

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KLASTERISASI PESERTA BPJS BERDASARKAN REKAM MEDIS MENGGUNAKAN ALGORITMA K-MEANS TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

ANA FITRI KHAIRANI

NIM. 11950125009



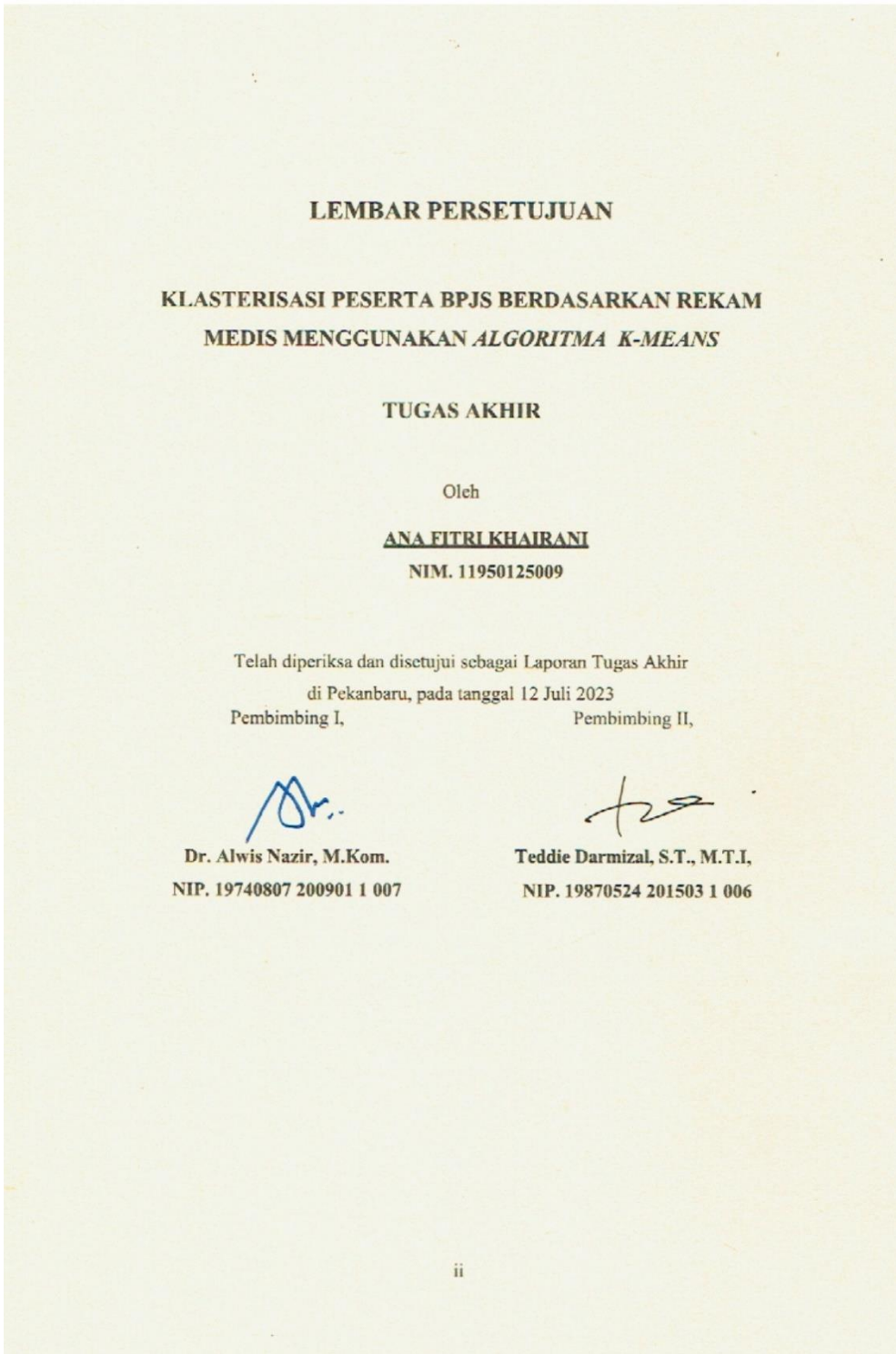
UIN SUSKA RIAU

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU**

2023

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PENGESAHAN

**KLASTERISASI PESERTA BPJS BERDASARKAN REKAM
MEDIS MENGGUNAKAN ALGORITMA K-MEANS**

Oleh

ANA FITRI KHAIRANI

NIM. 11950125009

Telah dipertahankan di depan sidang dewan penguji sebagai salah
satu syarat untuk memperoleh gelar Sarjana Teknik pada
Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 12 Juli 2023

Mengesahkan,

Ketua Jurusan,




Dr. Hartono, M.Pd.
NIP. 19640301 199203 1 003

Iwan Iskandar, M.T
NIP. 19821216 201503 1 003

DEWAN PENGUJI

Ketua : Muhammad Affandes, S.T., M.T
Pembimbing I : Dr. Alwis Nazir., M.Kom.
Pembimbing II : Teddie Darnizal, S.T., M.T.I
Penguji I : Yelfi Vitriani, S.Kom., MMSI
Penguji II : Yusra, MT.





LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

© Hak Cipta dan Hak Kekayaan Intelektual UIN Suska Riau

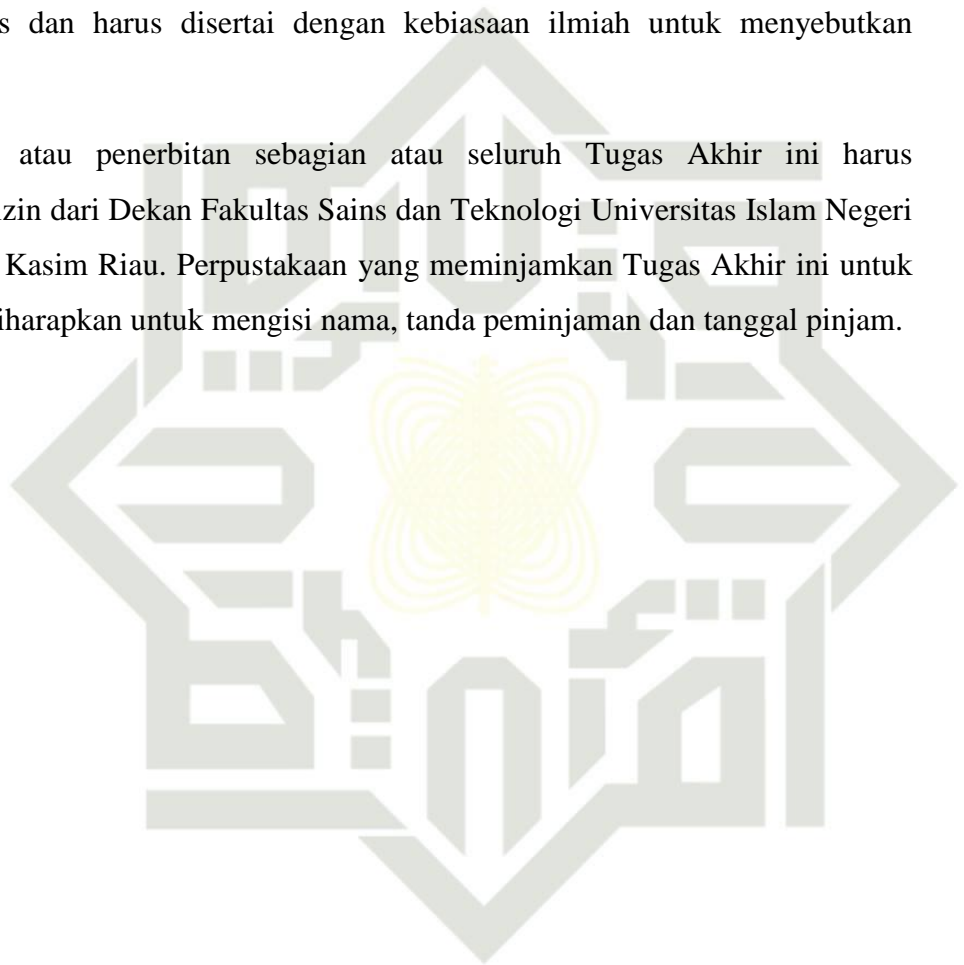
State Islamic University of Sultan Syarif Kasim Riau

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



UIN SUSKA RIAU



LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 12 Juli 2023

Yang membuat pernyataan,

ANA FITRI KHAIRANI

NIM. 11950125009

UIN SUSKA RIAU

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Alhamdulillahirabbil 'alamin

Dengan mengucapkan syukur pada Allah subhanahu wa ta'ala, dan shalawat serta salam kepada Nabi kita Muhammad shallallahu Alaihi wasallam, telah saya selesaikan Tugas Akhir ini...

Saya persembahkan Tugas Akhir Saya Ini Untuk Kedua Orang Tua, Keluarga, dan Teman-Teman...

Semoga Tugas Akhir in bermanfaat bagi pembacanya.

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini :

Nama : Ana Fitri Khairani
NIM : 11950125009
Tempat/Tgl.Lahir : Dumai, 08 Mei 2001
Fakultas : Sains dan Teknologi
Prodi : Teknik Informatika
Judul Skripsi : Klasterisasi Peserta BPJS Berdasarkan Rekam Medis
Menggunakan Algoritma K-Means

Menyatakan dengan sebenar-benarnya bahwa:

1. Penulisan Skripsi dengan judul sebagaimana tersebut diatas adalah hasil pemikiran penelitian saya sendiri.
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya.
3. Oleh karena itu, Skripsi saya ini, saya nyatakan bebas dari plagiat.
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan Skripsi saya tersebut, maka saya bersedia menerima sanksi sesuai perundang-undangan.

Demikian Surat Pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksa pihak manapun juga.

Pekanbaru, 12 Juli 2023

Yang membuat pernyataan,



ANA FITRI KHAIRANI

NIM. 11950125009

UIN SUSKA RIAU



Klasterisasi Peserta BPJS Berdasarkan Rekam Medis Menggunakan Algoritma K-Means

Ana Fitri Khairani*, Alwis Nazir, Teddie Darmizal, Yelfi Vitriani, Yusra

Fakultas Sains dan Teknologi, Prodi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim, Pekanbaru, Indonesia

Email: ^{1,*}11950125009@students.uin-suska.ac.id, ²alwis.nazir@uin-suska.ac.id, ³teddie.darmizal@uin-suska.ac.id, ⁴yelfi.vitriani@uin-suska, ⁵yusra@uin-suska.ac.id,

Email Penulis Korespondensi: 11950125009@students.uin-suska.ac.id

Abstrak— Riwayat pemeriksaan kesehatan pasien dengan BPJS (Badan Penyelenggara Jaminan Sosial) di RSUD Kota Dumai akan disimpan dalam bentuk rekam medis. Rekam medis merupakan berkas yang berisikan informasi jati diri pasien baik itu berisikan hasil kontrol, perawatan dan jasa lainnya. Tujuan penelitian ini untuk membantu pihak RSUD Dumai dalam melakukan pengelompokan peserta BPJS berdasarkan rekam medis untuk mengetahui penyakit, jenis kelamin dan kelompok BPJS yang paling dominan pada peserta BPJS Kota Dumai. Hasil pengelompokan penyakit kemudian diolah menggunakan teknik data mining. Pada penelitian ini difokuskan pada peserta BPJS PBI rawat inap data rekam medis pasien periode Januari hingga Desember 2022 yang kemudian dilakukan dengan K-Means Clustering sebagai algoritma pemrosesan data, terbagi 3 cluster. Cluster 0 didominasi jenis penyakit dengan kode O342 (Maternal care due to uterine scar from previous surgery), jenis kelamin perempuan dan rentang umur 21-40 tahun. Untuk cluster 1 didominasi jenis penyakit dengan kode E119 (Diabetes mellitus without complications), jenis kelamin perempuan dan rentang umur 41-60 tahun dan cluster 2 didominasi jenis penyakit dengan kode J180 (Bronchopneumonia, unspecified organism), jenis kelamin perempuan dan rentang umur 41-60 tahun

Kata Kunci: Rekam Medis; BPJS; Data Mining; Clustering; K-Means

Abstract—The history of the patient's medical examination with the BPJS (Social Security Administration Agency) at the Dumai City Hospital will be stored in the form of a medical record. The medical record is a file that contains patient identification information whether it contains the results of control, treatment and other services. The purpose of this research is to assist the Dumai Hospital in grouping BPJS participants based on medical records to find out the disease, gender and BPJS group that is most dominant in Dumai City BPJS participants. The results of grouping the disease will then be processed using data mining techniques. In this study, the focus was on BPJS PBI participants inpatient medical record data from January to December 2022, which were then processed using the K-Means Clustering algorithm with 3 clusters. Cluster 0 is dominated by disease types with code O342 (Maternal care due to uterine scar from previous surgery), female sex and age range 21-40 years. Cluster 1 is dominated by types of disease with code E119 (Diabetes mellitus without complications), female sex and age range 41-60 years and cluster 2 is dominated by disease types with code J180 (Bronchopneumonia, unspecified organism), female sex and age range 41-60 years.

Keywords: Medical Records; BPJS; Data Mining; Clustering; K-Means

1. PENDAHULUAN

Badan Penyelenggara Jaminan Sosial (BPJS) adalah badan hukum yang didirikan untuk menyelenggarakan program jaminan sosial[1][2]. Salah satu kategori dari BPJS adalah BPJS Kesehatan. Dengan BPJS Kesehatan, masyarakat akan lebih mudah memperoleh pelayanan dan fasilitas kesehatan di pusat pelayanan kesehatan yang sudah ditetapkan oleh pemerintah sebagai pemberi layanan kesehatan berbasis BPJS[3].

Rekam medis menurut Keputusan Menteri Kesehatan No. 269/ MENKES/ PER/ III/ 2008 adalah berkas yang berisikan catatan dan dokumentasi tentang identitas pasien, pemeriksaan, pengobatan, tindakan dan pelayanan lain yang telah diberikan kepada pasien. Data yang tercantum direkam medis harus memuat seluruh data pasien, baik data pribadi hingga data yang dapat menunjukkan diagnosis kesehatan pasien sehingga data ini dapat merepresentasikan kondisi pasien dengan akurat. Salah satu informasi pasien yang termuat didalam rekam medis adalah pengodean diagnosis dimana dilakukan pengodean pada tiap komponen data dengan menggunakan huruf, angka atau kombinasi keduanya sehingga memunculkan suatu kode yang merepresentasikan tiap komponen data[4]. Pengodean ini dilakukan oleh petugas yang disebut koder. Tugas dari koder adalah melakukan kodifikasi diagnosis ditulis oleh dokter yang merawat pasien sesuai dengan ICD-10 (International Statistical Classification of Diseases and Related Health Problem Tenth Revisions)[5][6].

ICD merupakan suatu sistem klasifikasi hasil kerjasama antara World Health Organization (WHO) bersama 10 pusat internasional(ICD-, 2000). Menurut WHO, ICD merupakan standar informasi kesehatan global untuk menentukan statistik mortalitas dan morbiditas. Pengodean dilakukan setelah diagnosis utama didapatkan, kemudian diagnosis tersebut dikode berdasarkan aturan yang semestinya. ICD-10 memiliki kode sebanyak 70.00 lebih. Selain itu, ICD-10 memiliki 3 volume (3 buku) [5]. ICD digunakan untuk menterjemahkan diagnosis penyakit dan masalah kesehatan dari kata-kata menjadi kode alfanumerik[7][8].

Rumah Sakit Umum Daerah (RSUD) Kota Dumai Provinsi Riau adalah salah satu lembaga kesehatan yang mayoritas pasiennya peserta BPJS Kesehatan. Setiap pasien yang berobat akan tersimpan semua identitas dan riwayat pelayanan kedalam bentuk rekam medis. Rekam Medis berisi informasi jati diri pasien, pemeriksaan, pengobatan, dan pelayanan yang telah diberikan kepada pasien[9][10].

Hak cipta dilindungi Undang-Undang. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber. Dilarang mengutip sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

RSUD Kota Dumai belum maksimal dalam upaya persiapan obat dan fasilitas yang diperuntukkan kepada pasien dikarenakan banyaknya pasien setiap harinya. Untuk memaksimalkan upaya tersebut, penelitian ini akan melakukan pengelompokan rekam medis peserta BPJS di RSUD Dumai yang difokuskan pada peserta BPJS PBI rawat inap sehingga pihak rumah sakit dapat mengetahui jenis penyakit, jenis kelamin, dan umur yang paling dominan pada peserta BPJS Kota Dumai sehingga pihak rumah sakit dapat melakukan persiapan maksimal terhadap stok obat maupun fasilitas lainnya untuk pasien dan pihak Dinas Kesehatan Kota Dumai dapat melakukan sosialisasi untuk pencegahan penyakit. Hasil dari pengelompokan rekam medis tersebut selanjutnya akan diolah menggunakan teknik data mining.

Data Mining atau Knowledge Discovery in Database (KDD) adalah suatu proses yang dilakukan dengan tujuan untuk menemukan keterkaitan atau hubungan yang memiliki makna dengan memperhatikan data yang besar, yang mana data ini akan disimpan dalam suatu penyimpanan dan akan dikelola secara statistic technic dan teknik pengelompokan titik pada data menjadi beberapa kelompok dan menempatkan titik data tersebut kedalam satu kelompok yang sama atau lebih mirip satu sama lain daripada di kelompok yang berbeda, dengan memanfaatkan informasi yang tersedia dengan titik data[13][14]. Penelitian ini menggunakan Algoritma Clustering K-Means. K-Means merupakan teknik clustering data non hierarki yang akan membagi data menjadi beberapa jenis cluster atau kelompok sehingga menghasilkan sekelompok data dengan karakteristik yang sama, sedangkan untuk data dengan karakteristik yang berbeda akan dikelompokkan ke dalam kelompok lainnya[15]. Pada penelitian sebelumnya dengan metode clustering k-means yaitu penelitian yang dilakukan oleh Jeri Wandana (2020) didapatkan hasil jenis penyakit dengan kode A09.9 (Diare/Disentri) mendominasi pada cluster 0, jenis penyakit yang lebih beragam pada cluster 1, dan jenis penyakit dengan kode K30 (Dyspepsia) lebih didominasi pada cluster 2[16]. Pada penelitian tersebut menggunakan data dari salah satu rumah sakit swasta di Pekanbaru dengan atribut sebanyak 7 yang terdiri dari kode penyakit, umur, jenis kelamin, jenis BPJS, Referred, kelas BPJS dan faskes, sedangkan penelitian ini menggunakan data sebanyak 5750 yang di dapatkan dari rumah sakit umum daerah Kota Dumai dengan atribut yang digunakan pada penelitian ini terdiri dari umur, jenis kelamin dan ICD-10 karena penelitian ini difokuskan pada pasien BPJS PBI rawat inap.

Penelitian ini dilakukan dengan tujuan untuk menerapkan clustering K-means untuk mendapatkan hasil yang baik sehingga dapat membantu pihak RSUD Dumai dalam memaksimalkan persediaan obat dan layanan lainnya.

2. METODOLOGI PENELITIAN

2.1. Clustering

Clustering adalah suatu teknik pengelompokan titik pada data menjadi beberapa kelompok dan menempatkan titik data tersebut kedalam satu kelompok yang sama atau lebih mirip satu sama lain daripada di kelompok yang berbeda, dengan memanfaatkan informasi yang tersedia dengan titik data[13][14]. Clustering adalah metode untuk mengategorikan atau pengelompokan sekelompok objek sesuai dengan atribut yang sama atau karakteristik ke dalam sejumlah groups[17].

2.2. K-Means

K-Means merupakan teknik clustering data non hierarki yang akan membagi data menjadi beberapa jenis cluster atau kelompok sehingga menghasilkan sekelompok data dengan karakteristik yang sama, sedangkan untuk data dengan karakteristik yang berbeda akan dikelompokkan ke dalam kelompok lainnya[15]. Algoritma K-means ini digunakan karena kesederhanaan algoritma tersebut dapat diaplikasikan pada berbagai bidang. Algoritma ini juga diakui sebagai salah satu algoritma teratas oleh IEEE[18][19].

2.3. Rapidminer

RapidMiner adalah platform perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Ini menyediakan beragam alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. RapidMiner dirancang untuk mudah digunakan dan memungkinkan pengguna untuk dengan mudah membangun dan menguji berbagai model, bahkan tanpa pengalaman pemrograman[20]

2.4. Data Penelitian

Data penelitian yang digunakan adalah data peserta BPJS PBI rawat inap periode januari-desember 2022 yang diperoleh dari Rumah Sakit Umum Daerah Kota Dumai. Tabel 1 merupakan data testing yang akan digunakan pada penelitian ini.

Tabel 1. Data penelitian

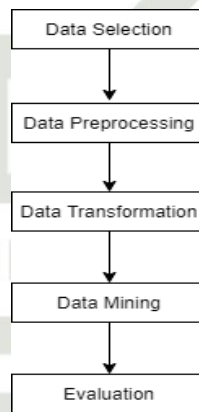
Pasien	Umur	Jns.Kelamin	Cara Masuk	Cara Bayar	Icd 10
1	31 Tahun	pr	r.inap	BPJS PBI	O210



Pasien	Umur	Jns.Kelamin	Cara Masuk	Cara Bayar	Icd 10
2	63 tahun	pr	r.inap	BPJS PBI	N185
3	50 tahun	pr	r.inap	BPJS PBI	J81
4	64 tahun	pr	r.inap	BPJS PBI	B909
5	29 tahun	pr	r.inap	BPJS PBI	O210
6	66 tahun	lk	r.inap	BPJS PBI	R630
7	38 tahun	Pr	r.inap	BPJS PBI	D179
8	14 tahun	lk	r.inap	BPJS PBI	D561
9	19 tahun	Pr	r.inap	BPJS PBI	D561
10	32 tahun	lk	r.inap	BPJS PBI	Q210
...
30	42 tahun	Pr	r.inap	BPJS PBI	E112

Tahapan Penelitian

Tahapan penelitian berisi tentang urutan dari awal hingga akhir penelitian. berikut ini adalah tahapan penelitian.



Gambar 1. Tahapan penelitian

a. Data Selection

Sebelum melakukan proses data mining algoritma k-means data akan melalui proses data selection yang merupakan langkah pertama pada KDD (Knowledge Discovery Database). Tahapan ini dilakukan pemilihan atau peneseleksian dari sekumpulan data yang relevan terhadap analisis untuk di terima dari koleksi data yang ada.

b. Data Preprocessing/Cleaning

Setelah dilakukan pemilihan data pada langkah data selection selanjutnya data akan melalui proses data preprocessing. Tahapan ini dilakukan pembersihan data seperti menghapus data duplikat dan penanganan informasi yang hilang (missing value)

c. Data Transformation

Setelah melewati proses data preprocessing selanjutnya dilakukan data transformation untuk mengubah data kedalam tipe yang sesuai dengan data mining. Tahapan ini bertujuan untuk menelaraskan data antara 0 dan 1 sehingga tidak terjadi outlier.

d. Data Mining

Pada proses data mining ini dilakukan pengolahan dataset yang didapatkan dari RSUD Kota Dumai diolah dengan menggunakan algoritma K-Means.

e. Evaluation

Tahapan ini akan dilakukan pengujian dengan rapidminer. Output yang dihasilkan dari proses data mining berupa gambar maupun tabel dan akan dievaluasi menggunakan teknik evaluasi Davies Bouldin Index (DBI).

3. HASIL DAN PEMBAHASAN

3.1 Data Selection

Tabel 2 merupakan contoh data selection rekam medis pasien BPJS PBI rawat inap RSUD Kota Dumai tahun 2022.

Tabel 2. Contoh Dataset

Pasien	Umur	Jns.Kelamin	Icd 10
1	31 Tahun	pr	O210
2	63 tahun	pr	N185

Hak Cipta Dilindungi Undang-Undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pasien	Umur	Jns.Kelamin	Icd 10
3	50 tahun	pr	J81
4	64 tahun	pr	B909
5	29 tahun	pr	O210
6	-	lk	-
7	38 tahun	Pr	D179
8	14 tahun	lk	D561
9	19 tahun	Pr	D561
10	32 tahun	lk	Q210
...
30	42 tahun	Pr	E112

3.2 Data Preprocessing

Setelah dilakukan data selection selanjutnya dilakukan penanganan informasi yang hilang (missing value) dan duplikasi data, seperti tabel berikut.

Tabel 3. dataset setelah dilakukan penanganan missing value

Pasien	Umur	Jns.Kelamin	Icd 10
1	31 Tahun	pr	O210
2	63 tahun	pr	N185
3	50 tahun	pr	J81
4	64 tahun	pr	B909
5	29 tahun	pr	O210
6	66 tahun	lk	R630
7	38 tahun	Pr	D179
8	14 tahun	lk	D561
9	19 tahun	Pr	D561
10	32 tahun	lk	Q210
...
30	42 tahun	Pr	E112

3.3 Data Transformation

Pada tahapan ini, data yang awalnya berjenis non numerik akan diubah menjadi data numerik, seperti tabel berikut.

Tabel 4. Transformasi data rekam medis

Pasien	Umur	Jns.Kelamin	Icd 10
1	2	2	472
2	4	2	419
3	3	2	268
4	4	2	52
5	2	2	472
6	4	1	602
7	2	2	83
8	1	1	93
9	1	2	93
10	2	1	728
...
30	3	2	113

3.4 Data Mining

Pada tahap ini dilakukan perhitungan clustering data mining K-Means. Untuk mencari jarak tiap data ke pusat centroid digunakan jarak Euclidean, C0 diambil dari data pasien 2, C1 dari data pasien 19, dan C2 dari data pasien 28, seperti pada tabel berikut.

Rumus Euclidean Distance

$$d(x,y) = \sqrt{\sum_{i=0}^n (x_i - y_i)^2} \tag{1}$$

Keterangan:

- x = data
- y = centroid
- n = jumlah data



Tabel 5. Centroid awal

Centroid Awal			
Pasien	Umur	Jenis Kelamin	ICD-10
C0	4	2	419
C1	3	2	114
C2	1	2	289

- Jarak atribut ke centroid 0 diketahui melalui perhitungan dengan data urutan 1.
 $C0 = \sqrt{(2 - 4)^2 + (2 - 2)^2 + (472 - 419)^2} = 53,037$
- Jarak atribut ke centroid 1 didapatkan dari perhitungan dengan data urutan 1
 $C1 = \sqrt{(2 - 3)^2 + (2 - 2)^2 + (472 - 114)^2} = 358,001$
- Jarak atribut ke centroid 2 didapatkan dari perhitungan dengan data urutan 1
 $C2 = \sqrt{(2 - 1)^2 + (2 - 2)^2 + (472 - 289)^2} = 183,002$
- Lakukan perhitungan hingga data urutan 30
- Selanjutnya, pengelompokan data dilakukan dengan melihat jarak terdekat dengan pusat cluster seperti pada tabel berikut

Tabel 6. Hasil Iterasi 1

C0	C1	C2	Jarak terdekat	Cluster
53,037	358,001	183,003	53,037	0
0	305,002	130,034	0	0
151,003	154	21,09	21,09	2
367	62,008	237,019	62,008	1
53,037	358,001	183,003	53,037	0
183,003	488,002	313,016	183,003	0
336,006	31,016	206,002	31,016	1
326,015	21,118	196,003	21,118	1
196,023	21,095	196	21,095	1
309,008	614,002	439,003	309,008	0
...
306,002	1	176,011	1	1

Berdasarkan dari hasil iterasi 1, maka diperoleh C0 berjumlah 6 anggota, C1 berjumlah 17 anggota, serta C2 berjumlah 7 anggota. Penentuan centroid baru diperoleh dari nilai rata-rata cluster yang bersangkutan, seperti pada tabel berikut.

Tabel 7. Centroid Baru

Centroid Awal			
Pasien	Umur	Jenis Kelamin	ICD-10
C0	2,5	9,167	547,3
C1	1,941	1,5294	84,23
C2	2,428	1,7142	256,85

Hitunglah iterasi kedua dengan menggunakan hasil hasil centroid baru

- Jarak atribut ke centroid 0 diketahui melalui perhitungan dengan mengambil dari 1.
 $C0 = \sqrt{(2 - 2,5)^2 + (2 - 9,167)^2 + (472 - 547,3)^2} = 75,641$
- Jarak atribut ke centroid 1 didapatkan dari perhitungan dengan data urutan 1
 $C1 = \sqrt{(2 - 1,941)^2 + (2 - 1,5294)^2 + (472 - 84,23)^2} = 387,770$
- Jarak atribut ke centroid 2 didapatkan dari perhitungan dengan data urutan 1
 $C2 = \sqrt{(2 - 2,2428)^2 + (2 - 1,7142)^2 + (472 - 84,23)^2} = 215,150$
- Lakukan perhitungan hingga data urutan 30
- Selanjutnya, pengelompokan dilakukan dengan melihat jarak terdekat dengan pusat cluster, seperti pada tabel berikut.

Tabel 8. Hasil Iterasi 2

C0	C1	C2	Jarak terdekat	cluster
75,641	387,770	215,150	75,641	0
128,059	334,777	162,157	128,059	0
279,392	183,774	11,168	11,168	2
495,354	32,29	204,856	32,29	1
75,64	387,77	215,151	75,64	0

Hak Cipta Dilindungi Undang-Undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

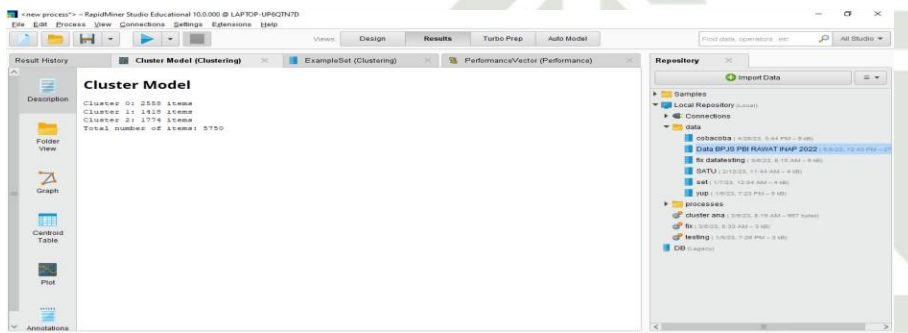


C0	C1	C2	Jarak terpendek	cluster
55,32	517,77	345,154	55,32	0
464,356	1,318	173,85	1,318	1
454,376	8,836	163,85	8,836	1
164,014	8,832	163,856	8,832	1
180,885	643,77	471,151	180,885	0
...
434,359	28,79	143,85	28,79	1

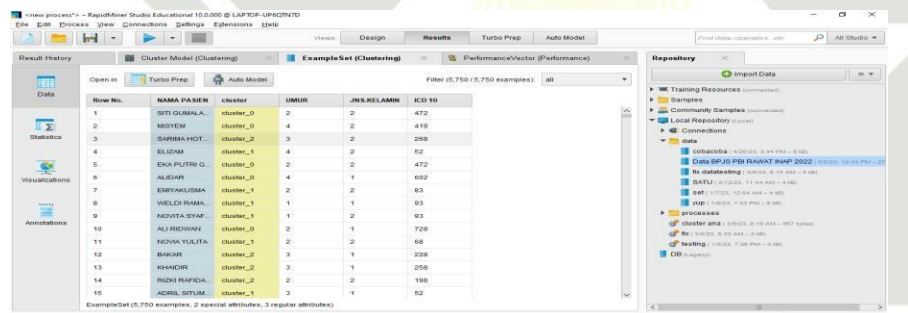
Dari hasil yang didapatkan pada iterasi 2, dapat dilihat bahwa iterasi 1 dan iterasi 2 memiliki hasil yang sama, yaitu C0 berjumlah 6 anggota, C1 berjumlah 17 anggota, serta C2 berjumlah 7 anggota maka proses berhenti.

3. Pengujian

Langkah selanjutnya setelah mendapatkan hasil akhir dilakukan pengujian dengan menggunakan Rapidminer 10.0, seperti pada gambar 2.



Gambar 2. Hasil Cluster model pada Rapidminer.

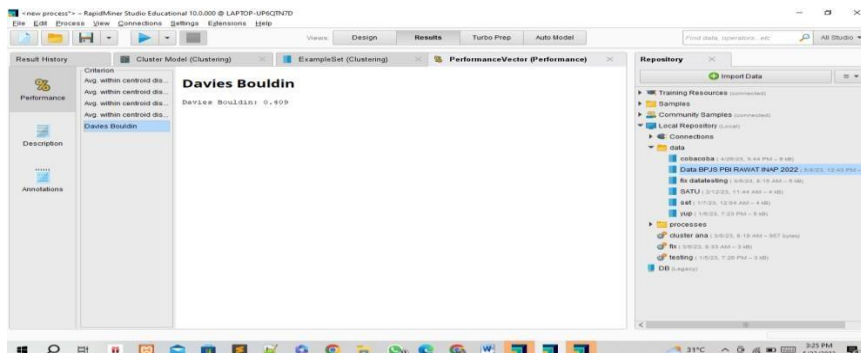


Gambar 3. Hasil Cluster model Rapidminer

Pada gambar 3 dapat dilihat hasil pengujian menggunakan Rapidminer 10.0 diketahui bahwa data yang digunakan adalah benar. Karena perhitungan manual dan rapidminer menghasilkan hasil yang sama.

3.6 Evaluation

Pada tahap evaluasi ini dilakukan uji kelayakan untuk mengetahui apakah algoritma k-means adalah algoritma yang pada penelitian ini. Evaluasi data yang digunakan pada penelitian ini adalah Index Davies Bouldin, berikut ini hasil pengujian menggunakan rapidminer 10.0



Gambar 4. Nilai Davies Bouldin Index (DBI) Pada Rapidminer

Hak Cipta Diindungi Undang-Undang
 1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan satu masalah.
 b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Diarangi mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Dari gambar 4 dapat dilihat bahwa hasil evaluasi membentuk tiga cluster dengan nilai DBI 0,409. K=3 merupakan nilai yang paling mendekati 0

4. KESIMPULAN

Setelah dilakukan proses clustering data dengan metode K-Means didapati hasil terdapat 3 cluster data rekam medis pasien pengguna BPJS PBI rawat inap dengan karakteristik yang berbeda. Pada cluster 0 dengan anggota sebanyak 2558 didominasi penyebaran jenis penyakit kode O342 (Maternal care due to uterine scar from previous surgery), jenis kelamin perempuan dan rentang umur 21-40 tahun, cluster 1 dengan anggota sebanyak 1418 didominasi penyebaran jenis penyakit dengan kode E119 (Diabetes mellitus without complications), jenis kelamin perempuan dan rentang umur 41-60 tahun, dan cluster 2 dengan anggota sebanyak 1774 didominasi penyebaran jenis penyakit dengan kode J180 (Bronchopneumonia, unspecified organism), jenis kelamin perempuan dan rentang umur 41-60 tahun.

REFERENCES

- [1] UU RI Nomor 24, UU no 24 tahun 2011 tentang BPJS. 2011, pp. 1–13.
- [2] R. Kusuma, A. Basniwati, L. Guna Nugraha, and S. Hariati, "Hak Peserta Badan Penyelenggara Jaminan Sosial Kesehatan," *Palar | Pakuan Law Rev.*, vol. 7, no. 2, pp. 194–205, 2021, doi: 10.33751/palar.v7i2.3242.
- [3] L. Masyfufah and I. Artikel, "Clustering Guna Menunjang Program Jaminan Balong Bendo Sidoarjo Clustering Bpjs ' S Inpatients By Using K -Means Clustering Method To Support the National Health Insurance Program At Anwar Medika Hospital Balong Bendo Sidoarjo," *J. Wiyata*, vol. 8, no. 1, pp. 8–22, 2021.
- [4] V. A. Isnaini, "Strategi Perbaikan Ketidaktepatan Kodefikasi Berkas Rekam Medis Pasien Rawat Jalan Berdasarkan ICD-10 dengan PDCA di Puskesmas Sukodono Lumajang," *Pros. Semin. Rekam Med. dan Inf. Kesehat.*, vol. 1, no. 1, pp. 7–8, 2019, [Online]. Available: <https://publikasi.polije.ac.id/index.php/prosidingrmd/article/view/1536>.
- [5] Y. Mardi, N. Mayasari, and V. G. Putri, "Sosialisasi Literasi Digital dalam Pengodean Penyakit Berdasarkan ICD-10 Versi 2016," *J. Community Engagem. Heal.*, vol. 3, no. 2, pp. 275–283, 2020, [Online]. Available: <https://jceh.org/index.php/JCEH/article/view/80>.
- [6] Kementerian Kesehatan Republik Indonesia, "Peraturan Menteri Kesehatan Republik Indonesia Nomor 27 Tahun 2014 tentang Petunjuk Teknis Sistem Indonesia Case Base Groups (INA-CBGs)," no. 795, pp. 1–49, 2014.
- [7] F. Teknologi and D. A. N. Informatika, "Rawat Jalan Berbasis Web Pada Poliklinik Polrestabes Surabaya," 2019.
- [8] R. Oashttamadea, "Pelatihan Pengodean Tindakan Medis Berdasarkan ICD 9-CM pada Kasus Obstetri dan Ginekologi di RSIA Sayang Ibu Batusangkar Abstrak," vol. 1, pp. 1–5, 2021.
- [9] Y. S. Sy, "Klasterisasi Pasien Rawat Inap Peserta BPJS Berdasarkan Jenis Penyakit Menggunakan Algoritma K-Means," *J. Sistim Inf. dan Teknol.*, vol. 5, pp. 33–37, 2022, doi: 10.37034/jsisfotek.v5i2.162.
- [10] PERMENKES RI No 269/MENKES/PER/III/2008, "permenkes ri 269/MENKES/PER/III/2008," *Permenkes Ri No 269/Menkes/Per/Iii/2008*, vol. 2008, p. 7, 2008.
- [11] J. Informatika, D. Reayasa, K. Jakakom, K. Telkomsel, R. Cell, and S. Candra, "Penerapan Data Mining Algoritma Naive Bayes Clasifier Untuk Mengetahui Minat Beli Pelanggan Terhadap Jurnal Informatika Dan Reayasa Komputer (JAKAKOM)," vol. 1, no. April, pp. 149–154, 2022.
- [12] Lea and Jenie, "Penerapan Algoritma Apriori Untuk Menentukan Impor Barang Pada PT. Pafa Mandiri Sakti," *J. Process.*, vol. 15, no. 1, pp. 54–61, 2020, doi: 10.33998/processor.2020.15.1.767.
- [13] V. Herlinda and D. Darwis, "Analisis Clustering Untuk Recredesialing Fasilitas Kesehatan Menggunakan Metode Fuzzy C-Means," *Darwis, Dartono*, vol. 2, no. 2, pp. 94–99, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/JTSI>.
- [14] A. A. Aldino, D. Darwis, A. T. Prastowo, and C. Sujana, "Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency," *J. Phys. Conf. Ser.*, vol. 1751, no. 1, 2021, doi: 10.1088/1742-6596/1751/1/012038.
- [15] K. H. K. Al Atros, A. Robi Padri, O. Nurdiawan, A. Faqih, and S. Anwar, "Model Klasifikasi Analisis Kepuasan Pengguna Perpustakaan Online Menggunakan K-Means dan Decission Tree," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, pp. 323–329, 2021, doi: 10.30865/jurikom.v8i6.3680.
- [16] J. Wandana, S. Defit, and S. Sumijan, "Klasterisasi Data Rekam Medis Pasien Pengguna Layanan BPJS Kesehatan Menggunakan Metode K-Means," *J. Inf. dan Teknol.*, vol. 2, pp. 4–9, 2020, doi: 10.37034/jidt.v2i4.73.
- [17] B. Harahap, "Penerapan Algoritma K-Means Untuk Menentukan Bahan Bangunan Laris (Studi Kasus Pada UD. Toko Bangunan YD Indarung)," *Reg. Dev. Ind. Heal. Sci. Technol. Art Life*, pp. 394–403, 2019, [Online]. Available: <https://ptki.ac.id/jurnal/index.php/readystar/article/view/82>.
- [18] M. B. Nugroho, I. E. Hendrawan, and P. P. Purwantoro, "Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI," *Nuansa Inform.*, vol. 16, no. 1, pp. 125–133, 2022, doi: 10.25134/nuansa.v16i1.5294.
- [19] G. Gustientiedina, M. H. Adiya, and Y. Desnelita, "Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan," *J. Nds. Teknol. dan Sist. Inf.*, vol. 5, no. 1, pp. 17–24, 2019, doi: 10.25077/teknosi.v5i1.2019.17-24.
- [20] M. B. Nahjan, N. Heryana, and A. Voutama, "Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell," *J. Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 1–4, 2023.

2. Diararang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.