

Analisis Data Rekam Medis Pasien Berdasarkan Usia Di Puskesmas Sekernan Ilir Jambi

Widya Martha Bella¹, Reny Wahyuning Astuti², Novhirtamely Kahar³

^{1,2,3}Program Studi Teknik Informatika/Universitas Nurdin Hamzah

e-mail: ¹widyamrtha20@gmail.com, ²rer3ny4atuti@gmail.com, ³novmely@gmail.com

Abstrak

Analisis merupakan proses penting yang digunakan untuk memecah suatu keseluruhan menjadi komponen-komponen terpisah guna mengenali tanda-tanda dari setiap komponen serta memahami fungsi mereka dalam keseluruhan yang terpadu. Dalam konteks Puskesmas Sekernan Ilir di Kabupaten Muaro Jambi, sebuah aplikasi telah berhasil dikembangkan menggunakan alat bantu KNIME dan algoritma data mining K-Means. Aplikasi ini bertujuan untuk mengelompokkan diagnosa berdasarkan umur dan penyakit pasien. Proses dimulai dengan memasukkan data berupa nomor rekam medis, nama pasien, jenis kelamin, umur, alamat, dan diagnosa ke dalam aplikasi. Hasil dari analisis tersebut berupa informasi mengenai cluster diagnosa berdasarkan faktor umur dan penyakit, yang secara jelas menampilkan pola penyakit yang dominan pada rentang umur tertentu. Hasil yang diharapkan dari Analisa data rekam medis ini adalah dapat memberi kemudahan untuk mengetahui pengelompokkan penyakit dengan metode data mining, sehingga dapat mengetahui dari informasi yang dihasilkan apa saja penyakit yang sering diderita oleh pasien yang berobat pada Puskesmas Sekernan Ilir Jambi.

Kata kunci: Analisis ,Data Mining, KNIME, K-Means, Puskesmas.

Abstract

Analysis is an essential process employed to dissect a whole into separate components in order to recognize the characteristics of each component and comprehend their functions within the integrated whole. In the context of Puskesmas Sekernan Ilir in Muaro Jambi Regency, an application has been successfully developed using the KNIME tool and the K-Means data mining algorithm. This application aims to cluster diagnoses based on patients' age and diseases. The process commences by inputting data such as medical record numbers, patient names, gender, age, address, and diagnosis into the application. The outcome of this analysis is in the form of information concerning diagnosis clusters based on age and disease factors, which distinctly showcase prevalent disease patterns within specific age ranges. The anticipated result of this medical record data analysis is to facilitate the understanding of disease categorization through data mining methods, thereby discerning from the generated information the diseases commonly suffered by patients seeking treatment at Puskesmas Sekernan Ilir in Jambi.

Keywords: Analysis, Data Mining, KNIME, K-Means, Public Health Center.

1. Pendahuluan

Pusat Kesehatan Masyarakat (Puskesmas) adalah salah satu sarana pelayanan kesehatan masyarakat yang berfungsi sebagai unit pelaksana teknis dinas kabupaten/kota yang bertanggung jawab menyelenggarakan pembangunan kesehatan di suatu wilayah kerja salah satunya adalah Puskesmas Sekernan Ilir

Rekam medis merupakan berkas yang berisi tentang identitas pasien, catatan pemeriksaan, catatan pengobatan dan catatan pelayanan kesehatan lainnya (Peraturan Menteri Kesehatan No.55 Tahun 2013). Rekam Medis mempunyai tujuan untuk menunjang tercapainya tertib administrasi dalam rangka upaya peningkatan pelayanan kesehatan di puskesmas.

Seiring dengan perkembangan teknologi informasi dan pengolahan data, sejumlah penelitian sebelumnya telah menggaris bawahi pentingnya analisis data rekam medis berdasarkan usia dalam upaya pemahaman lebih mendalam tentang profil kesehatan populasi. Beberapa penelitian telah membahas tentang penerapan teknik data mining, seperti penggunaan algoritma K-Means, dalam mengklasifikasikan pola penyakit berdasarkan faktor usia. Namun, dalam konteks Puskesmas Sekernan Ilir Jambi, masih belum adanya fitur pembantu untuk menentukan hasil *cluster* penyakit berdasarkan usia.

Oleh karena itu, dalam penelitian ini, kami membangun sebuah aplikasi berbasis Analisis Data Rekam Medis Berdasarkan Usia dengan menggunakan metode *data mining*, khususnya algoritma K-Means. Aplikasi ini diharapkan dapat memberikan informasi yang lebih akurat dan efektif mengenai penyakit yang sering diderita oleh pasien yang berobat di Puskesmas Sekernan Ilir Jambi.

2. Metode Penelitian

2.1 Metodologi Penelitian

Metode ini disebut juga dengan penelitian langsung, karena peneliti akan berhadapan langsung dengan pihak yang terkait di tempat melakukan penelitian. Dalam hal ini peneliti mencari data-data yang diperlukan dengan cara:

1. Metode Observasi

Menggunakan pendekatan *cross sectional*. Dalam penelitian ini subyek yang digunakan adalah petugas rekam medis di bagian pendaftaran pada Puskesmas Sekernan Ilir Jambi.

2. Metode Wawancara

Peneliti melakukan tanya jawab secara langsung pelaksanaan penyimpanan dan pengambilan dokumen rekam medis pada Puskesmas Sekernan Ilir Jambi.

Menurut hasil wawancara dan observasi yang dilakukan data yang digunakan dari tahun 2021-2022 dengan 500 *records* yang memiliki 6 atribut. Adapun 6 atribut dari data itu adalah:

Tabel 2.1. Atribut Rekam Medis

Attribut	Keterangan
No Rekam Medis	Nomor Rekam Medis Pasien
Nama	Nama Pasien
Jenis Kelamin	Jenis Kelamin Pasien
Umur	Umur Pasien
Alamat	Alamat Tempat Tinggal Pasien
Diagnosa	Penyakit Pasien

Ada beberapa tahapan yang dilakukan dalam penelitian analisis data rekam medis berdasarkan usia di puskesmas sekernan ilir yaitu sebagai berikut:

1. Data Rekam Medis Pasien Puskesmas Sekernan Ilir Jambi

Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian. Teknik dalam pengumpulan data dalam penelitian merupakan faktor penting demi keberhasilan penelitian. Metode pengumpulan data adalah suatu cara yang dilakukan untuk mengumpulkan data. Teknik pengumpulan data yang digunakan adalah observasi dan wawancara.

2. Seleksi data berdasarkan umur dan penyakit

Agar masalah penelitian yang dipilih benar-benar tepat, maka masalah perlu dievaluasi atau dianalisis penyeleksian data. Seleksi data pada penelitian ini yaitu mencari umur dan penyakit pasien pada Puskesmas Sekernan Ilir.

3. Hasil Seleksi data dengan metode *K-Means*
Langkah selanjutnya yaitu melihat hasil seleksi yang telah dilakukan dengan metode *K-Means* apakah data nya telah cocok dan benar.
4. Analisa hasil data rekam medis menggunakan *K-Means*
Analisa hasil dilakukan agar kecocokan pada data benar atau tidaknya sehingga peneliti dapat melihat hasil dari menggunakan metode *K-Means*.
5. Menguji Hasil
Menguji hasil analisis dan perhitungan dari algoritma *K-Means Clustering* menghasilkan kelompok penyakit berdasarkan usia pada Puskesmas Sekernan Ilir Jambi. Pada analisis akhir.
6. Membuat hasil pengujian kesimpulan
Pada tahap ini dilakukan pembuatan laporan yang disusun berdasarkan hasil penelitian dengan menggunakan teknik pengumpulan data primer dan data sekunder sehingga menjadi laporan penelitian yang dapat memberikan gambaran utuh tentang sistem yang sedang dibangun.

2.2 Analisis

Analisis adalah kata yang sering terdengar pada suatu evaluasi kegiatan. Analisis sering dilakukan untuk memperoleh kesimpulan mengenai pelaksanaan kegiatan tersebut (Sahyani, 2023) [1]. Menurut kamus besar bahasa Indonesia, yang dimaksud dengan analisis adalah penyelidikan dan penguraian terhadap suatu masalah untuk mengetahui keadaan yang sebenarnya dan proses pemecahan masalah yang dimulai dengan dugaan dan kebenarannya (Magdalena *et al.*, 2020) [2].

2.3 Rekam Medis

Menurut peraturan Menteri Kesehatan Nomor: 749a/MenKes/Per/XII/1989 tentang rekam medis, rekam medis adalah berkas yang berisikan catatan dan dokumen tentang identitas pasien, pemeriksaan, pengobatan, tindakan, dan pelayanan lain kepada pasien pada sarana pelayanan Kesehatan (Abduh, 2021) [3]. Rekam medis adalah siapa, apa, di mana dan bagaimana perawatan pasien selama di puskesmas. Untuk melengkapi rekam medis harus dimiliki data yang cukup tertulis dalam rangkaian kegiatan guna menghasilkan suatu *diagnosis*, jaminan, pengobatan dan hasil akhir (SK Men PAN No. 135 tahun 2002) (Handayani & Ramadhan, 2020) [4].

2.4 Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terikat dari berbagai *database* besar (Ariansyah *et al.*, 2018) [5]. Berdasarkan pengertian data mining yang telah dijelaskan di atas, maka data mining merupakan pengetahuan yang tersembunyi di dalam *database* yang di proses untuk menemukan pola dan teknik statistik matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi pengetahuan dari *database* tersebut (Utomo and Mesran, 2020) [6].

2.5 K-Means

Algoritma *K-means* menggunakan proses secara berulang-ulang untuk mendapatkan basis data *cluster*. Dibutuhkan jumlah *cluster* awal yang diinginkan sebagai masukan dan menghasilkan titik *centroid* akhir sebagai output (Indraputra & Fitriana, 2020) [7]. Metode *K-means* akan memilih pola *k* sebagai titik awal *centroid* secara acak atau random. Jumlah iterasi untuk mencapai *cluster centroid* akan dipengaruhi oleh calon *cluster centroid* awal secara random. Sehingga didapat cara dalam pengembangan algoritma dengan menentukan *centroid cluster* yang dilihat dari kepadatan data awal yang tinggi agar mendapatkan kinerja yang lebih tinggi (Andika & Anisa, 2020) [8].

Berikut ini adalah tahap-tahap algoritma *k-means clustering* :

1. *Input* data yang digunakan dalam *clustering*. Data ini digunakan untuk menentukan nilai rata-rata *data point* yang berada dalam satu *cluster* dan menentukan jarak dari setiap *data point* ke *centroid*
2. Alokasi ke *cluster* secara acak, dalam tahap ini pertama kalinya data point dialokasikan ke *cluster* secara acak tanpa ada kriteria tertentu.
3. Hitung *centroid data point* yang ada pada setiap *cluster*. Nilai *centroid* pada *k-means* digunakan sebagai pusat *cluster*. Dengan menentukan anggota *cluster* secara acak pada tahap sebelumnya, maka terbentuk iterasi awal sebagai pusat *cluster* acak.
4. Alokasi ke *centroid* terdekat, pada tahap ini hasil *centroid* dari setiap *cluster* sudah diketahui, kemudian *datapoint* dialokasikan pada *centroid* terdekat berdasarkan nilai jarak *similarity data point* terhadap *centroid*. Jarak *similarity* dari *data point* ke *centroid* pada masing masing *cluster* diperoleh dari perhitungan *euclidean distance*. Kemudian nilai jarak setiap *data point* ke *centroid cluster* dibandingkan, dan *data point* menjadi anggota dari *cluster* berdasarkan jarak *data point* ke *centroid* terdekat.
5. Konvergen, mengalokasikan data point ke *centroid* dengan nilai jarak terdekat, dengan menguji apakah *cluster* yang terbentuk telah membentuk *cluster* yang konvergen atau tidak. *Cluster* dinyatakan konvergen jika anggota dari masing masing.

3. Hasil dan Pembahasan

3.1 Penerapan Metode K-Means Secara Manual

Berikut ini merupakan tahapan proses perhitungan menggunakan metode *K-Means* yang dilakukan pada penelitian ini antara lain sebagai berikut:

1. Menyiapkan data awal 500 records

Tabel 3.1 Dataset

No	No RM	Nama Pasien	Umur (Th)	Diagnosa
1	165	M. Panes	1	3
2	68	Nesya	1	38
3	300	Arumi Narsya	2	25
4	306	Zahra	5	10
5	783	Febriansyah	5	27
6	79	Almair	5	27
7	216	M. Fahri	7	38
8	626	M. Bian	7	32
9	4421	Icha	7	18
10	227	Azira	9	3

2. Menentukan banyak nya *cluster* dan titik *centroid* awal

Tabel 3.1 Titik Centroid awal

Row ID	Umur	Diagnosa
cluster_0	4.686	28.483
cluster_1	15.071	6.706
cluster_2	19.911	38.156
cluster_3	31.417	27.854
cluster_4	62.182	33.667
cluster_5	42.896	37.188
cluster_6	48.27	22.349
cluster_7	46.983	6.033

3. Setelah itu dilakukan perhitungan secara manual di *Microsoft Excell* dan tools *KNIME* untuk menentukan titik *centroid* baru apakah berubah atau tetap sama dan melihat berapa kali terjadi iterasi.

Tabel 3.2 Hasil *Clustering* ke-2

	X (Umur)	Y (Diagnosa)
C1	4.69	28.48
C2	15.07	6.71
C3	19.91	38.16
C4	31.40	27.87
C5	62.03	33.59
C6	42.73	37.02
C7	48.27	22.35
C8	46.90	5.95

Tabel 3.3 Hasil *Clustering* ke-3

	X (Umur)	Y (Diagnosa)
C1	4.69	28.48
C2	15.07	6.71
C3	19.91	38.16
C4	31.40	27.87
C5	62.03	33.59
C6	42.73	37.02
C7	48.27	22.35
C8	46.90	5.95

Tabel 3.4 Hasil *Clustering* ke-4

	X (Umur)	Y (Diagnosa)
C1	4.69	28.48
C2	15.07	6.71
C3	19.91	38.16
C4	31.40	27.87
C5	62.03	33.59
C6	42.73	37.02
C7	48.27	22.35
C8	46.90	5.95

4. Menghitung data secara manual
Data Ke-1

$$\text{Jarak Ke C1} = \sqrt{(1 - 4,68)^2 + (3 - 28,48)^2}$$

$$\text{Jarak Ke C2} = \sqrt{(1 - 15,07)^2 + (3 - 6,70)^2}$$

$$\text{Jarak Ke C3} = \sqrt{(1 - 19,91)^2 + (3 - 38,15)^2}$$

$$\text{Jarak Ke C4} = \sqrt{(1 - 31,41)^2 + (3 - 27,85)^2}$$

$$\text{Jarak Ke C5} = \sqrt{(1 - 62,18)^2 + (3 - 33,66)^2}$$

$$\text{Jarak Ke C6} = \sqrt{(1 - 42,89)^2 + (3 - 37,18)^2}$$

$$\text{Jarak Ke C7} = \sqrt{(1 - 48,27)^2 + (3 - 22,34)^2}$$

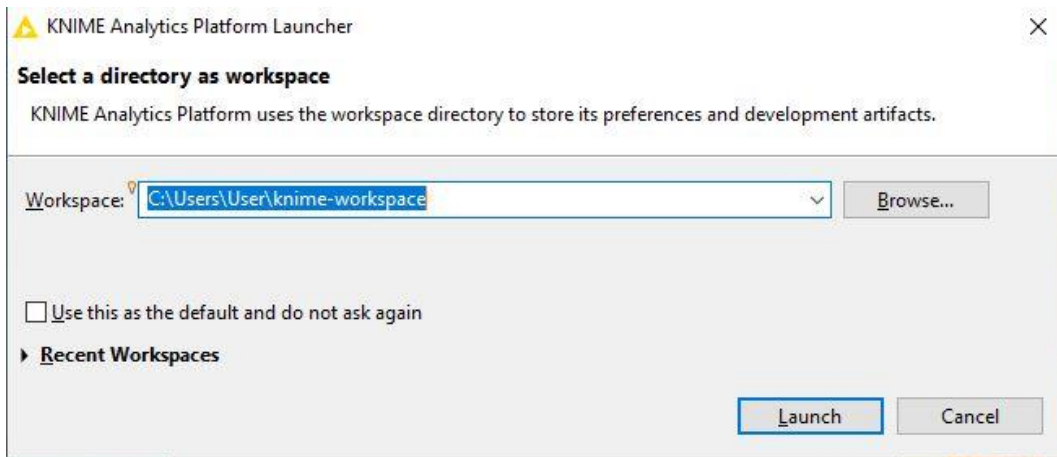
$$\text{Jarak Ke C8} = \sqrt{(1 - 46,98)^2 + (3 - 6,03)^2}$$

Iterasi berhenti pada iterasi ke-4 karena nilai *centroid* nya sudah tidak berubah sehingga langsung dilanjutkan dengan menggunakan *tools* KNIME

3.2 Implementasi Menggunakan KNIME

1. Tampilan Awal

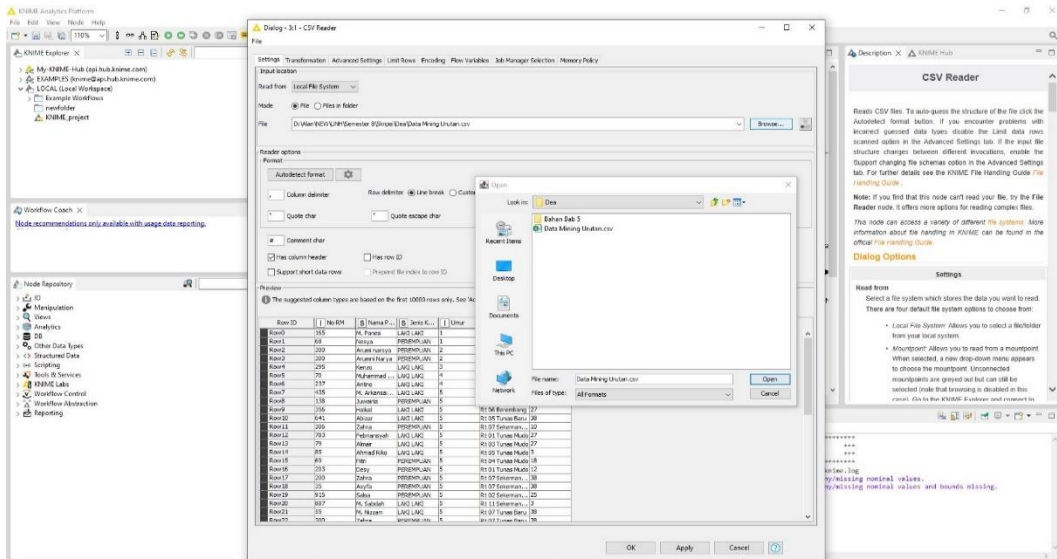
Pada tampilan awal ini terdapat pemilihan *workspace* dimana ingin di simpan file nya dan ketika sudah terpilih pilih *launch*.



Gambar 1. Tampilan Awal

2. Proseses *Import* Data

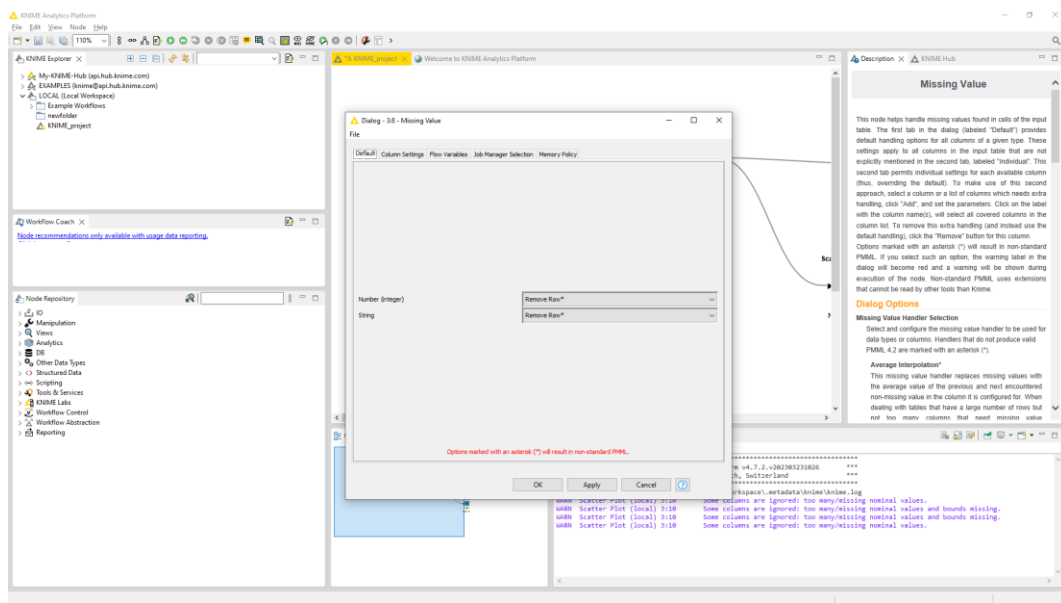
Saat workspace sudah terbuka kita import CSV Reader pada menu Node Repository. Pada proses ini dilakukan proses *import* data dari *Microsoft Excel* dengan aplikasi KNIME.



Gambar 2. Proses *Import* Data

3. Proseses *Missing Value*

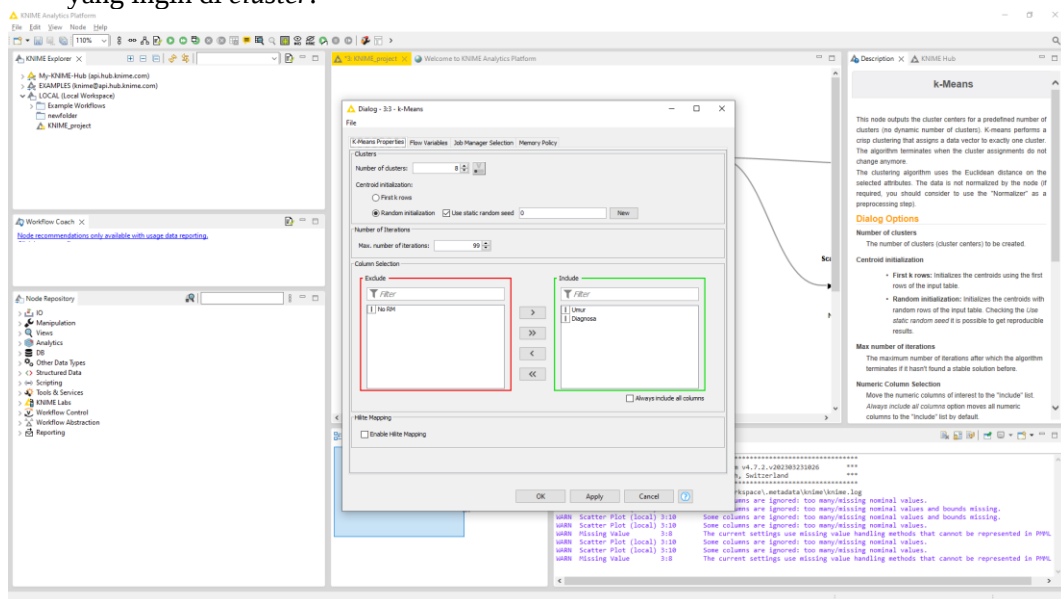
Setelah melakukan *import* data selanjutnya melakukan Proses *Missing Value* berguna untuk mencari kesalahan pada data yang telah di input, apakah terdapat data yang rusak atau lain sebagainya.



Gambar 3. Proses Missing Value

4. Proses Clustering K-Means

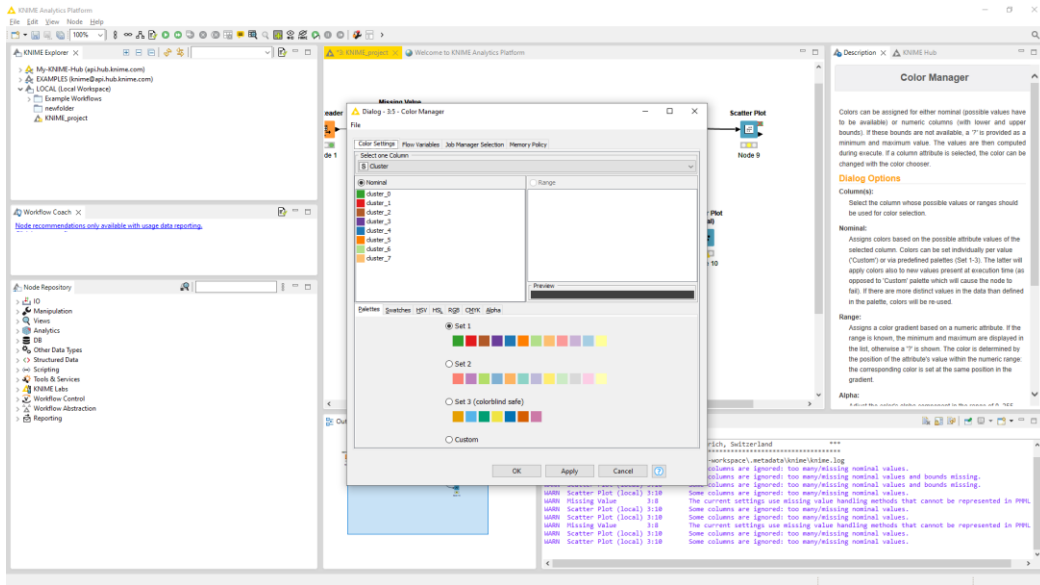
Selanjutnya melakukan melakukan *clustering K-Means* menggunakan *tools* KNIME yang mana kita telah memasukkan *tools K-Means* pada KNIME dan dapat kita lihat sebagai berikut. Pada *configure* kita bisa melakukan jumlah cluster dan *column* mana yang ingin di *cluster*.



Gambar 4. Proses Clustering K-Means

5. Pemilihan Color Manager

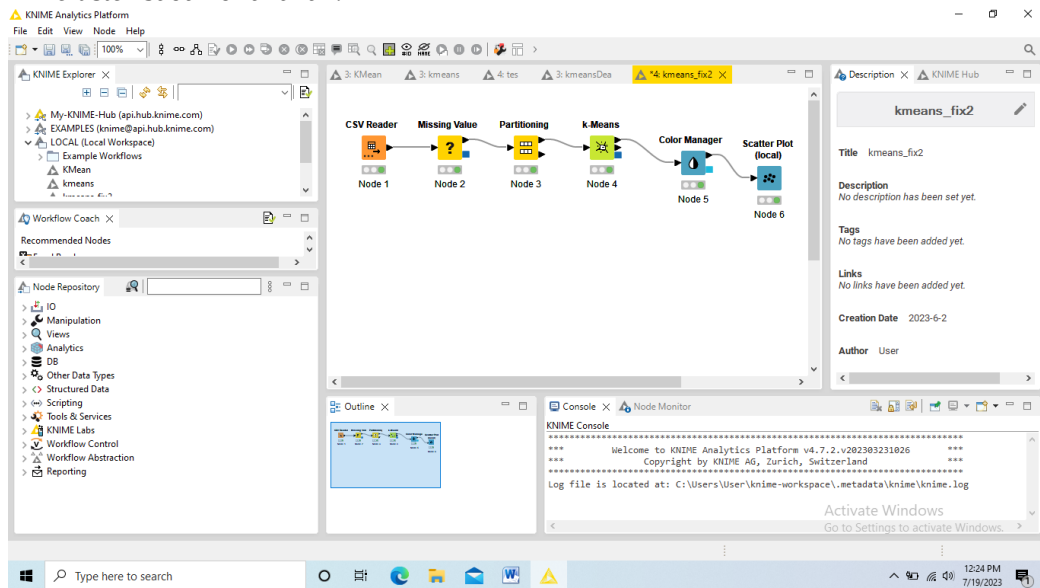
Selanjutnya pemilihan warna pada tiap titik *cluster* menggunakan *tools Color Manager* berguna untuk memilih warna dari tiap *cluster* yang telah dilakukan, sebagaimana dapat dilihat pada gambar.



Gambar 5. Pemilihan *Color Manager*

6. Tampilan KNIME

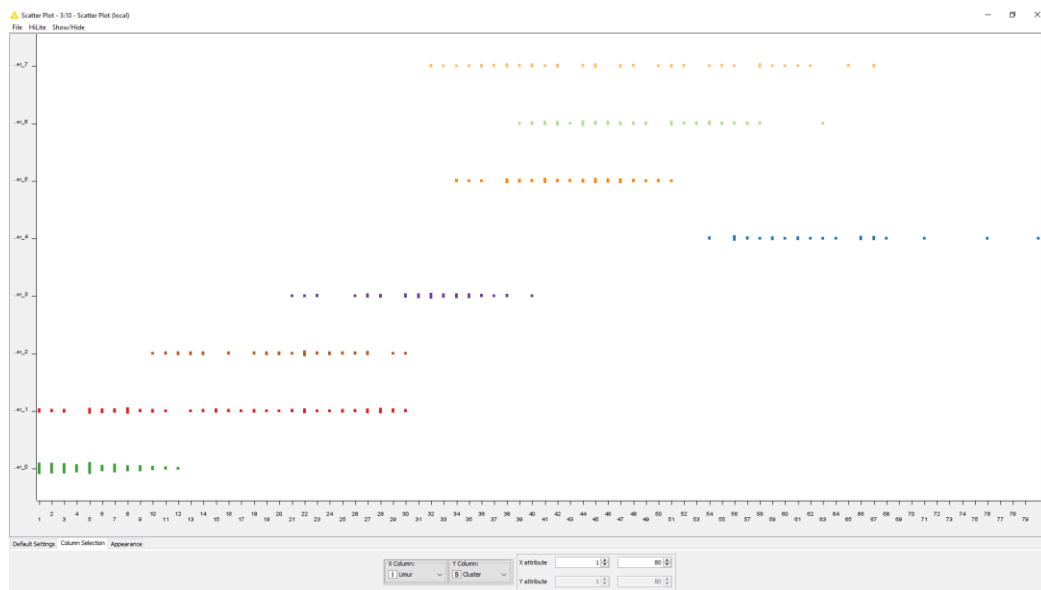
Ini adalah tampilan *nodes-nodes* yang telah di *configure* maka lampu dari *nodes* tersebut menunjukkan warna hijau yang berarti semua *nodes* telah tersambung dan proses *cluster* sudah dilakukan.



Gambar 6. Tampilan KNIME

7. Output Cluster K-Means

Output Cluster K-Means menggunakan tools Scatter Plot untuk membaca data menjadi plot agar data mudah untuk dibaca.



Gambar 7. Output Cluster K-Means

4. Kesimpulan

Dari 500 records data yang di olah menggunakan metode *K-Means* dengan bantuan tools KNIME menggunakan 8 cluster antara tools dan pencarian manual di *Microsoft Excel* menghasilkan akurasi sebesar 99,2% dengan hasil sebagai berikut:

Hasil yang didapat setelah melakukan perhitungan dan berdasarkan iterasi terakhir yaitu :

1. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 0-12 tahun cluster 1 memiliki penyakit terbanyak Ispa
2. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 1-30 tahun cluster 2 memiliki penyakit terbanyak C.Cold
3. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 10-30 tahun cluster 3 memiliki penyakit terbanyak Scabies
4. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 21-38 tahun cluster 4 memiliki penyakit terbanyak Ispa
5. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 54-80 tahun cluster 5 memiliki penyakit terbanyak Rujuk
6. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 33-51 tahun cluster 6 memiliki penyakit terbanyak Hipertensi
7. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 39-51 tahun cluster 7 memiliki penyakit terbanyak Hipertensi
8. Berdasarkan perhitungan iterasi terakhir dan berdasarkan dari kedelapan cluster, umur 32-67 tahun cluster 8 memiliki penyakit terbanyak C.Cold

Daftar Pustaka

- [1] Sahyani, N. 2023 *Analisis Strategi Pembelajaran Cooperative Learning Model Teams Games Tournament dalam Penanaman Sikap Sosial Siswa di Kelas V SD Negeri 17 Krui Pesisir Barat* (Tesis doktor, UIN Raden Intan Lampung).
- [2] Magdalena, I., Sundari, T. and Nurkamilah, S., 2020. Nasrullah, and Dinda Ayu Amalia.“. *Analisis Bahan Ajar*’, *Nusantara: Jurnal Pendidikan Dan Ilmu Sosial*, 2, pp.311-26..
- [3] Abduh, R. 2021 'Kajian Hukum Rekam Medis Sebagai Alat Bukti Malapraktik Medis', *De Lega Lata: Jurnal Ilmu Hukum*, 6(1), pp. 221-234.

- [4] Handayani, N. and Ramadhan, Y.E., 2020. Analisis Sistem Informasi Pendataan Rekam Medis Pasien Pengguna Bpjs Pada Klinik Rohmatan Nur Al Amim. *JIKA (Jurnal Informatika)*, 4(3), pp.1-8.
- [5] Ariansyah, T., Purwadi, P., & Yakub, S. 2018 'Implementasi Data Mining Untuk Mengestimasi Kebutuhan Persediaan Roti Panggang Di Junction Cafe Dengan Menggunakan Metode Regresi Linier Berganda', *Jurnal Cyber Tech*, 1(8).
- [6] Utomo, D.P. and Mesran, M., 2020. Analisis komparasi metode klasifikasi data mining dan reduksi atribut pada data set penyakit jantung. *Jurnal Media Informatika Budidarma*, 4(2), pp.437-444.
- [7] Indraputra, R. A., & Fitriana, R. 2020 'K-Means Clustering Data COVID-19', *Jurnal Teknik Industri*, 10(3), pp. 275-282.
- [8] Andika, T.H. and Anisa, N.S., 2020. Sistem Identifikasi Citra Daun Berbasis Segmentasi Dengan Menggunakan Metode K-Means Clustering. *vol*, 2, pp.9-17.