



PENGELOMPOKAN TEKANAN DARAH LANSIA DENGAN ALGORITMA K-MEANS DI KP.LEBAK JERO

Elsa Safutri¹, Ade Irma Purnamasari², Agus Bahtiar³, Edi Wahyudin⁴

^{1,2}Program Studi Teknik Informatika, STMIK IKMI Cirebon

³Program Studi Sistem Informasi, STMIK IKMI Cirebon

⁴Program Komputerisasi Akuntansi, STMIK IKMI Cirebon

Jln. Perjuangan No.10B, Karyamulya. Kec. Kesambi. Kota Cirebon, Jawa Barat 45135

[1elsasafutri02@gmail.com](mailto:elsasafutri02@gmail.com) , [2irma2974@yahoo.com](mailto:irma2974@yahoo.com) , [3agusbahtiar038@gmail.com](mailto:agusbahtiar038@gmail.com) , ⁴

ediwahyudin@gmail.com

Abstract

Hypertension is one of the main health problems often experienced by the elderly and requires appropriate treatment. However, limited resources at Posbindu are often an obstacle in effectively identifying hypertension risk groups. This study aims to design a model for clustering the risk of hypertension in the elderly at Posbindu Kp. Lebak Jero using the k-means algorithm, in support of more targeted health interventions. The data analyzed included systolic, diastolic blood pressure, age, gender, and body weight, based on Posbindu records during August-September 2024. The analysis process is carried out through the stages of Knowledge Discovery in Database (KDD), which includes collection, pre-processing, data transformation, clustering, and evaluation. To evaluate the quality of the clustering results, the Davies-Bouldin Index (DBI) was used. The results showed that the optimal DBI was 0.881 with k=2. The first cluster (Cluster_0) consists of 75 low-risk elderly (19 people with normal blood pressure and 56 pre-hypertension), with an age range of 45-59 years (46 people), 60-69 years (22 people), and >70 years (7 people). The second cluster (Cluster_1) consisted of 71 high-risk elderly (28 people with grade 1 hypertension and 43 people with grade 2 hypertension), with an age range of 45-59 years (26 people), 60-69 years (30 people), and >70 years (15 people). Systolic blood pressure in Cluster_0 was in the range of 80-143 mmHg and diastolic 80-90 mmHg, while in Cluster_1 systolic 140-200 mmHg and diastolic 78-130 mmHg. This study shows that the k-means algorithm is able to map hypertension risk groups and assist monitoring and support the implementation of more effective and targeted health interventions at Posbindu Kp. Lebak Jero.

Keywords : *K-means, Data Clustering, Blood Pressure, Posbindu, Davies-Bouldin Index (DBI).*

Abstrak

Hipertensi merupakan salah satu masalah kesehatan utama yang sering dialami oleh lansia dan memerlukan penanganan yang tepat. Namun, keterbatasan sumber daya di Posbindu sering kali menjadi hambatan dalam mengidentifikasi kelompok risiko hipertensi secara efektif. Penelitian ini bertujuan untuk merancang model pengelompokan risiko hipertensi pada lansia di Posbindu Kp. Lebak Jero menggunakan algoritma k-means, dalam mendukung intervensi kesehatan yang lebih terarah. Data yang dianalisis mencakup tekanan darah sistolik, diastolik, usia, jenis kelamin, dan berat badan, berdasarkan catatan Posbindu selama bulan Agustus-September 2024. Proses analisis dilakukan melalui tahapan *Knowledge Discovery in Database (KDD)*, yang meliputi pengumpulan, pra-pemrosesan, transformasi data, pengelompokan, dan evaluasi. Untuk mengevaluasi kualitas hasil pengelompokan, digunakan *Davies-Bouldin Index (DBI)*. Hasil penelitian menunjukkan bahwa *DBI* optimal sebesar 0,881 dengan k=2. Cluster pertama (*Cluster_0*) terdiri dari 75 lansia berisiko rendah (19 orang dengan tekanan darah normal dan 56 orang pra-hipertensi), dengan rentang usia 45-59 tahun (46 orang), 60-69 tahun (22 orang), dan >70 tahun



(7 orang). Cluster kedua (*Cluster_1*) terdiri dari 71 lansia berisiko tinggi (28 orang hipertensi tingkat 1 dan 43 orang hipertensi tingkat 2), dengan rentang usia 45-59 tahun (26 orang), 60-69 tahun (30 orang), dan >70 tahun (15 orang). Tekanan darah sistolik pada *Cluster_0* berada pada kisaran 80-143 mmHg dan diastolik 80-90 mmHg, sementara pada *Cluster_1* sistolik 140-200 mmHg dan diastolik 78-130 mmHg. Penelitian ini menunjukkan bahwa algoritma *k-means* mampu memetakan kelompok risiko hipertensi dan membantu pemantauan serta mendukung pelaksanaan intervensi kesehatan yang lebih efektif dan terarah di Posbindu Kp. Lebak Jero.

Kata kunci : *K-means*, Pengelompokan Data, Tekanan Darah, Posbindu, *Davies-Bouldin Index (DBI)*.

1. PENDAHULUAN

Kemajuan teknologi informasi yang telah memberikan dampak besar pada berbagai aspek kehidupan, termasuk bidang medis. Pengenalan teknologi informasi di sektor kesehatan memungkinkan pengelolaan data yang lebih baik. Salah satu penerapan teknologi ini yaitu dengan pemanfaatan data *Mining* dalam mendukung pengambilan keputusan berbasis data.[1]

Data *Mining* adalah suatu proses untuk mencari suatu pola atau suatu model yang sempurna, bermanfaat serta dapat dimengerti dari sebuah database. Pola atau informasi yang didapatkan dari database dapat digunakan untuk membantu dalam mengambil suatu keputusan di waktu yang akan datang[2][3].

Salah satu metode *data Mining* yang efektif adalah algoritma *k-means clustering*, yang dapat digunakan untuk mengidentifikasi pola dalam data kompleks. Dalam konteks kesehatan lansia, pengelompokan data tekanan darah berdasarkan hasil pemeriksaan dan karakteristik individu merupakan pendekatan penting untuk memudahkan pemantauan dan meningkatkan akurasi penentuan risiko kesehatan[4][5].

Tekanan darah tinggi merupakan salah satu masalah kesehatan utama di kalangan lansia yang meningkat seiring bertambahnya usia dimana memerlukan penanganan yang optimal untuk mencegah komplikasi serius. Namun, dalam praktiknya, data kesehatan lansia seringkali tidak terkelola dengan baik, terutama di tingkat pelayanan kesehatan masyarakat seperti Pos Pembinaan Terpadu (Posbindu). Posbindu sebagai pusat Pembangunan terpadu yang berperan penting dalam memantau kesehatan masyarakat, khususnya kelompok rentan seperti lansia. Tantangan yang dihadapi adalah bagaimana mengelola data kesehatan tersebut

agar dapat menghasilkan pola risiko yang spesifik, sehingga mempermudah perencanaan intervensi [6] [7].

Berdasarkan data yang diperoleh dari Posbindu Kp.Lebak Jero dari laporan kesehatan pada bulan Agustus sampai September 2024, diperoleh 189 pasien lansia dengan variasi tekanan darah dengan atribut lain seperti umur, jenis kelamin, serta berat badan. Data ini menunjukkan pentingnya pendekatan berbasis teknologi seperti algoritma *k-means* untuk mengelompokkan data tersebut. Metode ini telah terbukti efektif dalam berbagai penelitian sebelumnya, misalnya pengelompokan penyakit untuk mengetahui pola/jenis gangguan kesehatan dengan jumlah banyak maupun sedikit.[8], distribusi obat [9], mengelompokkan penderita penyakit Kronis pada warga lansia RW 07 kelurahan Pulogebang [10]. Namun, belum banyak penelitian yang secara spesifik menggunakan algoritma ini untuk menganalisis data tekanan darah lansia di Posbindu.

Selain itu, meskipun algoritma *k-means* telah banyak diterapkan di berbagai bidang kesehatan, penerapannya pada pengelompokan risiko hipertensi lansia di Posbindu masih terbatas. Penelitian sebelumnya menunjukkan bahwa algoritma ini dapat membantu mengelompokkan data secara efisien dan menghasilkan wawasan baru yang relevan untuk mendukung perencanaan intervensi [11]. Namun, dalam konteks lokal seperti Posbindu Kampung Lebak Jero, pendekatan berbasis data masih kurang dimanfaatkan.

Hipertensi pada lansia merupakan salah satu penyebab utama penurunan kualitas hidup yang dapat meningkatkan risiko penyakit kardiovaskular, stroke, hingga kematian dini. Meskipun Posbindu menyediakan layanan untuk memantau tekanan darah, pengelolaan data yang



tidak terstruktur sering kali menjadi kendala. Akibatnya, sulit untuk mendeteksi pola risiko atau menentukan prioritas intervensi kesehatan yang tepat. Berdasarkan pengamatan langsung di Posbindu Kp.Lebak Jero, data kesehatan pasien masih dikelola secara manual, sehingga analisis lebih lanjut untuk mendukung pengambilan keputusan tidak dapat dilakukan secara optimal.

Selain itu, meskipun algoritma *k-means* telah banyak diterapkan di berbagai bidang kesehatan, penerapannya pada pengelompokan risiko hipertensi lansia di Posbindu masih terbatas. Penelitian sebelumnya menunjukkan bahwa algoritma ini dapat membantu mengelompokkan data secara efisien dan menghasilkan wawasan baru yang relevan untuk mendukung perencanaan intervensi [11]. Namun, dalam konteks lokal seperti Posbindu Kampung Lebak Jero, pendekatan berbasis data masih kurang dimanfaatkan.

Penelitian ini bertujuan untuk meningkatkan model pengelompokan tekanan darah pada lansia menggunakan algoritma *k-means* di Posbindu Kp.Lebak Jero, dengan identifikasi kelompok risiko dan intervensi yang lebih tepat sasaran. Kontribusi utama penelitian terletak pada pengembangan metode yang dapat meningkatkan efisiensi dan akurasi pengelompokan pasien hipertensi berdasarkan faktor risiko yang berbeda. Pentingnya penelitian ini terletak pada potensi penerapan algoritma *k-means* pada pengelolaan data kesehatan yang lebih terstruktur, sehingga dapat meningkatkan kualitas pelayanan kesehatan di Posbindu.

Penelitian ini diharapkan dapat memberikan kontribusi pada pengelolaan data kesehatan yang lebih efektif di Posbindu, sehingga meningkatkan kualitas pelayanan kesehatan bagi lansia. Selain itu, hasil penelitian ini juga dapat menjadi dasar bagi pengembangan program kesehatan berbasis data pada skala yang lebih luas, serta memperkaya literatur tentang penerapan algoritma *k-means* dalam bidang kesehatan. Dengan mendukung analisis yang lebih akurat

Penelitian ini menggunakan pendekatan *Knowledge Discovery in Database (KDD)*, yang meliputi pengumpulan data, *preprocessing*, transformasi data, pengelompokan dengan algoritma *k-means*, dan evaluasi hasil *cluster* menggunakan *DBI Software* seperti *RapidMiner* akan digunakan untuk mempermudah proses analisis dan visualisasi data. Dengan metode ini, diharapkan penelitian ini tidak hanya

menghasilkan pengelompokan yang lebih akurat tetapi juga menjadi rujukan dalam pengembangan teknologi informasi kesehatan untuk lansia yang terbukti efektif pada pengelompokan kesehatan [12] [4].

2. TINJAUAN PUSTAKA

Penelitian terkait penggunaan algoritma *k-means* dalam pengelompokan data dibidang kesehatan telah dilakukan dengan berbagai pendekatan. Seperti penelitian [5] membahas pengelompokan jenis penyakit pasien di Puskesmas Cigugur Tengah dengan algoritma *k-means*, menggunakan metode Euclidean Distance. Penelitian ini menunjukkan bahwa algoritma *k-means* dapat membantu memahami distribusi penyakit untuk pengambilan keputusan. Demikian pula [6] menerapkan algoritma *k-means* dalam memetakan wilayah dengan prevalensi hipertensi berdasarkan data spasial. Hasil penelitian ini menunjukkan keberhasilan *k-means* dalam mengidentifikasi daerah risiko hipertensi tinggi dengan visualisasi spasial.

Studi lain, seperti yang dilakukan oleh [9] menggunakan algoritma *k-means* untuk mengelompokkan data obat berdasarkan tingkat pemakaian dengan evaluasi *cluster DBI* menghasilkan 0.513, yang menunjukkan kualitas optimal. Penelitian oleh [13] juga menekankan perbandingan algoritma *k-means* dan *k-medoids*, dengan hasil menunjukkan bahwa *k-means* memiliki kinerja lebih optimal berdasarkan nilai *DBI* terkecil 0,265.

Dalam konteks hipertensi, penelitian [3] ini digunakan metode *clustering* dengan algoritma *K-means* untuk mengelompokkan provinsi berdasarkan tingkat kemiskinannya dengan bantuan *Software Weka*. Hasil penelitian menunjukkan terdapat 3 *cluster* provinsi di Indonesia berdasarkan tingkat kemiskinannya. Penelitian lainnya oleh [14] mengelompokkan pasien hipertensi berdasarkan tingkat keparahan menggunakan variabel tekanan darah dan berat badan, menghasilkan 2 *cluster* dengan mayoritas pasien dengan kategori hipertensi berat.

Beberapa studi juga mengintegrasikan metode *k-means* untuk pengelompokan data pasien rawat inap atau rawat jalan. Seperti penelitian [15] mengidentifikasi pola prevalensi penyakit di Kabupaten Ngawi, yang menghasilkan 4 *cluster* berbasis karakteristik penyakit. Penelitian [16] juga menggunakan metodologi *CRISP-DM* dalam mengelompokkan kasus

penyakit RSUD Kota Bandung berdasarkan usia, memberikan wawasan distribusi penyakit berdasarkan demografi.

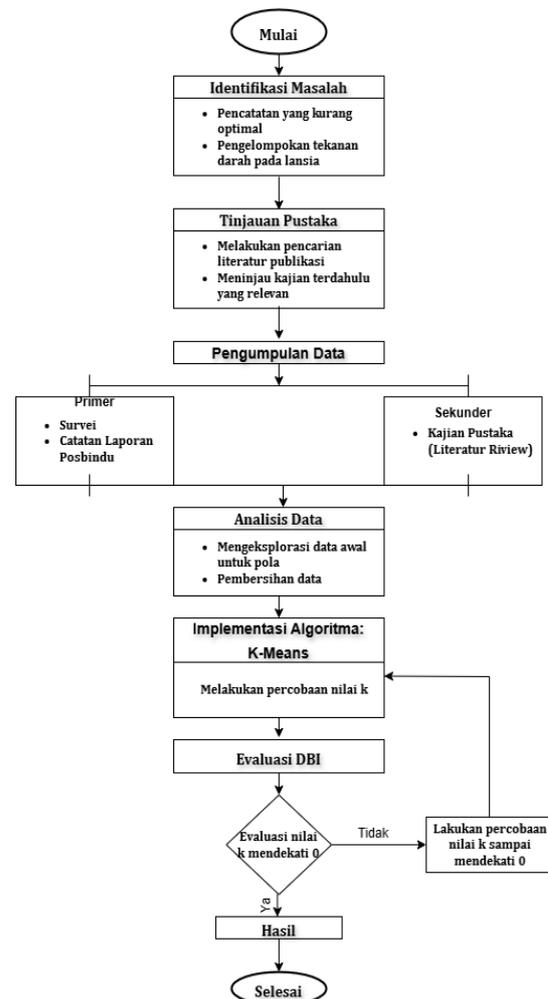
Pada penelitian ini kurangnya terletak pada penerapan algoritma *k-means* dalam pengelompokan risiko tekanan darah lansia secara spesifik di Posbindu, yang dapat memberikan strategi berbasis data untuk pencegahan dan penanganan hipertensi. Penelitian sebelumnya lebih banyak berfokus pada dataset rawat jalan, prevalensi penyakit, atau stok obat tanpa penekanan spesifik pada lansia di lingkungan posbindu.

Maka dari itu penelitian ini melakukan pendekatan baru dengan menggunakan algoritma *k-means* untuk mengelompokkan risiko tekanan darah pada lansia di Posbindu Kp. Lebak Jero. Penggunaan indikator seperti nilai *Davies-Bouldin Index* untuk mengevaluasi efektivitas pengelompokan memberikan kontribusi tambahan terhadap pengembangan model pengelompokan di bidang kesehatan lansia.

3. METODOLOGI PENELITIAN

Penelitian ini menggunakan metode kuantitatif yang dimana proses yang dilakukan dengan menggunakan data berupa angka-angka dengan teknik statistik [17]

Tahapan penelitian yang diterapkan pada penelitian ini meliputi beberapa tahapan untuk hasil yang maksimal, sebagai berikut.



Gambar 1. Diagram Tahapan Penelitian
Dari tahapan diatas dapat dijelaskan bahwa

3.1 Identifikasi Masalah

Malakukan analisis dan identifikasi masalah secara menyeluruh setelah topik ditentukan menggunakan algoritma *k-means*.

3.2 Tinjauan Pustaka

Tinjauan Pustaka dilakukan dengan beberapa tahap, diantaranya:

a. Formulasi permasalahan

Formulasi Permasalahan mencari dan meninjau literatur atau kumpulan tulisan dan publikasi ilmiah yang relevan dengan topik permasalahan,

b. Pencarian literatur

pencarian literatur tentang topik yang akan memebrikan Gambaran umum tentang model.



c. Evaluasi data

d. Analisis data interpretasi

3.3 Pengumpulan Data

Data yang digunakan merupakan data primer yang diperoleh secara langsung dari Posbindu Kp.Lebak Jero, yang meliputi tekanan darah sistolik,diastolik, usia, umur, jenis kelamin, dan berat badan lansia. Untuk mengetahui prevalensi hipertensi dalam populasi tertentu data yang dikumpulkan mulai dari bulan Agustus sampai bulan September 2024 sebanyak 189 orang lansia yang akan dianalisis. Sedangkan data sekunder yang dipakai yaitu data yang diperoleh dari jurnal dan website pengkategorian hipertensi menurut JNC-VII 2003. Langkah- langkah pengumpulan data adalah sebagai berikut :

a.Verifikasi data dilakukan untuk mengecek konsistensi catatan tekanan darah yang sudah terdokumentasi.

b. Pengkategorian tekanan darah berdasarkan ketentuan.

c.Pengkategorian lansia dengan mengelompokkan umur lansia berdasarkan rentang usia sesuai dengan Permenkes RI: Pra-lansia: 45-59 tahun, Lansia: 60-69 tahun, dan lansia risiko tinggi :> 70 tahun.

3.3 Analisis Data

Analisis data dalam penelitian ini dilakukan untuk mengevaluasi kualitas pengelompokan data tekanan darah lansia menggunakan algoritma *k-means*. Pengelompokan ini bertujuan untuk memberikan pemahaman yang lebih mendalam tentang pola tekanan darah lansia, sehingga dapat membantu tenaga kesehatan menentukan strategi intervensi yang tepat untuk setiap kelompok risiko. Objek analisis yaitu data tekanan darah lansia yang diperoleh dari pemeriksaan di Posbindu Kp. Lebak Jero, mencakup tekanan darah sistolik, diastolik, dan informasi relevan lainnya. Proses analisis dilakukan menggunakan aplikasi *RapidMiner*, yang menyediakan berbagai alat untuk *preprocessing* data, transformasi, pengelompokan, dan evaluasi kualitas *cluster*.

Sebelum analisis dilakukan, data melalui tahap persiapan, di mana dataset dibersihkan dari nilai yang hilang dan anomali menggunakan operator seperti *Remove Duplicate*. Selanjutnya, pada tahap transformasi, data nominal diubah menjadi data numerik menggunakan operator *Nominal to Numerical* untuk memastikan format data sesuai untuk proses *clustering*. Setelah data siap, algoritma *k-means* diterapkan untuk mengelompokkan data berdasarkan kesamaan antar atribut. Hasil pengelompokan kemudian dievaluasi menggunakan *Davies-Bouldin Index (DBI)*, yang mengukur seberapa baik *cluster* yang terbentuk terpisah satu sama lain. Nilai *DBI* yang lebih rendah menunjukkan *cluster* yang lebih berkualitas.

Dengan pendekatan ini, analisis data dilakukan secara sistematis dan terukur. Proses ini memberikan wawasan yang akurat tentang pola tekanan darah lansia, sehingga mendukung pengambilan keputusan yang lebih efektif dalam upaya pencegahan dan penanganan risiko kesehatan.

3.4. Implementasi Algoritma *K-means*

Penerapan algoritma *k-means clustering* pada data lansia di Posbindu akan dilakukan setelah data melalui tahap *preprocessing*. Pemilihan algoritma *clustering* yang tepat akan disesuaikan dengan karakteristik data dan tujuan dari analisis. Algoritma ini akan digunakan untuk mengelompokkan lansia ke dalam kelompok-kelompok berdasarkan kesamaan yang dimiliki [4]. Algoritma *k-means* diuji sebanyak 10 kali untuk menemukan *cluster* yang optimal.

3.5 Evaluasi *DBI*

Evaluasi *Davies-Bouldin Index (DBI)* digunakan untuk menentukan jumlah *cluster* optimal, dengan nilai *DBI* yang lebih kecil (mendekati 0) menunjukkan *cluster* yang lebih baik serta hasil pengelompokan divisualisasikan secara grafis untuk memudahkan pemahaman pola antar *cluster* [11].

Metode penelitian yang menganalisis data menggunakan data *mining* menggunakan proses *KDD* [18]. Dibawah ini diagram fase *KDD*.



Gambar 2. Proses KDD

Adapun deksripsi pada tahap diatas dijelaskan pada tabel dibawah ini:

Tabel 1. Deskripsi Proses KDD

No	Tahapan	Deskripsi
1.	Data	kumpulan informasi yang akan digunakan sebelum dilakukan proses <i>Knowledge Discovery in Database (KDD)</i>
2.	<i>Data Selection</i>	Data yang diperoleh dan dikumpulkan yang digunakan dan relevan untuk analisis
3	<i>Preprocessing Data</i>	Pada tahap ini dilakukan proses penghilangan data data yang tidak lengkap, tidak konsisten atau mengandung noise, hanya data bersih yang digunakan
4	<i>Transformation Data</i>	Pengubahan data dan pembuatan atribut baru yang akan digunakan dalam analisis
5	<i>Data Mining</i>	Diterapkan algoritma <i>k-means</i> dengan pembentukan <i>cluster</i> yang terbentuk berdasarkan pola.
6	Evaluasi	Evaluasi dilakukan menggunakan Davies Bouildin Index (<i>DBI</i>) untuk mengetahui seberapa efektif model tersebut

4. HASIL DAN PEMBAHASAN

Untuk mengelompokkan data tekanan darah pada lansia di Posbindu Kp.Lebak Jero menggunakan metode Knowledge Discovery in databse menggunakan software *RapidMiner AI Studio 2024.1.0*.

4.1 Selection data

Data yang diproses adalah laporan kesehatan dari Posbindu yang berjumlah 189 orang lansia, dengan 5 atribut yang mencakup jenis kelamin, umur, berat badan, tekanan darah sistolik dan diastolik. Dibawah ini terdapat dataset sebelum diproses.

Tabel 2. Dataset

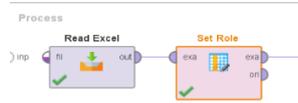
N o	Nama	Jenis Kela min	Um ur	Ber at Badan	Sist olik	Diast olik
1	Hj. Oom	P	82	54	180	130
2	Yati	P	70	45	140	80
3	Waki dah	P	59	64	160	90
....
1 8 8	Kos wara	L	64	47	162	125
1 8 9	Yuha nah	P	80	46	157	86

Selanjutnya tahap pertama dalam *clustering* yaitu memasukkan operator *read excel* yang ada pada aplikasi *RapidMiner*. Dengan fungsi membaca informasi pengelompokkan berdasarkan atribut yang digunakan.



Gambar 3. Read Excel

Setelah *read excel* dijalankan, maka sistem akan menampilkan preview data yang dimana akan menampilkan data yang akan digunakan. Selanjutnya menambahkan operator *Set Role* yang digunakan untuk menandai atribut yang akan menjadi id dari dataset tersebut.



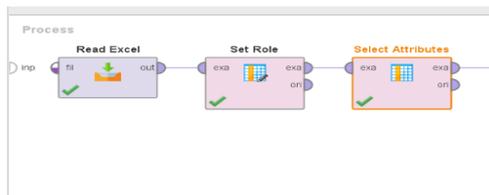
Gambar 4. Set Role

Setelah memasukkan operator *Set Role*, lalu edit list pada bagian parameter san pilih id yang digunakan. Dibawah ini preview dari operator *Set Role*.

Row No.	Nama	Jenis Kela...	Umur	Berat Badan	Sistolik	Diastolik
1	Hj Omm	P	82	54	180	130
2	Yati	P	70	45	140	80
3	Wakidah	P	59	64	160	90
4	Oth	P	59	53	160	90
5	Neni	P	59	62	110	70
6	Och	P	75	55	160	90
7	Dede Juariah	P	49	54	160	90
8	Yayah Mard...	P	51	58	160	100
9	Nenah	P	79	43	130	80
10	Nunung M	P	63	55	120	80
11	Nia Kurniawati	P	67	58	150	90
12	Halimah	P	60	64	160	90
13	Odah	P	64	44	170	90

Gambar 5. preview Set Role

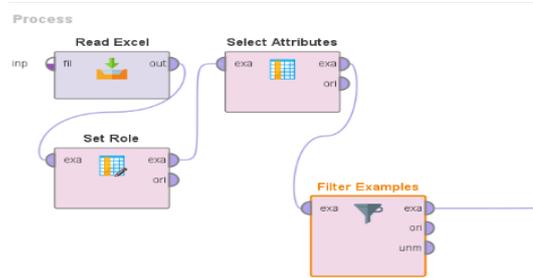
Tahap berikutnya yaitu memasukkan operator *Select Attribute* yang memiliki peran untuk menentukan atribut yang akan digunakan dalam proses *clustering data Mining*.



Gambar 6. Operator Select Attribute

4.2 Preprocessing Data

Preprocessing data merupakan tahapan pembersihan data hanya data yang relevan dan sesuai yang akan digunakan dalam pemodelan. Dalam tahap ini terdapat dua tahapan, pertama dengan menambahkan operator *Filter Examples* dengan tujuan memisahkan data lansia berdasarkan batas minimum usia 45 tahun. Dapat dilihat pada gambar dibawah ini.



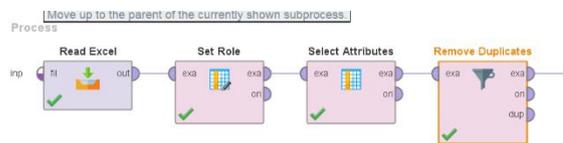
Gambar 7. Operetor Filter Examples

Setelah operator dimasukkan dan dijalankan dataset yang berawal 189 lansia menjadi 146 lansia. Ini membuktikan bahwa operator tersebut berjalan dan hanya menampilkan data dengan usia lebih dari 45 tahun.

Row No.	Nama	cluster	Jenis Kela...	Jenis Kela...	Umur	Berat Badan	Sistolik	Diastolik
1	Hj Omm	cluster_1	1	0	82	54	180	130
2	Yati	cluster_0	1	0	70	45	140	80
3	Wakidah	cluster_1	1	0	59	64	160	90
4	Oth	cluster_1	1	0	59	53	160	90
5	Neni	cluster_0	1	0	59	62	110	70
6	Och	cluster_1	1	0	75	55	160	90
7	Dede Juariah	cluster_1	1	0	49	54	160	90
8	Yayah Mard...	cluster_1	1	0	51	58	160	100
9	Nenah	cluster_0	1	0	79	43	130	80
10	Nunung M	cluster_0	1	0	63	55	120	80
11	Nia Kurniawati	cluster_1	1	0	67	58	150	90
12	Halimah	cluster_1	1	0	60	64	160	90
13	Odah	cluster_1	1	0	64	44	170	90

Gambar 8. Hasil Filter Examples

Tahap kedua yaitu memasukkan operator *Remove Duplicate* untuk memfilter dan menghapus data yang sama. Dengan cara mengatur pada bagian parameter memilih "all" pada bagian *attribute filter type* maka semua data akan di analisis menggunakan operator tersebut

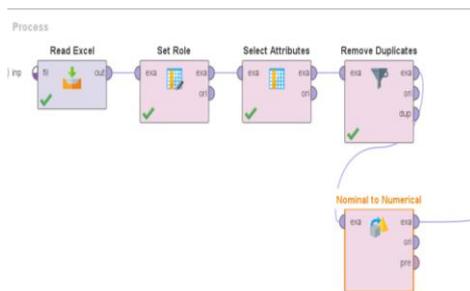


Gambar 9. operator Remove Duplicate

Setelah dimasukkan operator tersebut jumlah data tetap sama, menandakan bahwa data sudah akurat dan bersih siap untuk dilakukan *clustering k-means*.

4.3 Transformation Data

Tranformasi data digunakan untuk mengubah atribut ke format numerik. Proses ini penting karena algoritma *k-means* hanya dapat memproses data dalam bentuk numerik. Proses transformasi terdiri dari dua tahap. Pertama memasukkan operator *Nominal to Numerical* dengan atribut filter 'single' karea hanya satu yang akan diubah.



Gambar 10. Operator *Nominal to Numerical*

Setelah operator tersebut dimasukkan maka selanjutnya adalah memilih atribut yang akan diubah, disini atribut yang akan diubah yaitu dengan nama kolom jenis kelamin. Hasil dari operator tersebut dapat dilihat pada gambar dibawah ini,

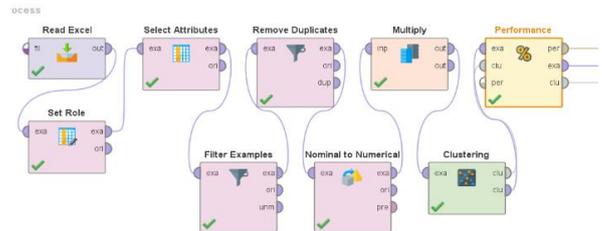
Row No.	Nama	cluster	Jenis Kela...	Jenis Kela...	Umr	Berat Badan	Sistolik	Diastolik
1	Hj Omm	cluster_1	1	0	82	54	180	130
2	Yali	cluster_0	1	0	70	45	140	80
3	Wakidah	cluster_1	1	0	59	64	160	90
4	Oth	cluster_1	1	0	59	53	160	90
5	Nani	cluster_0	1	0	59	62	110	70
6	Och	cluster_1	1	0	75	55	160	90
7	Dede Juatiah	cluster_1	1	0	49	54	160	90
8	Yayeh Mard...	cluster_1	1	0	51	58	160	100
9	Nenah	cluster_0	1	0	79	43	130	80
10	Nunung M	cluster_0	1	0	63	55	120	80
11	Nia Kumawati	cluster_1	1	0	67	58	150	90
12	Hallimah	cluster_1	1	0	60	64	160	90
13	Orleh	cluster_1	1	0	64	44	170	90

Gambar 11. Hasil Operator *Nominal to Numerical*

4.4 Data Mining

Pada proses penelitian ini menggunakan algoritma *k-means clustering* untuk mengelompokkan tekanan darah lansia berdasarkan tekanan darah sistolik dan diastolic serta menggunakan bantuan *RapidMiner*. *RapidMiner* menggunakan operator *k-means* yang dilakukan mulai dari k=2 sampai 10 kali dan operator performace dengan mengatur panel parameter sebelah kanan pada bagian *main criterion: Davies Bouildin* dan centang kotak

maximize, proses ini meliputi perhitungan nilai jarak dan nilai Davies Bouildin Index (*DBI*) sebagai metode untuk mengevaluasi kinerja sistem atau perangkat komunikasi secara efisien.



Gambar 12. Proses *K-means Clustering* di *RapidMiner*

4.5 Interpretation / Evaluasi

Dalam proses pengelompokan data, evaluasi model dilakukan untuk menentukan *cluster* terbaik. Dalam metode ini, nilai *Indeks Davies-Bouldin (DBI)* lebih rendah yang menghasilkan kualitas *cluster* yang optimal. Dalam model *k-means* yang diterapkan, dalam pemodelan ini nilai k yang bervariasi diuji mulai dari k=2 hingga k=10. Dapat dilihat pada tabel dibawah ini.

Tabel 3. Anggota *Cluster*

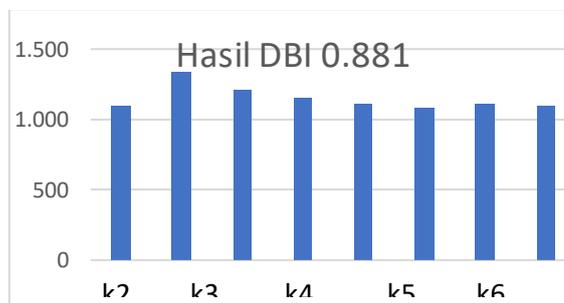
K	Anggota Cluster	DBI	Avg
2	Cluster 0: 75 items Cluster 1: 71 items Total number of items: 146	0.881	442.070
3	Cluster 0: 54 items Cluster 1: 45 items Cluster 2: 47 items Total number of items: 146	1.096	330.606
4	Cluster 0: 57 items Cluster 1: 36 items Cluster 2: 24 items Cluster 3: 29 items Total number of items: 146	1.332	292.425



5	Cluster 0: 34 items Cluster 1: 35 items Cluster 2: 19 items Cluster 3: 21 items Cluster 4: 37 items Total number of items: 146	1.211	243.187
6	Cluster 0: 42 items Cluster 1: 30 items Cluster 2: 7 items Cluster 3: 18 items Cluster 4: 20 items Cluster 5: 29 items Total number of items: 146	1.160	220.786
7	Cluster 0: 29 items Cluster 1: 23 items Cluster 2: 22 items Cluster 3: 30 items Cluster 4: 18 items Cluster 5: 3 items Cluster 6: 21 items Total number of items: 146	1.116	199.859
8	Cluster 0: 25 items Cluster 1: 8 items Cluster 2: 5 items Cluster 3: 25 items Cluster 4: 19 items Cluster 5: 23 items Cluster 6: 25 items Cluster 7: 16 items Total number of items: 146	1.079	181.488
9	Cluster 0: 20 items	1.117	164.933

	Cluster 1: 21 items Cluster 2: 14 items Cluster 3: 16 items Cluster 4: 20 items Cluster 5: 8 items Cluster 6: 22 items Cluster 7: 5 items Cluster 8: 20 items Total number of items: 146		
10	Cluster 0: 15 items Cluster 1: 29 items Cluster 2: 26 items Cluster 3: 7 items Cluster 4: 14 items Cluster 5: 11 items Cluster 6: 16 items Cluster 7: 2 items Cluster 8: 5 items Cluster 9: 21 items Total number of items: 146	1.100	161.406

Hasil evaluasi menggunakan DBI menunjukkan gambar k=2 sampai k-10.



Gambar 13. Grafik Hasil Evaluasi DBI

Hasil analisis pada pemilihan cluster terbaik menggunakan penilaian DBI menunjukkan bahwa 2 cluster lebih baik daripada jumlah cluster

lainnya sebab menghasilkan nilai *DBI* paling kecil dan mendekati 0 dengan nilai *DBI* 0.881. Hasil nilai *centeroid* yang diperoleh dari masing-masing atribut menunjukkan:

Attribute	cluster_0	cluster_1
Jenis Kelamin = P	0.880	0.831
Jenis Kelamin = L	0.120	0.169
Umur	58.213	62.352
Berat Badan	55.133	57.282
Sistolik	119.147	159.634
Diastolik	78.040	94.789

Gambar 14. Nilai Centeroid

Dibawah ini merupakan hasil ExampleSet Clustering dari hasil pengelompokan.

Row No.	Nama	cluster	Jenis Kela...	Jenis Kela...	Umur	Berat Badan	Sistolik	Diastolik
1	Hj Omm	cluster_1	1	0	82	54	180	130
2	Yati	cluster_0	1	0	70	45	140	80
3	Wakidah	cluster_1	1	0	59	64	160	90
4	Och	cluster_1	1	0	59	53	160	90
5	Neni	cluster_0	1	0	59	62	110	70
6	Och	cluster_1	1	0	75	55	160	90
7	Dede Juariah	cluster_1	1	0	49	54	160	90
8	Yayah Mard...	cluster_1	1	0	51	58	160	100
9	Nenah	cluster_0	1	0	79	43	130	80
10	Nunung M	cluster_1	1	0	63	55	120	80
11	Nia Kurniawati	cluster_1	1	0	67	58	150	90
12	Halimah	cluster_1	1	0	60	64	160	90
13	Cekah	cluster_1	1	0	64	54	170	90

Gambar 15. ExampletSet Clustering

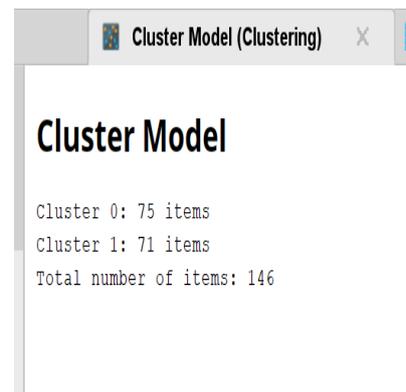
4.6 Knowledge

Hasil penelitian menggunakan algoritma *k-means clustering* menggunakan *tools RapidMiner* dan hasil evaluasi *Davies-Bouldin Index* dapat dilihat sebagai berikut:

a. Nilai Davies Bouldin Index (*DBI*)

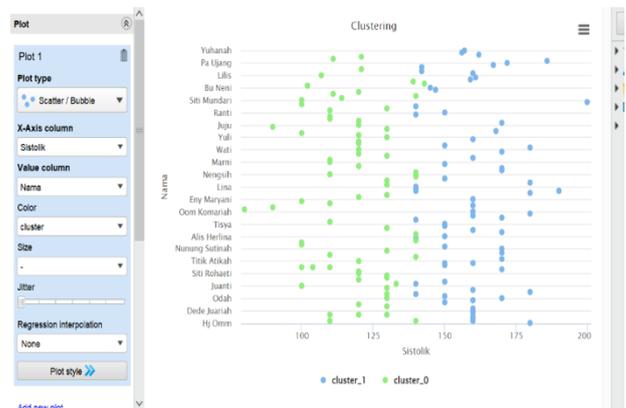
Pengelompokan dilakukan dengan penerapan algoritma *K-means clustering* menggunakan perangkat lunak *RapidMiner* dan evaluasi *DBI* diperoleh nilai *cluster* optimal dan terbaik pada $K=2$ dengan hasil nilai *DBI* terkecil yaitu 0.881 dengan anggota *Cluster_0*: 75 items dan *Cluster 1*: 71 item. Hal ini menunjukkan bahwa

pengelompokan dengan dua *cluster* menghasilkan *cluster* optimal dan memiliki Tingkat pemisahan terbaik. Dapat diamati pada gambar dibawah ini.



Gambar 16. Hasil Cluster

b. Plot



Gambar 17. Hasil Plot

Dari gambar *Plot* visualisasi diatas dapat dijelaskan bahwa grafik menunjukkan dua kelompok yang diberi warna berbeda. Pada *Cluster_0* dengan warna hijau menunjukkan cenderung dengan nilai sistolik yang lebih rendah (normal) sedangkan pada *Cluster_1* dengan warna biru yang menunjukkan kelompok dengan tekanan darah sistolik tinggi (hipertensi). Rentang tekanan darah sistolik yang terlihat pada grafik berkisar antara 100 sampai 200. Ini menunjukkan bahwa adanya variasi tekanan darah yang



signifikan diantara individu-individu dalam dataset ini.

5. KESIMPULAN DAN SARAN

Penerapan algoritma *K-means* untuk pengelompokan tekanan darah lansia di Posbindu Kp. Lebak Jero berhasil mencapai hasil optimal pada $k=2$ dengan nilai *Davies-Bouldin Index (DBI)* sebesar 0,881 setelah 10 kali percobaan. Hasil pengelompokan membagi data menjadi dua klaster, yaitu *Cluster_0* yang terdiri dari 75 lansia dengan risiko hipertensi lebih rendah, mencakup kategori normal dan pra-hipertensi, serta *Cluster_1* yang terdiri dari 71 lansia dengan risiko hipertensi lebih tinggi, mencakup hipertensi Tingkat 1 dan Tingkat 2. Model ini menunjukkan efektivitas dalam mengelompokkan lansia berdasarkan tingkat tekanan darah sistolik, meskipun variabel lain seperti jenis kelamin, umur, dan berat badan hanya memberikan sedikit perbedaan

6. UCAPAN TERIMA KASIH

Terimakasih saya ucapkan kepada semua pihak yang telah membantu dalam berjalannya proses penelitian ini. Penelitian ini tentu tidak akan berjalan tanpa ada dorongan dan motivasi dari berbagai pihak. Untuk itu, izinkan saya mengucapkan terimakasih kepada Allah Swt. , kedua orang tua, sahabat, STMIK IKMI Cirebon, Posbindu Kp.Lebak Jero serta Tim Penerbit MISI STMIK Lombok serta pihak lain yang telah memberikan dukungan dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA:

- [1] A. E. Rahayu, K. Hikmah, N. Y. Ningsih, and A. C. Fauzan, "Penerapan K-Means Clustering Untuk Penentuan Klasterisasi Beasiswa Bidikmisi Mahasiswa," *Ilk J. Comput. Sci. Appl. Informatics*, vol. 1, no. 2, pp. 82–86, 2019, doi: 10.28926/ilkomnika.v1i2.23.
- [2] M. R. Nugroho, I. E. Hendrawan, and Puwanto, "Penerapan Algoritma K-Means Untuk Klasterisasi Data Obat Pada Rumah Sakit ASRI," *J. NUANSIA Inform.*, vol. 16, no. 1, pp. 125–133, 2022, [Online]. Available: <https://journal.uniku.ac.id/index.php/ilkom>
- [3] A. Bahauddin, A. Fatmawati, and F. P. Sari, "Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means," *J. Manaj. Inform. dan Sist. Inf.*, vol. 4, no. 1, pp. 1–8, 2021, doi: 10.36595/misi.v4i1.216.
- [4] D. S. Qirom, A. Faqih, and G. Dwilestari, "Implementasi Algoritma K-Means Untuk Klasterisasi Pasien Hipertensi Berdasarkan Karakteristik Pasien," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 2056–2063, 2024, doi: 10.36040/jati.v8i2.8314.
- [5] C. A. Sugianto, A. H. Rahayu, and A. Gusman, "Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Cigugur Tengah," *J. Inf. Technol.*, vol. 2, no. 2, pp. 39–44, 2020, doi: 10.47292/joint.v2i2.30.
- [6] H. A. Arbi and R. A. Putri, "Visualisasi Data Pemetaan Daerah Hipertensi Menggunakan Algoritma K-Means," *J. Teknol. Sist. Inf. dan Apl.*, vol. 6, no. 4, pp. 631–638, 2023, doi: 10.32493/jtsi.v6i4.33923.
- [7] Saikin and Kusri, "Model Data Mining Untuk Karakteristik Data Traveller Pada Perusahaan Tour and Travel," *J. Manaj. Inform. dan Sist. Inf.*, vol. 2, no. 2, pp. 61–68, 2019, doi: 10.36595/misi.v2i2.105.
- [8] A. Hidayah, W. Lestari, and E. Purwanto, "Implementasi Data Mining Algoritma K-Means untuk Clustering Penyakit di RS Panti Waluyo Surakarta," *DutaCom*, vol. 15, no. 2, pp. 80–90, 2022, doi: 10.47701/dutacom.v15i2.2009.
- [9] Y. Tis Asy Aria, M. Julkarnain, and F. Hamdani, "Penerapan Algoritma K-Means Clustering Untuk Data Obat," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 1, pp. 649–657, 2023, doi: 10.30865/klik.v4i1.1117.
- [10] W. Utomo, "Penerapan Algoritma K-Means Untuk Pengelompokan Penyakit Kronis pada Warga Lansia (Studi Kasus Pada: Posyandu Lansia RW 07) Wargijono," *J. Media Inform. Budidarma*, vol. 4, no. 4, pp. 1153–1161, 2020, doi: 10.30865/mib.v4i4.2410.
- [11] Azkar and Kusri, "Klasterisasi Pasien Rawat Jalan di Puskesmas dengan Menggunakan Metode Algoritma Clustering K-Means," *KLIK Kaji. Ilm.*



- Inform. dan Komput.*, vol. 4, no. 5, pp. 2628–2636, 2024, doi: 10.30865/klik.v4i5.1832.
- [12] F. Rahmadayanti, I. Anggraini, and T. Susanti, “Pengklasterisasian Data Penyakit Hipertensi dengan Menggunakan Metode K-Means,” *J. Inf. Syst. Res.*, vol. 4, no. 2, pp. 737–741, 2023, doi: 10.47065/josh.v4i2.2905.
- [13] M. Herviany, S. P. Delima, T. Nurhidayah, and Kasini, “Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Daerah Rawan Tanah Longsor Pada Provinsi Jawa Barat,” *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 34–40, 2021, doi: 10.57152/malcom.v1i1.60.
- [14] U. Qalsum and W. Abidin, “Klasifikasi Penyakit Hipertensi Menggunakan Metode K-Means Clustering,” *J. MSA (Mat. dan Stat. serta Apl.*, vol. 11, no. 2, pp. 124–128, 2024, doi: 10.24252/msa.v11i2.45291.
- [15] H. Dilawati, H. Widiyanto, and A. Kuswiadji, “Klasterisasi Data Rekam Medis Pasien Menggunakan Metode K-Means Clustering Di Rumah Sakit Widodo Ngawi,” *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 5, no. 2, pp. 139–147, 2024, [Online]. Available: doi: <https://doi.org/10.37148/bios.v5i2.134>
- [16] I. Soliani and S. Juanita, “Grouping the Prevalence of Disease Cases By Age in Bandung City Hospitals Using K-Means,” *J. Tek. Inform.*, vol. 3, no. 6, pp. 1647–1654, 2022, doi: 10.20884/1.jutif.2022.3.6.430.
- [17] D. S. Saputri, G. M. Putra, and M. F. Larasati, “Implementation of the K-Means Clustering Algorithm for the Covid-19 Vaccinated Village in the Ujung Padang Sub-District,” *J. Tek. Inform.*, vol. 3, no. 2, pp. 261–267, 2022, [Online]. Available: <https://doi.org/10.20884/1.jutif.2022.3.2.165>
- [18] I. Setiawan, “Knowledge Discovery In Databases (KDD) Terhadap Customer Reviews Pada Situs E-Commerce,” 2018. [Online]. Available: [http://edocs.ilkom.unsri.ac.id/2885/1/Knowledge Discovery in Databases - 09031281621045.docx](http://edocs.ilkom.unsri.ac.id/2885/1/Knowledge%20Discovery%20in%20Databases%20-%2009031281621045.docx)