

Analisis Algoritma K-Means Clustering Pada Tingkat Inflasi Selama Covid-19

Ummy Sholihah¹, Fendy², Firnaldo³

^{1,2,3}Program Studi : Sistem Informasi, Universitas Darwan Ali

Email : ummy.shlhh@gmail.com¹, fendyendy353@gmail.com², fernandoando238@gmail.com³

ABSTRACT— COVID-2019 is a virus that has caused an uproar in the world. COVID-19 can infect the human respiratory system. On December 31, 2019, the WHO China Country Office reported a case of pneumonia of unknown cause in Wuhan City, China. Inflation is an increase in the price of goods and services that occurs within a certain period of time. In this study, using the data that has been obtained, a grouping of inflation rates will be made in areas exposed to the COVID-19 virus during 2019-2021, especially in the city of Sampit where this case study was conducted. The formulation of the problem was carried out to obtain a complete picture of the scope of the problem and the right steps in solving it and to make this research directed towards knowing the inflation rate that occurred in East Kotawaringin. Literature study is carried out by searching and collecting previous literature and studying it with the aim of obtaining supporting references in the form of research journals and books on data mining. Data sources were taken from the publication of the Official Statistical News (BRS) through the official website of the Central Bureau of Statistics for East Kotawaringin Regency. Data cleaning is a process for eliminating or finding missing values in a record and eliminating noise. K-Means Clustering is one of the simplest clustering algorithms. Next, implementation is carried out using software to see the comparison and suitability of the final decision results obtained from the K-Means Clustering algorithm using the Weka application. Conclusions were drawn from the research that has been done, where these conclusions can be used as a reference in knowing the inflation rate before and after Covid-19 in East Kotawaringin Regency based on the expenditure group. From the results of the research above, it can be concluded that three groups of inflation rates have been obtained, namely C0 = moderate, C1 = slow, and C2 = high.

Keywords— Clustering, Covid-19, Data Mining, Inflation, K-Means.

ABSTRAK— COVID-2019 merupakan virus yang sempat membuat gempar dunia. COVID-19 dapat menginfeksi sistem pernapasan manusia. Pada 31 Desember 2019, WHO China Country Office melaporkan kasus pneumonia yang tidak diketahui penyebabnya di Kota Wuhan, China. Inflasi adalah kenaikan harga barang dan jasa yang terjadi dalam kurun waktu tertentu. Pada penelitian ini melalui data yang telah didapatkan akan dibuat pengelompokan tingkat laju inflasi di daerah yang terpapar virus COVID-19 selama tahun 2019-2021 khususnya daerah kota Sampit dimana studi kasus ini dilakukan. Perumusan masalah dilakukan untuk memperoleh gambaran yang lengkap tentang lingkup masalah dan langkah yang tepat dalam pemecahannya serta membuat penelitian ini terarah untuk mengetahui tingkat inflasi yang terjadi di Kotawaringin Timur. Studi literatur dilakukan dengan pencarian dan pengumpulan literatur terdahulu dan mempelajarinya dengan tujuan yaitu mendapatkan referensi pendukung baik berupa jurnal penelitian maupun buku-buku mengenai data mining. Sumber data diambil dari publikasi Berita Resmi Statistik (BRS) melalui website resmi Badan Pusat Statistik Kabupaten Kotawaringin Timur. Data cleaning adalah suatu proses untuk menghilangkan atau menemukan missing value pada suatu record dan menghilangkan noise. K-Means Clustering merupakan salah satu algoritma dalam clustering yang paling sederhana. Selanjutnya dilakukan implementasi dengan menggunakan perangkat lunak untuk melihat perbandingan dan kesesuaian dari hasil keputusan akhir yang didapatkan dari algoritma K-Means Clustering dengan menggunakan aplikasi Weka. Ditarik kesimpulan dari penelitian yang telah dilakukan, dimana kesimpulan ini dapat menjadi acuan dalam mengetahui tingkat inflasi sebelum dan setelah Covid-19 di Kabupaten Kotawaringin Timur berdasarkan kelompok pengeluarannya. Dari hasil penelitian di atas dapat ditarik kesimpulan telah didapatkan tiga kelompok tingkat kelajuan inflasi yaitu C0 = sedang, C1 = lambat, dan C2 = tinggi.

Kata kunci— Clustering, Covid-19, Data Mining, Inflasi, K-Means.

I. PENDAHULUAN

COVID-2019 merupakan virus yang sempat membuat gempar dunia. COVID-19 dapat menginfeksi sistem pernapasan manusia. Virus ini dapat menyebabkan pernapasan berat seperti pneumonia. Ketika menyerang manusia virus ini menyebabkan infeksi saluran pernafasan seperti flu, pusing, hingga penyakit yang lebih berat seperti

Middle East Respiratory Syndrome (MERS) dan Severe Acute Respiratory Syndrome (SARS) [1].

Pada 31 Desember 2019, WHO China Country Office melaporkan kasus pneumonia yang tidak diketahui penyebabnya di Kota Wuhan, China. Pada tanggal 9 Maret 2020 WHO secara resmi mengumumkan COVID-19 sebagai pandemi. Pada 2 Maret 2020 COVID-19 dilaporkan masuk ke Indonesia selang 2 bulan tepatnya pada 9 Mei 2020 COVID-19 di Indonesia mengalami

peningkatan kasus terinfeksi COVID-19. Hingga akhir tahun 2021 kasus terkonfirmasi terinfeksi virus corona mencapai 4.262.720 jiwa. Untuk mewaspadai bertambahnya jumlah jiwa yang terinfeksi virus COVID-19 maka pemerintah menetapkan PPKM terutama pada daerah yang memiliki angka kematian yang tinggi.

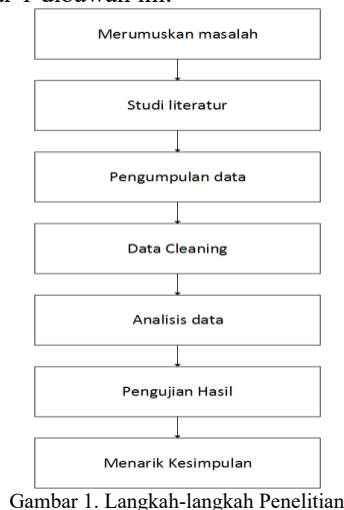
Pandemi covid-19 tidak hanya berdampak terhadap kondisi kesehatan manusia namun juga mengganggu kesehatan ekonomi dunia. Akibat pandemi covid-19 Indonesia cukup terhantam keras dengan penyebaran covid-19. Tingginya tingkat pengangguran sebagai dampak dari pemutusan hubungan kerja, pembatasan ruang gerak yang menyebabkan tersendatnya roda perekonomian, turunnya pendapatan masyarakat, permintaan yang tidak stabil, hingga tingginya bahan pangan. Kondisi ini menyebabkan tingginya tingkat inflasi dan terjadinya instabilitas ekonomi dalam kurun waktu yang tidak dapat dipastikan [2].

Inflasi adalah kenaikan harga barang dan jasa yang terjadi dalam kurun waktu tertentu. Inflasi merupakan indikator untuk melihat tingkat perubahan dan dianggap terjadi jika proses kenaikan harga berlangsung secara terus-menerus dan saling pengaruh-memengaruhi [3]. Inflasi didefinisikan sebagai kenaikan umum tingkat harga dalam suatu perekonomian selama periode waktu tertentu, yang mengakibatkan penurunan berkelanjutan dalam daya beli uang [4].

Pada penelitian ini melalui data yang telah didapatkan akan dibuat pengelompokan tingkat laju inflasi di daerah yang terpapar virus COVID-19 selama tahun 2019-2021 khususnya daerah kota Sampit dimana studi kasus ini dilakukan. Metode yang akan digunakan dalam penelitian ini yaitu metode K-Means Clustering. K-Means merupakan salah satu algoritma Unsupervised Learning yang digunakan dalam mengelompokan data ke dalam beberapa kelompok dengan sistem partisi [5].

II. METODOLOGI PENELITIAN

Metodologi penelitian akan menjelaskan mengenai tahapan-tahapan penelitian secara sistematis sehingga dapat digunakan sebagai pedoman yang jelas dalam menyelesaikan dan membuat analisis dari hasil penelitian yang dilakukan. Adapun metode penelitian dapat dilihat pada Gambar 1 dibawah ini:



Gambar 1. Langkah-langkah Penelitian

Kerangka kerja digambarkan sebagai langkah-langkah atau cara sistematis atau metode yang dilakukan pada penelitian ini, agar langkah-langkah yang diambil peneliti dalam penelitian dapat lebih mudah dipahami. Untuk lebih jelasnya berikut ini dibuat rincian secara lebih detail mengenai metodologi pemecahan masalah dari tahapan kerangka kerja pada Gambar 1.

A. Merumuskan Masalah

Langkah pertama yang harus dilakukan yaitu menentukan masalah yang akan menjadi latar belakang dilakukannya penelitian ini. Perumusan masalah dilakukan untuk memperoleh gambaran yang lengkap tentang lingkup masalah dan langkah yang tepat dalam pemecahannya serta membuat penelitian ini terarah untuk mengetahui tingkat inflasi yang terjadi di Kotawaringin Timur.

B. Studi Literatur

Studi literatur dilakukan dengan pencarian dan pengumpulan literatur terdahulu dan mempelajarinya dengan tujuan yaitu mendapatkan referensi pendukung baik berupa jurnal penelitian maupun buku-buku mengenai data mining, algoritma K-Means dan software Weka serta literatur yang membahas tentang data perkembangan ekonomi selama persebaran virus corona untuk memperkuat pengetahuan dasar dan teori pendukung.

C. Pengumpulan Data

Tahapan ini dapat dilakukan dengan mencari tahu data apa saja yang dibutuhkan untuk menunjang perhitungan dalam penelitian ini. Sumber data diambil dari publikasi Berita Resmi Statistik (BRS) melalui website resmi Badan Pusat Statistik Kabupaten Kotawaringin Timur. Data tersebut merupakan data yang berisi nilai IHK dan juga inflasi yang terjadi selama berlangsungnya covid-19 di tahun 2019 hingga 2021 beserta dengan kelompok pengeluarannya. Data yang digunakan pada penelitian ini hanya data inflasi tahun kalender 2019, 2020 dan 2021. Untuk rincian data dapat dilihat pada Gambar 3, Gambar 4, dan Gambar 5 di bawah ini.

Kelompok Pengeluaran	Indeks Harga Konsumen (IHK)			Inflasi November 2019 ^(*)	Inflasi Tahun Kalender 2019 ^(**)	Inflasi Tahun ke Tahun ^(***)
	November 2018	Desember 2018	November 2019			
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Umum	137,04	139,06	140,88	0,26	1,31	2,80
1 BAHAN MAKANAN	141,84	144,46	148,05	1,36	2,49	4,38
2 MAKANAN JADI, MINUMAN, ROKOK & TEMBAKAU	142,52	142,57	145,22	0,39	1,86	1,89
3 PERUMAHAN, AIR, LISTRIK, GAS & BAHAN BAKAR	130,65	130,84	131,58	0,01	0,57	0,71
4 SANDANG	121,12	121,57	125,55	0,02	3,27	3,66
5 KESEHATAN	123,43	123,51	126,78	0,20	2,65	2,71
6 PENDIDIKAN, REKREASI DAN OLAH RAGA	150,52	150,36	156,83	0,03	4,30	4,19
7 TRANSPORT, KOMUNIKASI DAN JASA KEUANGAN	138,96	146,62	143,25	-1,02	-2,30	3,09

Gambar 2. IHK dan Inflasi/Deflasi Kota Sampit Tahun 2019 Menurut Kelompok Pengeluaran

Kelompok Pengeluaran	Indeks Harga Konsumen (IHK)			Inflasi November 2020 (%)	Inflasi Tahun Kalendar 2020 (%)	Inflasi Tahun ke Tahun (%)	Andil Inflasi November 2020 (%)
	November 2019	Desember 2019	November 2020				
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
UMUM	103,41	103,82	105,32	0,52	1,44	1,85	0,52
MAKANAN, MINUMAN DAN TEMBAKAU	104,77	105,04	107,71	0,89	2,54	2,81	0,33
PAKAIAN DAN ALAS KAKI	104,73	104,73	105,02	0,00	0,28	0,28	0,00
PERUMAHAN, AIR, LISTRIK, GAS DAN BAHAN BAKAR LAINNYA	99,73	99,71	99,34	0,35	-0,37	-0,39	0,05
PERLENGKAPAN, PERALATAN DAN PEMELIHARAAN RUTIN RUMAH TANGGA	102,72	102,76	103,31	-0,24	0,54	0,57	-0,01
KESEHATAN	111,94	111,97	112,42	0,12	0,40	0,43	0,00
TRANSPORTASI	100,78	103,38	104,10	1,12	0,70	3,29	0,13
INFORMASI, KOMUNIKASI, DAN JASA KEUANGAN	98,78	98,66	97,50	0,01	-1,18	-1,30	0,00
REKREASI, OLARAHAGA, DAN BUDAYA	99,59	99,59	99,92	0,00	0,33	0,33	0,00
Pendidikan	107,04	107,04	105,37	0,00	-1,56	-1,56	0,00
Penyediaan Makanan dan Minuman/Restoran	107,61	107,61	108,41	0,00	0,74	0,74	0,00
Perawatan Pribadi dan Jasa Lainnya	106,02	106,16	114,96	0,31	8,29	8,43	0,02

Gambar 3. IHK dan Inflasi/Deflasi Kota Sampit Tahun 2020 Menurut Kelompok Pengeluaran

Kelompok Pengeluaran	IHK November 2020	IHK Desember 2020	IHK November 2021	Tingkat Inflasi November 2021 ^a (%)	Tingkat Inflasi Tahun Kalendar 2021 ^b (%)	Tingkat Inflasi Tahun ke Tahun ^c (%)	Andil Inflasi November 2021 (%)
Umum (Headline)	105,32	105,50	109,65	0,32	3,93	4,11	0,32
Makanan, Minuman, dan Tembakau	107,71	108,03	111,44	0,53	3,16	3,46	0,19
Pakaian dan Atas Kaki	105,02	104,99	106,07	0,00	1,03	1,00	0,00
Perumahan, Air, Listrik, dan Bahan Bakar Rumah Tangga	99,34	99,36	113,66	0,26	14,39	14,42	0,04
Perengkapan, Peralatan, dan Pemeliharaan Rutin Rumah Tangga	103,31	103,35	106,78	0,39	3,32	3,36	0,02
Kesehatan	112,42	112,45	138,81	0,00	23,44	23,47	0,00
Transportasi	104,10	104,62	101,08	0,20	-3,38	-2,90	0,02
Informasi, Komunikasi, dan Jasa Keuangan	97,50	97,56	97,76	0,22	0,21	0,27	0,01
Rekreasi, Olahraga, dan Budaya	99,92	99,92	101,23	1,05	1,31	1,31	0,01
Pendidikan	105,37	105,37	109,95	0,00	4,35	4,35	0,00
Penyediaan Makanan dan Minuman/Restoran	108,41	108,51	110,99	0,24	2,29	2,38	0,02
Perawatan Pribadi dan Jasa Lainnya	114,96	114,66	116,24	0,02	1,38	1,11	0,00

Gambar 4. IHK dan Inflasi/Deflasi Kota Sampit Tahun 2021 Menurut Kelompok Pengeluaran

D. Data Cleaning

Data cleaning adalah suatu proses untuk menghilangkan atau menemukan missing value pada suatu record dan menghilangkan noise. Proses ini mencakup membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lainnya yang lebih relevan dan diperlukan untuk KDD (Knowledge Discovery In Databases), seperti data atau informasi eksternal.

E. Analisis Data

Untuk analisis data digunakan salah satu metode clustering yaitu K-Means. K-Means Clustering merupakan salah satu algoritma dalam clustering yang paling sederhana. K-means clustering melakukan proses pemodelan menggunakan unsupervised learning dan mengelompokkan data menggunakan sistem partisi. Secara umum, terdapat dua jenis clustering yang digunakan dalam proses pengelompokkan data, yaitu Hierarchical dan Non-Hierarchical.

K-Means merupakan salah satu metode clustering Non-Hierarchical atau Partitional Clustering yang secara umum dilakukan dengan algoritma sebagai berikut:

1. Menentukan jumlah cluster (k).
2. Menentukan titik pusat (centroid) cluster awal secara random.
3. Menghitung jarak centroid dengan setiap data. Ukuran jarak yang digunakan adalah Euclidean distance dengan rumus berikut [6].

$$D(i, f) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

4. Data akan dimasukkan ke dalam cluster berdasarkan jarak centroid terdekat. Lalu hitung kembali centroid cluster yang baru. Centroid cluster merupakan rata-rata semua data dalam sebuah cluster, dapat dihitung menggunakan rumus berikut.

$$V(i, j) = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

5. Ulangi penghitungan jarak centroid dengan setiap data (kembali pada langkah ke-3) jika masih terdapat data yang berpindah cluster atau apabila perubahan nilai centroid berada di atas nilai threshold yang dibentuk atau jika perubahan nilai pada objective function yang digunakan di atas nilai threshold yang ditentukan. Kembali ke langkah 3, jika masih terdapat data yang berpindah cluster atau apabila perubahan nilai centroid berada di atas nilai threshold yang dibentuk atau jika perubahan nilai pada objective function yang digunakan di atas nilai threshold yang ditentukan [7].

Jadi, pada tahap ini dilakukan analisa terhadap data ekonomi khususnya inflasi yang terjadi sebelum dan setelah Covid-19 di Kotawaringin Timur. Data tersebut akan melalui perhitungan menggunakan algoritma K-Means Clustering.

Pada analisis data dilakukan perhitungan manual menggunakan metode yang telah diterangkan pada paragraf sebelumnya dan selanjutnya dilakukan analisis data lagi menggunakan software WEKA. Waikato Environment for Knowledge Analysis (WEKA) adalah sebuah perangkat lunak yang memiliki banyak algoritma machine learning untuk keperluan data mining. WEKA juga memiliki banyak tools untuk pengolahan data, mulai dari pre-processing, classification, regression, clustering, association rules, dan visualization.

Weka digunakan untuk mempercepat pengolahan data dengan algoritma K-Means Clustering pada pertumbuhan ekonomi sebelum dan sesudah tersebarnya virus corona di Kalimantan Tengah khususnya di Kota Sampit yang mana data tersebut di peroleh melalui situs web Badan Pusat Statistik Kabupaten Kotawaringin Timur.

F. Menguji Hasil

Proses ini dilakukan dengan menganalisa apakah hasil pengolahan dan perhitungan yang dilakukan secara manual sudah benar. Untuk melihat perbandingan dan kesesuaian dari hasil keputusan akhir yang didapatkan, akan digunakan salah satu aplikasi yaitu Weka.

G. Menarik Kesimpulan

Pada tahap ini akan ditarik kesimpulan dari penelitian yang telah dilakukan, dimana kesimpulan ini dapat menjadi acuan dalam mengetahui tingkat inflasi sebelum dan setelah Covid-19 di Kabupaten Kotawaringin Timur berdasarkan kelompok pengeluarannya.

III. DESAIN, HASIL DAN PEMBAHASAN

Pada data yang telah dikelompokkan menjadi satu tabel seperti yang tertera pada Tabel 1. Diketahui terdapat beberapa data yang kosong pada atribut 2019 dengan

kelompok pengeluarannya yaitu, Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga, Penyediaan Makanan dan Minuman/Restoran, serta Perawatan Pribadi dan Jasa Lainnya. Dari tiga data yang masih kosong tersebut maka terlebih dahulu akan dilakukan proses data cleaning dengan menghitung nilai missing value.

TABEL I
 DATA KELOMPOK PENGELUARAN

Kelompok Pengeluaran	2019	2020	2021
Makanan, Minuman dan Tembakau	1.86	2.54	3.16
Pakaian dan Alas Kaki	3.27	0.28	1.03
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	0.57	-0.37	14.39
Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga		0.54	3.32
Kesehatan	2.65	0.40	23.44
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	-2.30	-0.48	-3.17
Pendidikan, Rekreasi, Olahraga dan Budaya	4.30	-1.23	5.66
Penyediaan Makanan dan Minuman/Restoran		0.74	2.29
Perawatan Pribadi dan Jasa Lainnya		8.29	1.38

Nilai untuk mengganti missing value atribut 2019:

$$x(2019) = \frac{\sum_1^9 x(2019)}{6} = \frac{10,35}{6} = 1,73$$

TABEL II
 HASIL SETELAH MENGHITUNG MISSING VALUE

Kelompok Pengeluaran	2019	2020	2021
Makanan, Minuman dan Tembakau	1.86	2.54	3.16
Pakaian dan Alas Kaki	3.27	0.28	1.03
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	0.57	-0.37	14.39
Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	1.73	0.54	3.32
Kesehatan	2.65	0.40	23.44
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	-2.30	-0.48	-3.17
Pendidikan, Rekreasi, Olahraga dan Budaya	4.30	-1.23	5.66
Penyediaan Makanan dan Minuman/Restoran	1.73	0.74	2.29
Perawatan Pribadi dan Jasa Lainnya	1.73	8.29	1.38

A. Perhitungan K-Means Clustering

Perhitungan K-Means Clustering dimulai dengan menentukan terlebih dahulu jumlah cluster yang digunakan pada data laju infansi yaitu sebanyak 3 Cluster. Cluster tersebut diantaranya (C0) sedang, (C1) rendah dan (C2) tinggi. Selanjutnya menentukan titik awal centroid secara random untuk digunakan pada perhitungan clustering.

TABEL III
 TITIK AWAL CENTROID

Titik Centroid	X	Y	Z
C0	1.86	2.54	3.16
C1	1.73	0.54	3.32
C2	0.57	-0.37	14.39

Menghitung jarak setiap data yang ada terhadap setiap pusat cluster:

a. Jarak antara data kelompok pengeluaran pertama dengan pusat cluster pertama.

$$C0 = \sqrt{(1,86 - 1,86)^2 + (2,54 - 2,54)^2 + (3,16 - 3,16)^2} = 0$$

b. Jarak antara data kelompok pengeluaran pertama dengan pusat cluster kedua.

$$C1 = \sqrt{(1,86 - 1,73)^2 + (2,54 - 0,54)^2 + (3,16 - 3,32)^2}$$

$$C1 = \sqrt{(0,13)^2 + (2)^2 + (-0,16)^2}$$

$$C1 = \sqrt{4,0425}$$

$$C1 = 2,01$$

c. Jarak antara data kelompok pengeluaran pertama dengan pusat cluster ketiga.

$$C2 = \sqrt{(1,86 - 0,57)^2 + (2,54 - (-0,37))^2 + (3,16 - 14,39)^2}$$

$$C2 = \sqrt{(1,29)^2 + (2,91)^2 + (-11,23)^2}$$

$$C2 = \sqrt{136,2451}$$

$$C2 = 11,67$$

Data-data pada kelompok pengeluaran akan dihitung dengan semua pusat cluster (titik centroid) yang ada.

TABEL IV
 HASIL ITERASI 1

Kelompok Pengeluaran	C0	C1	C2
Makanan, Minuman dan Tembakau	0	2.01	11.67
Pakaian dan Alas Kaki	3.41	2.77	13.65
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	11.67	11.17	0

Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	2.01	0	11.17
Kesehatan	20.41	20.14	9.32
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	8.15	7.71	17.79
Pendidikan, Rekreasi, Olahraga dan Budaya	5.14	3.90	9.53
Penyediaan Makanan dan Minuman/Restoran	2.00	1.05	12.21
Perawatan Pribadi dan Jasa Lainnya	6.02	7.99	15.67

Untuk mengetahui data yang dimiliki termasuk ke dalam cluster C0, C1 atau C2 yaitu dengan menentukan jarak terpendek dari hasil perhitungan pada Tabel 3.

TABEL V
 JARAK TERPENDEK DAN PENGELOMPOKKAN MASING-MASING CLUSTER

Kelompok Pengeluaran	Jarak Terpendek	Cluster
Makanan, Minuman dan Tembakau	0	C0
Pakaian dan Alas Kaki	2.77	C1
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	0	C2
Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	0	C1
Kesehatan	9.32	C2
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	7.71	C1
Pendidikan, Rekreasi, Olahraga dan Budaya	3.90	C1
Penyediaan Makanan dan Minuman/Restoran	1.05	C1
Perawatan Pribadi dan Jasa Lainnya	6.02	C0

Iterasi pertama menghasilkan kelompok C0 sebanyak 2 cluster, kelompok C1 sebanyak 5 cluster dan kelompok C2 sebanyak 2 cluster. Langkah selanjutnya adalah menentukan titik centroid baru dengan cara menjumlahkan kelompok dengan cluster yang sama dan dibagi dengan total data pada masing-masing cluster. Sehingga didapatkan hasil untuk titik centroid baru seperti yang telah dimuat pada Tabel 5.

TABEL VI
 TITIK CENTROID BARU

Titik Centroid	X	Y	Z
C0	1.80	5.42	2.27
C1	1.75	-0.03	1.83
C2	1.61	0.02	18.92

Langkah selanjutnya yaitu melakukan kembali iterasi berikutnya dengan langkah yang sama seperti iterasi pertama. Untuk iterasi kedua gunakan titik centroid baru.

TABEL VII
 HASIL ITERASI 2

Kelompok Pengeluaran	C0	C1	C2
Makanan, Minuman dan Tembakau	3.01	2.90	15.96
Pakaian dan Alas Kaki	5.48	1.75	17.96
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	13.49	12.62	4.66
Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	4.99	1.60	15.60
Kesehatan	21.77	21.64	4.66
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	9.01	6.44	22.43
Pendidikan, Rekreasi, Olahraga dan Budaya	7.87	4.76	13.58
Penyediaan Makanan dan Minuman/Restoran	4.68	0.90	16.64
Perawatan Pribadi dan Jasa Lainnya	3.01	8.33	19.39

TABEL VIII
 JARAK TERPENDEK DAN PENGELOMPOKKAN MASING-MASING CLUSTER

Kelompok Pengeluaran	Jarak Terpendek	Cluster
Makanan, Minuman dan Tembakau	2.90	C1
Pakaian dan Alas Kaki	1.75	C1
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	4.66	C2
Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	1.60	C1
Kesehatan	4.66	C2
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	6.44	C1
Pendidikan, Rekreasi, Olahraga dan Budaya	4.76	C1
Penyediaan Makanan dan Minuman/Restoran	0.90	C1
Perawatan Pribadi dan Jasa Lainnya	3.01	C0

Iterasi masih harus dilakukan karena masih terjadi perbedaan hasil antara iterasi 1 dan iterasi 2. Sebelumnya, titik centroid akan kembali dicari tahu terlebih dahulu agar dapat digunakan pada perhitungan iterasi 3.

TABEL IX
 TITIK CENTROID BARU

Titik Centroid	X	Y	Z
C0	1.73	8.29	1.38
C1	1.77	0.40	2.05
C2	1.61	0.02	18.92

TABEL X
 HASIL ITERASI 3

Kelompok Pengeluaran	C0	C1	C2
Makanan, Minuman dan Tembakau	6.02	2.41	15.96
Pakaian dan Alas Kaki	8.17	1.83	17.97
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	15.67	12.42	4.66
Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	7.99	1.28	15.61
Kesehatan	23.45	21.41	4.65
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	10.67	6.67	22.44
Pendidikan, Rekreasi, Olahraga dan Budaya	10.75	4.71	13.59
Penyediaan Makanan dan Minuman/Restoran	7.60	0.42	16.65
Perawatan Pribadi dan Jasa Lainnya	0	7.92	19.39

TABEL XI
 JARAK TERPENDEK DAN PENGELOMPOKKAN MASING-MASING CLUSTER

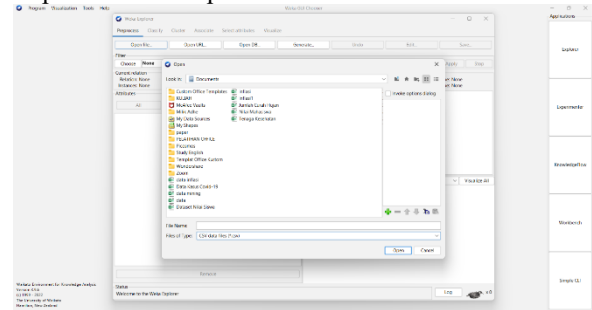
Kelompok Pengeluaran	Jarak Terpendek	Cluster
Makanan, Minuman dan Tembakau	2.41	C1
Pakaian dan Alas Kaki	1.83	C1
Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	4.66	C2
Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	1.28	C1
Kesehatan	4.65	C2
Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	6.67	C1
Pendidikan, Rekreasi, Olahraga dan Budaya	4.71	C1
Penyediaan Makanan dan Minuman/Restoran	0.42	C1
Perawatan Pribadi dan Jasa Lainnya	0	C0

Dalam metode K-Means, perhitungan akan berhenti apabila cluster pada suatu iterasi telah menghasilkan nilai yang sama seperti iterasi sebelumnya. Di sini iterasi

berhenti pada iterasi ketiga dengan hasil sama seperti iterasi 2. Untuk kelompok C0 sebanyak 1 cluster, kelompok C1 sebanyak 6 cluster dan kelompok C2 sebanyak 2 cluster.

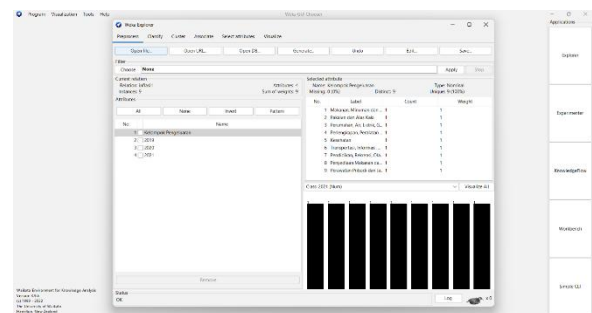
B. Clustering Menggunakan WEKA

Setelah melakukan perhitungan manual dengan menghitung tiap-tiap cluster dan juga menggunakan excel, selanjutnya akan dilakukan pengujian menggunakan aplikasi pengujian data mining, yaitu Weka. Tujuan dari penggunaan software ini adalah membandingkan hasil dengan perhitungan secara teoritis dengan hasil yang didapatkan melalui proses di Weka.

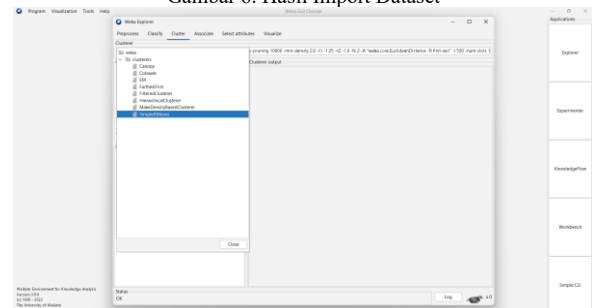


Gambar 5. Proses Import Dataset

Langkah pertama dalam menggunakan Weka yaitu melakukan import data yang akan dihitung. Sebelum memulai import data, file dari data tersebut perlu disimpan dalam format csv (Comma delimited) agar dapat diproses oleh Weka.

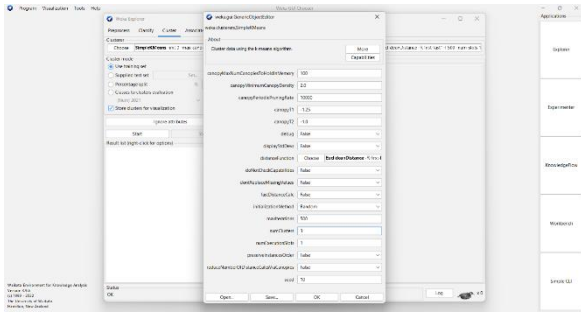


Gambar 6. Hasil Import Dataset



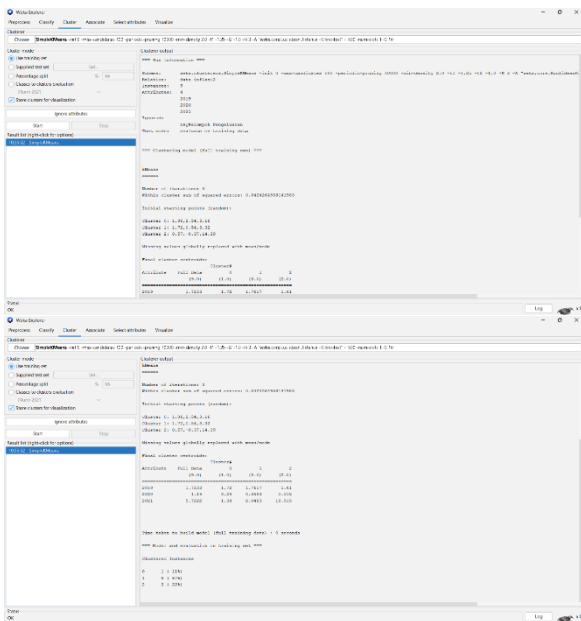
Gambar 7. Pemilihan Metode

Metode yang digunakan adalah cluster dengan cara pengklasterannya yaitu SimpleKMeans.



Gambar 8. Penentuan Jumlah Cluster

Di sini akan dilakukan penentuan jumlah cluster yang diinginkan. Karena dari data yang ada akan dikelompokkan menjadi tiga kelompok yaitu rendah, sedang dan tinggi maka untuk numClusters atau jumlah cluster yaitu sebanyak 3 cluster.



Gambar 9. Hasil Clustering

Dari proses klusterisasi pada Weka didapatkan hasil seperti pada Gambar 10 di atas ini. Dengan pusat awal centroid yang ditentukan secara random diketahui bahwa iterasi terjadi sebanyak 3 kali dengan menghasilkan nilai SSE (Sum of Squared Errors) yaitu 0,85172. Dari data yang telah diproses ini, didapatkan 3 kelompok dengan hasil sebagai berikut:

1. Untuk cluster 0 pada tahun 2019 menghasilkan pusat cluster akhir = 1.83, tahun 2020 = 8.29, tahun 2021 = 1.38 sebanyak 1 kelompok pengeluaran dari 9 kelompok
2. Cluster 1 pada tahun 2019 menghasilkan pusat cluster akhir = 1.90, tahun 2020 = 0.39, dan tahun 2021 = 2.04 sebanyak 6 kelompok pengeluaran dari 9 kelompok.
3. Cluster 2 pada tahun 2019 menghasilkan pusat cluster akhir = 1.61, tahun 2020 = 0.015, dan tahun 2021 = 18.92 sebanyak 1 kelompok pengeluaran dari 9 kelompok.

data inflasi.aff						
Relation: data inflasi2_clustered						
No.	1: Instance Number	2: I _j -Kelompok Pengeluaran	3: 2019	4: 2020	5: 2021	6: Cluster
	Nominal	Nominal	Numeric	Numeric	Numeric	Nominal
1	0.0	Makanan, Minuman dan Tembakau	1.86	2.54	3.16	cluster1
2	1.0	Pakaian dan Alas Kaki	3.27	0.28	1.03	cluster1
3	2.0	Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya	0.57	-0.37	14.39	cluster2
4	3.0	Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga	1.72	0.54	3.32	cluster1
5	4.0	Kesehatan	2.65	0.4	23.44	cluster2
6	5.0	Transportasi, Informasi, Komunikasi, dan Jasa Keuangan	-2.3	-0.48	-3.17	cluster1
7	6.0	Pendidikan, Rekreasi, Olahraga dan Budaya	4.3	-1.23	5.66	cluster1
8	7.0	Penyediaan Makanan dan Minuman/Restoran	1.72	0.74	2.29	cluster1
9	8.0	Perawatan Pribadi dan Jasa Lainnya	1.72	8.29	1.38	cluster0

Gambar 10. Hasil Akhir

Maka didapatkan hasil bahwa kelompok pengeluaran ‘Perawatan Pribadi dan Jasa Lainnya’ termasuk ke dalam cluster 0 yang mana artinya untuk jenis pengeluaran ini memiliki tingkat laju inflasi sedang. Sementara itu pada kelompok pengeluaran lain seperti ‘Makanan, Minuman dan Tembakau’, ‘Pakaian dan Alas Kaki’, ‘Perlengkapan, Peralatan dan Pemeliharaan Rutin Rumah Tangga’, ‘Transportasi, Informasi, Komunikasi dan Jasa Keuangan’, ‘Pendidikan, Rekreasi, Olahraga dan Budaya’, ‘Penyediaan Makanan dan Minuman/Restoran’ termasuk ke dalam cluster 1 dengan kelompok yang tingkat laju inflasinya rendah. Kemudian untuk cluster 2 sebagai kelompok dengan tingkat laju inflasi yang tinggi terdiri dari kelompok pengeluarannya yaitu ‘Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya’ dan ‘Kesehatan’.

IV. KESIMPULAN

Dari hasil penelitian di atas dapat ditarik kesimpulan telah didapatkan tiga kelompok tingkat kelajuan inflasi yaitu C0 = sedang, C1 = lambat, dan C2 = tinggi. Dari pengelompokan tersebut dapat diketahui laju inflasi di kota Sampit pada 3 tahun terakhir yaitu 2019-2021, pada tahun tersebut merupakan awal covid-19 diumumkan sebagai pandemi dan hal itu menyebabkan inflasi di kota Sampit meningkat khususnya pada bagian ‘Perumahan, Air, Listrik, Gas dan Bahan Bakar Lainnya’ dan ‘Kesehatan’.

V. REFERENSI

- [1] Z. Nabila, A. Rahman Isnain, and Z. Abidin, “ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS,” *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, 2021.
- [2] R. Amalin Lutfi and M. Panorama, “DAMPAK COVID-19 TERHADAP TINGKAT INFLASI (KELOMPOK PENGELUARAN MAKANAN, MINUMAN, TENAGA KERJA DAN KESEHATAN PERIODE 2010-2020),” *J. Ekon. dan Bisnis Islam*, vol. 1, 2021.
- [3] Y. Prayoga, H. S. Tambunan, and I. Parlina, “Penerapan Clustering Pada Laju Inflasi Kota Di Indonesia Dengan Algoritma K-Means,” *BRAHMANA J. Penerapan Kecerdasan Buatan*, vol. 1, no. 1, 2019, doi: 10.30645/brahmana.v1i1.4.
- [4] K. Sari and N. A. Sudibyo, “Implementasi Metode K-means untuk Mengelompokkan

- Tingkat Inflasi di Indonesia,” *E-Journal Biostatistics| Dep.*, 2020.
- [5] D. P. Sari, “IMPLEMENTASI ALGORITMA K-MEANS DALAM MENENTUKAN TINGKAT PENYEBARAN PANDEMI COVID-19 DI SUMATERA BARAT,” *Comput. Based Inf. Syst. J.*, vol. 9, no. 1, 2021, doi: 10.33884/cbis.v9i1.3646.
- [6] P. A. Ariawan, “Optimasi Pengelompokan Data Pada Metode K-means dengan Analisis Outlier,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 5, no. 2, 2019, doi: 10.25077/teknosi.v5i2.2019.88-95.
- [7] M. F. Rahmi, P. S. Prasetyo, R. Nurhabibah, R. Perdana, and W. O. Z. Madjida, “PENGELOMPOKKAN PROVINSI DI INDONESIA BERDASARKAN JUMLAH KASUS COVID-19 DAN FASILITAS KESEHATAN,” *J. Apl. Stat. Komputasi Stat.*, vol. 13, no. 1, 2021, doi: 10.34123/jurnalasks.v13i1.274.