi	62. Kota Palopo
30. Sumenep	63. Kota Gorontalo
31. Kota Kediri	64. Mamuju
32. Kota Malang	65. Kota Tual
33. Kota Probolinggo	

Dari Tabel 4 Terlihat bahwa pada tahun 2020 ada 13 kota pada *cluster* 0, 8 kota pada *cluster* 1, 1 kota pada *cluster* 2, 3 kota pada *cluster* 3 dan sisanya berada di *cluster* 4 yaitu 65 kota, dengan jumlah total 90 kota. Sehingga dapat diketahui pada tahun 2020 indonesia mengalami laju inflasi yang cukup rendah

Tabel 5 Hasil Clustering tahun 2021

Cluster 0	1. Kota Manado	4. Kota Tual
	2. Kota Kotamobagu	5. Kota Sorong
	3. Kota Gorontalo	6. Timika
Cluster 1	1 Kota Meulaboh	37. Kota Semarang
	2 Kota Banda Aceh	38. Kota Tegal
	3 Kota Lhokseumawe	39. Kota Yogyakarta
	4 Kota Sibolga	40. Jember
	5 Kota Pematang Siantar	41. Banyuwangi
	6 Kota Medan	42. Sumenep
	7 Kota Padangsidempuan	43. Kota Kediri
	8 Kota Gunungsitoli	44. Kota Malang
	9 Kota Padang	45. Kota Probolinggo
	10 Kota Bukittinggi	46. Kota Madiun
	11 Tembilahan	47. Kota Surabaya
	12 Kota Pekanbaru	48. Kota Tangerang
	13 Kota Dumai	49. Kota Cilegon
	14 Bungo	50. Kota Serang
	15 Kota Jambi	51. Kota Denpasar
	16 Kota Palembang	52. Kota Mataram
	17 Kota Lubuklinggau	53. Kota Bima
	18 Kota Bengkulu	54. Waingapu
	19 Kota Bandar Lampung	55. Kota Kupang
	20 Kota Metro	56. Kota Pontianak

21	Tanjung Pandan	57.	Kota Palangka Raya
22	Kota Pangkal Pinang	58.	Kota Baru
23	Kota Batam	59.	Tanjung
24	Kota Tanjung Pinang	60.	Kota Balikpapan
25	Dki Jakarta	61.	Kota Samarinda
26	Kota Bogor	62.	Tanjung Selor
27	Kota Sukabumi	63.	Luwuk
28	Kota Bandung	64.	Kota Palu
29	Kota Cirebon	65.	Bulukumba
30	Kota Bekasi	66.	Watampone
31	Kota Depok	67.	Kota Makassar
32.	Kota Tasikmalaya	68.	Kota Pare-Pare
33	Cilacap	69	Kota Palopo
34	Purwokerto	70	Kota Kendari
35	Kudus	71	Mamuju
36	Kota Surakarta		
Cluster 2	1. Manokwari		
Cluster 3	1. Singaraja	3.	Kota Jayapura
	2. Merauke		
Cluster 4	1. Maumere	6.	Kota Tarakan
	2. Sintang	7.	Kota Bua-Bau
	3. Kota Singkawang	8.	Kota Ambon
	4. Sampit	9.	Kota Ternate
	5. Kota Banjarmasin		

Dari Tabel 5 Terlihat bahwa pada tahun 2021 ada 6 kota pada *cluster 0*, 71 kota pada *cluster 1*, 1 kota pada *cluster 2*, 3 kota pada *cluster 3* dan sisanya berada di *cluster 4* yaitu 9 kota, dengan jumlah total 90 kota. Sehingga dapat diketahui pada tahun 2020 indonesia mengalami laju inflasi yang cukup tinggi.

Tabel 6 Hasil Clustering tahun 2022

Cluster 0	1. Kota Gunungsitoli	4.	Kota Kotamobagu
	2. Kota Pekanbaru	5.	Kota Gorontalo
	3. Singaraja		
Cluster 1	1. Kota Banda Aceh	30.	Kota Tangerang
	2. Kota Padangsidempuan	31.	Kota Cilegon
	3. Kota Palembang	32.	Kota Serang
	4. Kota Lubuklinggau	33.	Kota Denpasar
	5. Kota Bengkulu	34.	Kota Mataram
	6. Kota Bandar Lampung	35.	Kota Bima
	7. Kota Metro	36.	Kota Kupang
	8. Kota Tanjung Pinang	37.	Sintang
	9. Dki Jakarta	38.	Kota Pontianak
	10. Kota Bogor	39.	Kota Singkawang
	11. Kota Sukabumi	40.	Sampit
	12. Kota Cirebon	41.	Kota Palangka Raya
	13. Kota Bekasi	42.	Kota Baru
	14. Kota Depok	43.	Tanjung
	15. Kota Tasikmalaya	44.	Kota Banjarmasin

16. Cilacap	45. Kota Balikpapan
17. Purwokerto	46. Kota Samarinda
18. Kudus	47. Tanjung Selor
19. Kota Surakarta	48. Kota Tarakan
20. Kota Semarang	49. Kota Manado
21. Kota Tegal	50. Luwuk
22. Kota Yogyakarta	51. Kota Palu
23. Jember	52. Bulukumba
24. Banyuwangi	53. Watampone
25. Kota Kediri	54. Kota Makassar
26. Kota Malang	55. Kota Pare-Pare
27. Kota Probolinggo	56. Kota Palopo
28. Kota Madiun	57. Kota Sorong
29. Kota Surabaya	58. Kota Jayapura
Cluster 2	
1. Kota Kendari	1. Kota Ternate
2. Kota Ambon	2. Timika
3. Kota Tual	
Cluster 3	
1. Tanjung Pandan	2. Manokwari
Cluster 4	
1. Kota Meulaboh	11. Kota Jambi
2. Kota Lhokseumawe	12. Kota Pangkal Pinang
3. Kota Sibolga	13. Kota Batam
4. Kota Pematang Siantar	14. Kota Bandung
5. Kota Medan	15. Sumenep
6. Kota Padang	16. Waingapu
7. Kota Bukittinggi	17. Maumere
8. Tembilahan	18. Kota Bua-Bau
9. Kota Dumai	19. Mamuju
10. Bungo	20. Merauke

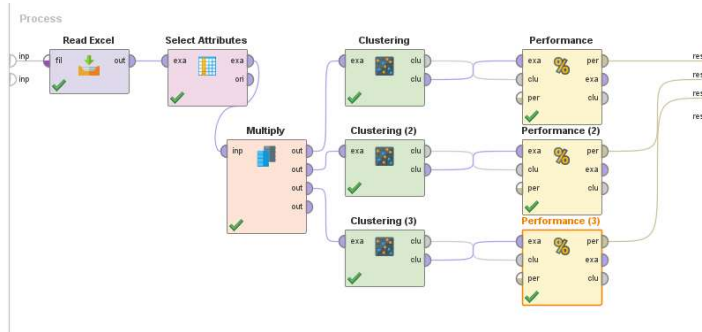
Dari Tabel 6 Terlihat bahwa pada tahun 2022 ada 5 kota pada *cluster* 0, 58 kota pada *cluster* 1, 5 kota pada *cluster* 2, 2 kota pada *cluster* 3 dan sisanya berada di *cluster* 4 yaitu 20 kota, dengan jumlah total 90 kota. Sehingga dapat diketahui pada tahun 2020 indonesia mengalami laju inflasi yang cukup tinggi.

B. Menentukan cluster optimum dan nilai DBI dari hasil pengelompokan

Davies Bouldin Indeks(DBI) adalah metode internal yang digunakan dalam evaluasi klaster. Keandalan internal DBI mengukur efektivitas pengelompokan dengan mengukur kuantitas dan kualitas data yang dikumpulkan.

Setelah mempersiapkan data, lanjutkan ke langkah berikutnya, pengolahan data. Data diolah menggunakan algoritma K-means. Pada penelitian ini jumlah *cluster* ditetapkan menjadi 3, 4, dan 5 dengan melakukan percobaan dengan jumlah *cluster* yang berbeda. Terakhir, jumlah cluster yang digunakan dalam penelitian ini ditentukan berdasarkan hasil evaluasi. Analisis data yang digunakan dalam penelitian ini

menggunakan RapidMiner sebagai perangkat lunak untuk melakukan proses analisis. Metode *clustering* dan algoritma K-Means pada aplikasi *Rapidminer* dapat dilihat Gambar 19.



Gambar 19 Model K-Means menentukan nilai DBI

Selanjutnya dilakukan evaluasi *cluster* untuk mengetahui berapa jumlah *cluster* dengan nilai DBI terendah, yaitu jumlah *cluster* dengan data yang paling mirip. Dari hasil percobaan dengan menggunakan 3, 4 dan 5 *cluster* pada *RapidMiner*, rekapitulasi nilai DBI yang diperoleh ditunjukkan dalam Tabel 7 untuk dataset pada tahun 2020, pada Tabel 8 untuk dataset pada tahun 2021 dan pada Tabel 9 untuk dataset pada tahun 2022.

Tabel 7 Davies Bouldin tahun 2020

Percobaan	Jumlah Cluster	Nilai DBI
1.	3	0.160
2.	4	0.133
3.	5	0.117

Berdasarkan hasil rekapitulasi *Davies Bouldin Index* (DBI) yang ditunjukkan pada Tabel, dengan hasil dari nilai k=3 yaitu 0.160, k=4 yaitu 0.133 dan k=5 yaitu 0.117, maka nilai k yang akan dipilih dan menjadi jumlah *cluster* adalah pada percobaan kedua dengan nilai k=5 karena memiliki nilai DBI paling kecil yaitu 0.117, karena semakin kecil nilai DBI maka semakin optimum cluster yang dihasilkan, sehingga pada penelitian ini cluster yang akan digunakan sebanyak 5 *cluster*. Adapun hasil clustering K-Means yang diperoleh dengan menggunakan 5 *cluster* sebagai berikut:



Gambar 20 Davies Bouldin tahun 2020

Tabel 8 Davies Bouldin tahun 2021

Percobaan	Jumlah Cluster	Nilai DBI
1.	3	0.145
2.	4	0.137
3.	5	0.129

Berdasarkan hasil rekapitulasi *Davies Bouldin Index* (DBI) yang ditunjukkan pada Tabel, dengan hasil dari nilai $k=3$ yaitu 0.145, $k=4$ yaitu 0.137 dan $k=5$ yaitu 0.129, maka nilai k yang akan dipilih dan menjadi jumlah cluster adalah pada percobaan kedua dengan nilai $k=5$ karena memiliki nilai DBI paling kecil yaitu 0.129, karena semakin kecil nilai DBI maka semakin optimum *cluster* yang dihasilkan, sehingga pada penelitian ini *cluster* yang akan digunakan sebanyak 5 *cluster*. Adapun hasil *clustering* K-Means yang diperoleh dengan menggunakan 5 cluster sebagai berikut



Gambar 21 Davies Bouldin tahun 2021

Tabel 9 Davies Bouldin tahun 2022

Percobaan	Jumlah Cluster	Nilai DBI
1.	3	0.167
2.	4	0.151
3.	5	0.145

Berdasarkan hasil rekapitulasi *Davies Bouldin Index* (DBI) yang ditunjukkan pada Tabel, dengan hasil dari nilai $k=3$ yaitu 0.167, $k=4$ yaitu 0.151 dan $k=5$ yaitu 0.145, maka nilai k yang akan dipilih dan menjadi jumlah *cluster* adalah pada percobaan kedua dengan nilai $k=5$ karena memiliki nilai DBI paling kecil yaitu 0.145, karena semakin kecil nilai DBI maka semakin optimum *cluster* yang dihasilkan, sehingga pada penelitian ini *cluster* yang akan digunakan sebanyak 5 *cluster*. Adapun hasil *clustering* K-Means yang diperoleh dengan menggunakan 5 *cluster* sebagai berikut:



Gambar 22 Davies Bouldin tahun 2022

KESIMPULAN DAN SARAN

Berdasarkan penjelasan pemodelan data mining dapat disimpulkan bahwa:

1. Dari hasil penelitian diketahui bahwa pada tahun 2020 ada 13 kota pada *cluster* 0, 8 kota pada *cluster* 1, 1 kota pada *cluster* 2, 3 kota pada *cluster* 3 dan sisanya berada di *cluster* 4 yaitu 65 kota, dengan jumlah total 90 kota. Sehingga dapat diketahui pada tahun 2020 indonesia mengalami laju inflasi yang cukup rendah. Pada tahun 2021 ada 6 kota pada *cluster* 0, 71 kota pada *cluster* 1, 1 kota pada *cluster* 2, 3 kota pada *cluster* 3 dan sisanya berada di *cluster* 4 yaitu 9 kota, dengan jumlah total 90 kota. Sehingga dapat diketahui pada tahun 2020 indonesia mengalami laju inflasi yang cukup tinggi. Pada tahun 2022 ada 5 kota pada *cluster* 0, 58 kota pada *cluster* 1, 5 kota pada *cluster* 2, 2 kota pada *cluster* 3 dan sisanya berada di *cluster* 4 yaitu 20 kota, dengan jumlah total 90 kota. Sehingga dapat diketahui pada tahun 2020 indonesia mengalami laju inflasi yang cukup tinggi.
2. Dari pencarian nilai DBI hasil mengelompokan laju inflasi di Indonesia menggunakan algoritma *K-Means clustering* dengan menggunakan lima *cluster* maka menghasilkan nilai *Davies Bouldin* atau DBI sebesar: 0.117 pada dataset inflasi tahun 2020, nilai DBI sebesar 0.129 pada tahun 2021 dan nilai dbi sebesar 0.145 pada tahun 2022.

Saran yang disampaikan oleh peneliti dalam penelitian ini adalah pada peneliti selanjutnya untuk menggunakan aplikasi dan algoritma yang berbeda dalam teknik data mining mengelompokkan laju inflasi kota-kota di Indonesia.

DAFTAR REFERENSI

- Darmansah, D. D. (2021). Analisis Penyebaran Penularan Virus Covid-19 di Provinsi Jawa Barat Menggunakan Algoritma K-Means Clustering. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(3), 1188–1199. <https://doi.org/10.35957/jatisi.v8i3.1034>
- Fuady, M., & Nugraha, J. (2017). Implementasi Metode K-Means Dan K-Medoids Untuk Mengelompokkan 82 Kota Di Indonesia Berdasarkan Indeks Harga Konsumen. *Prosiding Seminar Nasional Seri 7*, 327–337.
- Handayani, F. (2022). Aplikasi Aplikasi Data Mining Menggunakan Algoritma K-Means Clustering untuk Mengelompokkan Mahasiswa Berdasarkan Gaya Belajar. *Jurnal Teknologi Dan Informasi*, 12(1), 46–63. <https://doi.org/10.34010/jati.v12i1.6733>
- Hardiani, T. (2022). Analisis Clustering Kasus Covid 19 di Indonesia Menggunakan Algoritma K-Means. *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 11(2), 156–165. <https://doi.org/10.23887/janapati.v11i2.45376>
- Mulyani, R. (2020). Inflasi dan Cara Mengatasinya dalam Islam. *Lisyabab : Jurnal Studi Islam Dan Sosial*, 1(2), 267–278. <https://doi.org/10.58326/jurnallisyabab.v1i2.47>
- Muttaqin, M. R., & Defriani, M. (2020). Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa. *ILKOM Jurnal Ilmiah*, 12(2), 121–129. <https://doi.org/10.33096/ilkom.v12i2.542.121-129>
- Nabila, Z., Rahman Isnain, A., & Abidin, Z. (2021). ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS. *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 2(2), 100. <http://jim.teknokrat.ac.id/index.php/JTISI>
- Priyatman, H., Sajid, F., & Haldivany, D. (2019). Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(1), 62. <https://doi.org/10.26418/jp.v5i1.29611>
- Santosa, A. B. (2017). Analisis Inflasi di Indonesia. *Prosiding Seminar Nasional Multi Disiplin Ilmu & Call Papers UNISBANK Ke-3 (SENDI_U 3) 2017*, 445–452.
- Sudibyoy, N. A. (2020). Implementasi Metode K-means untuk Mengelompokkan Tingkat Inflasi di Indonesia Article Penerapan Multimedia Pembelajaran Interaktif Elektronika dengan Framework RAD (Rapid Application Development) Menggunakan HTML View project. <https://www.researchgate.net/publication/353767505>