

IDENTIFIKASI PEMILIHAN OBAT BAGI PASIEN HIPERTENSI DAN HIPERKOLESTEROLEMIA DENGAN METODE DESICION TREE

Soffin Thoriq Arfian*¹, Khayruraya Abrar J², Azfa Yashifa R³, Farhan Naufal M⁴,
Rajnaparamitha Kusumastuti⁵

¹²³⁴⁵Prodi S1 Informatika, STMIK Amikom Surakarta
¹²³⁴⁵Sukoharjo, Indonesia

Email: ¹soffin.10406@mhs.amikomsolo.ac.id, ²khayruraya.10400@mhs.amikomsolo.ac.id, ³azfa.10389@mhs.amikomsolo.ac.id, ⁴farhan.10397@mhs.amikomsolo.ac.id, ⁵rajna@dosen.amikomsolo.ac.id

Abstract

Hypertension also known as high blood pressure, is a medical condition in which arterial blood pressure is consistently lower than normal. Hypercholesterolemia, also known as high cholesterol, is a condition where the amount of cholesterol in the blood is too high. These two conditions often occur together and require appropriate drug management to reduce cardiovascular risk. This study aims to identify variables that influence drug use by patients with high blood pressure and cholesterol using the Desicion Tree classification method. The Desicion Tree classification method was used in the study to understand whether treatment might be appropriate for patients with the same condition in the future. The dataset consists of: Age, Gender, Blood Pressure, and cholesterol, sodium-potassium, and medication. The use of the Desicion Tree algorithm with RapidMiner shows an accuracy of 99%, which proves the effectiveness of the method in providing assistance in making higher decisions regarding determining drugs for patients. It is hoped that this research can help doctors in making decisions regarding drug selection for patients.

Keywords: High blood pressure, Cholesterol, Classification, Desicion Tree, RapidMiner

Abstraksi

Hipertensi juga dikenal sebagai tinggi darah, adalah suatu kondisi medis di mana tekanan darah arteri secara konsisten lebih rendah dari biasanya. hiperkolesterolemia juga Dikenal dengan sebutan kolesterol tinggi, merupakan suatu kondisi dimana jumlah kolesterol dalam darah terlalu tinggi. Kedua kondisi ini sering terjadi bersamaan dan memerlukan pengelolaan obat yang tepat untuk mengurangi resiko kardiovaskular. Penelitian ini bertujuan untuk mengidentifikasi variabel -variabel yang memengaruhi penggunaan obat oleh pasien dengan kondisi tekanan darah tinggi dan kolesterol menggunakan metode klasifikasi Desicion Tree. Metode klasifikasi Desicion Tree digunakan dalam penelitian Untuk memahami apakah pengobatan mungkin tepat untuk pasien dengan kondisi yang sama di kemudian hari. Dataset tersebut terdiri dari: Usia, Jenis kelamin, Tekanan Darah, dan kolesterol, natrium-kalium, dan obat. Penggunaan algoritma Desicion Tree dengan RapidMiner menunjukkan akurasi sebesar 99% yang membuktikan efektivitas metode dalam memberi bantuan dalam membuat keputusan yang lebih tinggi terkait penentuan obat untuk pasien. Penelitian ini diharapkan dapat membantu dokter dalam memilih membuat keputusan terkait pemilihan obat kepada pasien.

Kata Kunci: Tekanan darah tinggi, Kolesterol, Klasifikasi, Desicion Tree, RapidMiner

1. PENDAHULUAN

Kesehatan merupakan salah satu aset yang paling berharga bagi setiap individu. Menjaga kesehatan berarti melakukan tindakan pencegahan dan penanganan berbagai kondisi medis yang dapat mempengaruhi kualitas hidup. Dua kondisi medis yang sering muncul secara bersamaan dan memerlukan perhatian khusus adalah hipertensi (tekanan darah tinggi) dan hiperkolesterolemia (kolesterol tinggi).

Hipertensi dikenal sebagai salah satu penyebab utama kematian di seluruh dunia, sering disebut sebagai "*the silent killer*" karena sifatnya yang tanpa gejala namun mematikan. [1]. Kondisi ini dapat menyebabkan komplikasi serius seperti penyakit jantung, stroke, dan gagal ginjal jika tidak dikelola dengan baik. Sementara itu, kolesterol tinggi merupakan kondisi medis di mana kadar kolesterol dalam darah melebihi batas normal. Kolesterol tinggi dapat memicu penyakit jantung koroner karena menyebabkan penyumbatan pada pembuluh darah yang mengurangi aliran darah ke jantung [2].

Untuk mengelola kedua kondisi tersebut sering kali menjadi tantangan, karena melibatkan berbagai faktor seperti usia pasien, tekanan darah, kadar kolesterol, dan keseimbangan elektrolit dalam tubuh. Dalam praktiknya, keputusan pengobatan sering kali didasarkan pada pengalaman subjektif dokter atau panduan umum, yang belum tentu sesuai untuk semua pasien. Oleh karena itu, diperlukan pendekatan berbasis data untuk membantu dokter membuat keputusan yang lebih akurat dan objektif.

Teknologi *data mining* adalah proses yang digunakan untuk mengekstrak informasi baru dari sejumlah besar data yang sebelumnya tidak mungkin dipahami secara manual. Tujuan dari proses ini adalah menganalisis pola data yang ada agar dapat diubah menjadi informasi yang lebih berguna. Informasi tersebut diperoleh melalui proses analisis pola yang penting atau berdasarkan data yang terdapat pada kumpulan data tersebut [3]. Salah satu teknik penambangan data yang terkenal adalah *Decision Tree*, yang merupakan metode klasifikasi dan prediksi yang efektif. [4].

Decision tree adalah metode klasifikasi dan prediksi yang sangat efektif dan terkenal. Metode *Decision Tree* mengubah data yang kompleks menjadi pohon keputusan yang merepresentasikan aturan-aturan [5]. Keunggulan dari *Decision Tree* adalah kemampuannya dalam memvisualisasikan keterkaitan antar variabel, sehingga memudahkan pengambilan keputusan dalam situasi data yang kompleks [6].

Identifikasi faktor-faktor utama yang memengaruhi pemilihan obat pada pasien dengan hipertensi dan hiperkolesterolemia, menggunakan algoritma *Decision Tree*. Dengan memanfaatkan *dataset* dari Kaggle, penelitian ini juga bertujuan untuk mengembangkan model prediksi yang dapat membantu dokter dalam menentukan obat yang tepat berdasarkan atribut pasien. Penelitian ini memberikan kontribusi baik secara teoretis, dengan memperluas penerapan *Decision Tree* dalam analisis data medis, maupun secara praktis, dengan menyediakan alat bantu bagi pengambilan keputusan klinis.

2. TINJAUAN PUSTAKA

Hipertensi sering dikenal sebagai "*silent killer*" karena penderita sering kali tidak menunjukkan gejala selama bertahun-tahun, meskipun kondisi tersebut dapat menyebabkan komplikasi serius pada organ-organ vital, seperti jantung, otak, dan ginjal. Gejala seperti pusing, gangguan penglihatan, serta sakit kepala biasanya muncul ketika hipertensi sudah berkembang lebih lanjut dan tekanan darah mencapai tingkat berbahaya[7].

Sementara itu, kolesterol tinggi mengacu pada kondisi di mana kadar kolesterol berlebihan dalam darah cenderung menempel pada dinding bagian dalam pembuluh darah. Kelebihan LDL yang mengalami oksidasi dapat membentuk gumpalan, yang semakin membesar dan berakhir pada penyempitan pembuluh darah, dalam proses yang dikenal sebagai aterosklerosis. Kadar kolesterol yang tinggi ini merupakan masalah serius, karena menjadi salah satu faktor risiko utama berbagai penyakit tidak menular, seperti penyakit jantung, stroke, dan diabetes melitus[8].

Data mining adalah teknik yang digunakan untuk mengekstraksi pengetahuan tersembunyi dari data secara otomatis atau semi-otomatis. Dalam bidang kesehatan, data mining bermanfaat untuk mengungkap informasi penting, seperti pola penyakit, faktor risiko, dan hubungan antar variabel[9].

Pada penelitian yang dilakukan[10] mengartikan Decision tree sebagai algoritma pembelajaran mesin yang diawasi yang digunakan untuk menyelesaikan masalah klasifikasi. Algoritma ini bertujuan untuk menghasilkan model prediktif yang terstruktur dalam bentuk aturan-aturan yang jelas dan mudah diimplementasikan.

Dalam penelitian yang dilakukan oleh[11], Decision Tree telah digunakan untuk memprediksi hipertensi berdasarkan data klinis pasien, di penelitian ini penelitian membandingkan performanya dengan algoritma lain seperti K-Nearest Neighbor (KNN), Naive Bayes, dan Artificial Neural Network memiliki tingkat performa dengan nilai akurasi sebesar 94.7%, recall sebesar 91.5% dan precision sebesar 97.7%.

Pada penelitian yang dilakukan[12] bertujuan mengembangkan aplikasi berbasis kecerdasan buatan untuk membantu diagnosis hipertensi dengan akurasi optimal. Menggunakan metode decision tree dan random forest, keduanya mencapai akurasi 100%, menunjukkan efektivitas metode ini dalam mendukung diagnosis hipertensi.

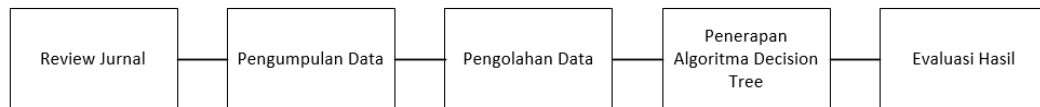
Pada penelitian yang dilakukan[13] Metode yang digunakan dalam kasus ini adalah algoritma Decision Tree, yang memberikan hasil akurasi tinggi saat menggunakan rasio data 80:20, dengan akurasi mencapai 98,71%. Selain itu, nilai Presisi, Recall, dan F1-Score masing-masing adalah 0,9872, 0,9872, dan 0,9867. Dengan rasio data 70:30, meskipun hasilnya sedikit lebih rendah, tetap menunjukkan performa yang baik dengan akurasi 98,28%, Presisi 0,9832, Recall 0,9828, dan F1-Score 0,9804.

Penelitian sebelumnya telah menggunakan berbagai algoritma, seperti Naïve Bayes dan KNN, untuk memprediksi hipertensi dan kolesterol. Namun, belum ada penelitian yang secara spesifik mengintegrasikan atribut rasio natrium-kalium (Na-Ka) dengan tekanan darah sebagai dasar dalam menentukan pemilihan obat. Penelitian ini

bertujuan untuk mengisi celah tersebut dengan menggunakan algoritma Decision Tree, yang memungkinkan analisis lebih mendalam dan keputusan berbasis data untuk membantu pemilihan obat yang lebih tepat.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimen, yang bertujuan untuk menguji efektivitas algoritma Decision Tree dalam menentukan jenis obat yang sesuai untuk pasien dengan hipertensi dan hiperkolesterolemia. Penelitian dilakukan menggunakan dataset sekunder yang diunduh dari platform Kaggle, yang berisi informasi medis pasien dan jenis obat yang diberikan. Dataset ini dianalisis dengan aplikasi RapidMiner untuk membangun model klasifikasi berbasis algoritma Decision Tree. Langkah penelitian meliputi Review Jurnal, pengumpulan data, Pengolahan data, penerapan algoritma decision tree dan Evaluasi Hasil, seperti gambar 1 berikut.



Gambar 1. Alur Pemrosesan Data

3.1. Review Jurnal

Pada tahap ini, dilakukan review jurnal untuk memahami temuan, metodologi, dan teori yang telah ada dalam bidang ini. Review jurnal juga bertujuan untuk mengetahui tren, kekurangan, dan peluang penelitian baru dalam bidang pemilihan obat untuk pasien hipertensi dan hiperkolesterolemia.

3.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari Kaggle Obat A, B, C, X, Y untuk Desicion Tree yang berisi 200 informasi data pasien, kondisi kesehatan, dan obat-obatan yang digunakan. Data ini mencakup atribut seperti Usia, Jenis kelamin, Tekanan Darah, dan Kolesterol, Natrium-Kalium, Obat, seperti yang di jelaskan pada tabel 1 berikut.

Tabel 1. Dataset Atribut

Usia	Usia Pasien (15 age- 74 age)
Seks	Jenis Kelamin Pasien (Male, Female)
BP	Tekanan Darah (High, Low)
Kolesterol	Tingkat Kolesterol (High, Normal)
Na_to_Ka	Natrium - Kalium
Obat	Obat yang bekerja pada pasien (Drug Y, Drug X, Other)

Data di atas adalah tabel dari dataset medis yang digunakan untuk menganalisis variabel-variabel yang mempengaruhi pemilihan obat untuk pasien. Berikut adalah penjelasan rinci dari masing-masing variabel dalam dataset tersebut:

1. Usia

Variabel yang menunjukkan usia pasien dalam rentang 15 hingga 74 tahun. Usia dapat mempengaruhi respons pasien terhadap pengobatan tertentu karena perubahan fisiologis yang terjadi seiring bertambahnya usia. Dalam analisis, usia bisa menjadi prediktor penting untuk menentukan jenis obat yang paling efektif untuk pasien.

2. Seks (Jenis Kelamin)

Variabel ini menunjukkan jenis kelamin pasien, dengan dua kategori yaitu Male (laki-laki) dan Female (perempuan). Jenis kelamin bisa mempengaruhi metabolisme obat, dosis yang diperlukan, dan kemungkinan efek samping. Oleh karena itu, penting untuk mempertimbangkan jenis kelamin dalam pemilihan obat yang sesuai.

3. BP (Tekanan Darah)

Variabel yang menunjukkan status tekanan darah pasien, dengan dua kategori yaitu High (tinggi) dan Low (rendah). Tekanan darah tinggi atau rendah dapat mempengaruhi keputusan tentang obat apa yang harus diberikan. Misalnya, pasien dengan tekanan darah tinggi mungkin memerlukan obat yang berbeda dari pasien dengan tekanan darah rendah untuk menghindari risiko komplikasi lebih lanjut.

4. Kolesterol

Variabel ini menunjukkan tingkat kolesterol pasien, dengan dua kategori yaitu High (tinggi) dan Normal. Kolesterol tinggi merupakan faktor risiko untuk berbagai kondisi kardiovaskular dan bisa mempengaruhi pemilihan obat. Dokter mungkin akan memilih obat yang juga bisa membantu menurunkan kolesterol atau setidaknya tidak memperburuk kondisi kolesterol pasien.

5. Na_to_Ka (Natrium - Kalium)

Variabel ini menunjukkan rasio antara natrium dan kalium dalam darah pasien. Rasio ini bisa memberikan informasi tentang keseimbangan elektrolit dalam tubuh, yang penting dalam pengelolaan beberapa kondisi medis. Rasio ini bisa mempengaruhi efektivitas dan pilihan obat tertentu.

6. Obat (Drug)

Variabel target yang menunjukkan obat yang bekerja pada pasien, dengan kategori yang berbeda seperti Drug Y, Drug X, dan Other. Variabel ini adalah output dari model klasifikasi, di mana berdasarkan input dari variabel lainnya (usia, jenis kelamin, tekanan darah, kolesterol, dan rasio Na/K), model mencoba memprediksi obat mana yang paling efektif untuk pasien.

Variabel-variabel di atas digunakan dalam model prediksi untuk menentukan obat yang paling tepat untuk setiap pasien berdasarkan karakteristik individual mereka. Dengan menggunakan teknik klasifikasi, model dapat memprediksi obat yang efektif berdasarkan pola yang terlihat dari variabel input (usia, jenis kelamin, tekanan darah, kolesterol, dan Na_to_Ka).

3.3. Pengolahan Data

Data yang digunakan dari dataset yang didapat dari sumber kaggle, berjumlah 200 dataset di bersikan menggunakan aplikasi excel, seperti digambarkan di tabel 2 berikut.

Tabel 2. Dataset Obat A, B, C, X, Y untuk Desicion Tree

Age	Sex	BP	Cholesterol	Na_to_K	Drug
23	F	HIGH	HIGH	25.355	drugY
47	M	LOW	HIGH	13.093	drugC
47	M	LOW	HIGH	10.114	drugC
28	F	NORMAL	HIGH	7.798	drugX
61	F	LOW	HIGH	18.043	drugY
22	F	NORMAL	HIGH	8.607	drugX
49	F	NORMAL	HIGH	16.275	drugY
41	M	LOW	HIGH	11.037	drugC
60	M	NORMAL	HIGH	15.171	drugY
43	M	LOW	NORMAL	19.368	drugY
47	F	LOW	HIGH	11.767	drugC
34	F	HIGH	NORMAL	19.199	drugY
43	M	LOW	HIGH	15.376	drugY
74	F	LOW	HIGH	20.942	drugY

3.4. Penerapan Algoritma Desicion Tree

Pada tahap ini, kami menggunakan aplikasi RapidManer dengan menggunakan algoritma decision tree yang diterapkan dengan mempertimbangkan teknik klasifikasi, stratified sampling, dan gain ratio.

a. Pembagian Data dengan Stratified Sampling

Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Teknik stratified sampling digunakan untuk Memastikan bahwa pembagian data pelatihan dan pengujian mempertahankan distribusi kelas yang sama seperti dalam dataset asli. Teknik Stratified sampling memastikan bahwa proporsi setiap kelas (misalnya, jenis obat) dalam dataset pelatihan dan pengujian konsisten dengan proporsi di dataset keseluruhan. Membantu dalam menghindari bias dan memastikan bahwa model dapat belajar dari semua kelas dengan representasi yang seimbang.

b. Pelatihan Model dengan Klasifikasi

Model decision tree dipakai dengan menggunakan data yang telah dibagi. Dalam konteks klasifikasi, model ini akan mempelajari bagaimana atribut seperti usia, jenis kelamin, tekanan darah, kolesterol, natrium-kalium, dan lainnya

mempengaruhi pemilihan obat. Decision tree akan membangun aturan berdasarkan atribut tersebut untuk mengklasifikasikan keputusan pemilihan obat.

c. Penggunaan Gain Ratio

Gain Ratio digunakan sebagai kriteria pembagian untuk menentukan atribut yang paling informatif dalam klasifikasi. Teknik ini memastikan bahwa atribut yang dipilih memberikan kontribusi maksimal terhadap hasil klasifikasi.

d. Evaluasi Model

Model kinerja Decision Tree dievaluasi menggunakan beberapa metrik berikut:

- **Akurasi** : Persentase prediksi yang benar dari total data pengujian. Memberikan gambaran umum tentang model kemampuan dalam mengklasifikasikan data yang tidak terlihat sebelumnya.
- **Presisi dan Recall** : Mengukur kinerja model untuk setiap kelas target (jenis obat). Presisi mengukur akurasi prediksi untuk setiap kelas, sedangkan recall mengukur kemampuan model untuk menemukan semua instansi kelas tersebut.
- **Kappa** : Menilai kesepakatan antara model prediksi dan label aktual, dengan korelasi kemungkinan kesepakatan acak. Kappa memberikan gambaran lebih jelas tentang model kekuatan kesepakatan.

e. Visualisasi Pohon Keputusan

Memvisualisasikan Pohon keputusan untuk memudahkan pemahaman keputusan yang diambil. Visualisasi ini akan memperlihatkan bagaimana atribut dan nilai-nilainya mempengaruhi hasil keputusan pemilihan obat

3.5. Evaluasi Hasil

Proses klasifikasi dilakukan dengan menggunakan perangkat lunak Rapid Miner, dengan menggunakan algoritma Decision Tree. Data yang sudah berhasil diproses digunakan untuk hasil evaluasi dari menggunakan algoritma Decision Tree adalah berdasarkan akurasi dan kappa.

4. HASIL DAN PEMBAHASAN

4.1. Proses Pengolahan Data

Penelitian ini bertujuan untuk mengimplementasikan algoritma Decision Tree dalam memprediksi pemilihan obat yang tepat untuk pasien berdasarkan beberapa atribut kesehatan. Data yang digunakan mencakup atribut seperti usia , jenis kelamin , tekanan darah , kadar Kolesterol , dan rasio natrium terhadap kalium .

Proses pengolahan data dimulai dengan mengkonversi dataset ke dalam format Excel, kemudian mengimpornya ke dalam RapidMiner untuk analisis lebih lanjut. Pada tahap ini, data disiapkan dengan benar untuk memastikan bahwa atribut drug diubah menjadi label, yang menandakan obat yang diprediksi untuk pasien. Validasi model dilakukan menggunakan metode stratified sampling untuk memastikan distribusi yang seimbang di setiap kelas. Keputusan pohon yang dihasilkan dengan parameter yang

disesuaikan untuk mencapai akurasi model yang maksimal. Berikut adalah hasil dari pemrosesan data:

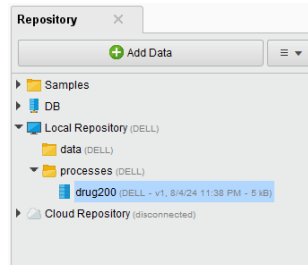
4.1.1. Perhitungan Rapid Minner

Proses perhitungan diawali dengan menginput dataset Obat A, B, C, X, Y untuk Pohon Keputusan ke dalam RapidMiner. Setelah itu masukan dataset, proses mengolah dataset agar mempermudah proses perhitungan. Pengelolaan dataset Age (C1), Sex (C2), BP (C3), Cholesterol (C4), Na_to_Ka (C5) dan memberikan label pada variabel class. Pertama, ubah format dataset ke excel seperti pada gambar 2.

Age	Sex	BP	Cholesterol	Na_to_Ka	Drug
23	F	HIGH	HIGH	25.355	drugY
47	M	LOW	HIGH	13.093	drugC
47	M	LOW	HIGH	10.114	drugC
28	F	NORMAL	HIGH	7.798	drugX
61	F	LOW	HIGH	18.043	drugY
22	F	NORMAL	HIGH	8.607	drugX
49	F	NORMAL	HIGH	16.275	drugY
41	M	LOW	HIGH	11.037	drugC
60	M	NORMAL	HIGH	15.171	drugY
43	M	LOW	NORMAL	19.368	drugY
47	F	LOW	HIGH	11.767	drugC
34	F	HIGH	NORMAL	19.199	drugY
43	M	LOW	HIGH	15.376	drugY
74	F	LOW	HIGH	20.942	drugY
50	F	NORMAL	HIGH	12.703	drugX
16	F	HIGH	NORMAL	15.516	drugY
69	M	LOW	NORMAL	11.455	drugX

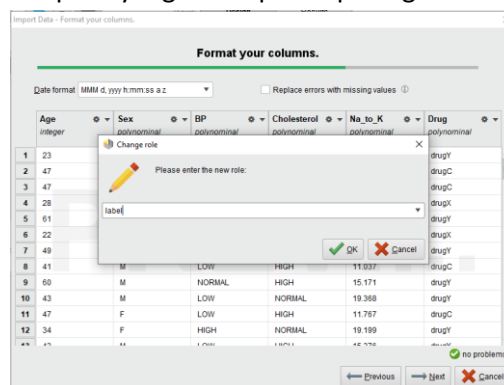
Gambar 2. Dataset dalam format excel

Selanjutnya, temukan data yang telah Anda miliki dalam format tabel XLS. Lakukan proses Importing Data ke dalam aplikasi RapidMiner dengan mencari dataset yang sudah diubah ke format Excel. Setelah itu, masukkan dataset tersebut ke dalam Local Repository, seperti yang ditunjukkan pada Gambar 3.



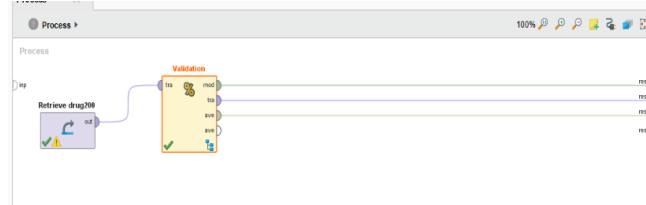
Gambar 3. Repository

Sebelumnya jangan lupa ubah role pada atribut drug dengan memilih change rolee dan ubah ke label. Seperti yang di tampilkan pada gambar 4.



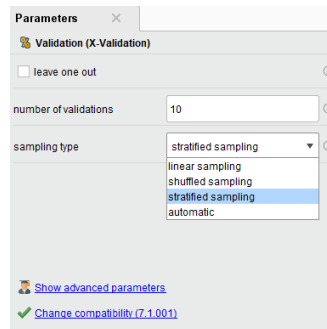
Gambar 4. Change role

Selanjutnya drag dataset Seperti yang di tampilkan pada gambar 5. Setelah itu buat validasi dengan mencari di operator pada validation dan pilih x-validation. Selanjutnya hubungkan dengan menarik garis dari tabel drug200 ke validation, selanjutnya untuk validation juga hubungkan dengan menarik garis ke result di sisi kanan.



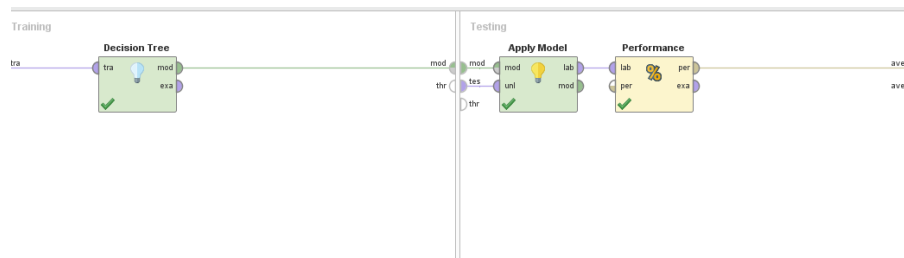
Gambar 5. Susunan data drug dan validation

Pada parameter validation ubah sampling type ke stratified sampling. Seperti yang di tampilkan pada gambar 6.



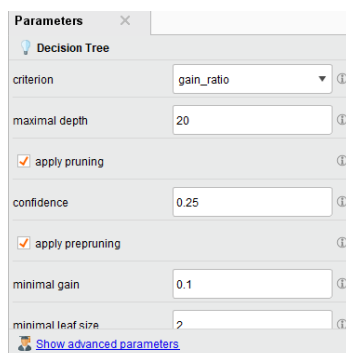
Gambar 6. Parameters validation

Selanjutnya atur pada validation dengan klit validadion 2x, sama buat desicion tree di operator dengan pilih modeling terus pilih predictive setelah itu pilih trees pilih desicion tree. Dan juga buat apply model di confidences, dan performance di segmentation. Dan hubung kan Seperti yang di tampilkan pada gambar 7.



Gambar 7. Susunan desicion tree, apply model, dan performance

Pada parameters desicion tree buat criterion gain_ratio, maximal depth 20, pada apply pruning confidence 0.25 dan apply prepruning minimail gain 0.1, minimal leaf size 2. Seperti yang di tampilkan pada gambar 8.



Gambar 8. Parameters desicion tree

Setelah mengatur parameter untuk setiap operator dan menyusun posisi operator dengan benar, klik ikon **Run**. Tunggu hingga RapidMiner menyelesaikan prosesnya dan menampilkan hasil keputusan decision tree, yang akan berupa tabel akurasi dan grafik pohon.

4.1.2. Hasil Perhitungan Menggunakan RapidMiner

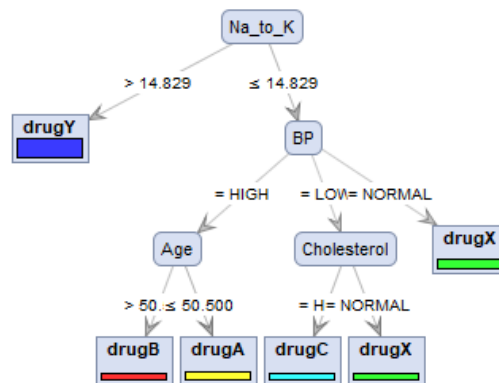
Berdasarkan pohon keputusan menghasilkan aturan prediksi yang jelas. Aturan ini menunjukkan hubungan antara atribut-atribut seperti Na_to_K, BP, Cholesterol, dan Age dengan pemilihan obat.

- Root Node (Node Akar)
 - Na_to_K > 14.829: rasio Natrium ke Kalium lebih dari 14.829 menunjukkan bahwa pasien mungkin mengalami ke elektrolit yang signifikan, yang membutuhkan obat DrugY.
- Na_to_K ≤ 14.829

Jika rasio natrium terhadap kalium lebih rendah, maka faktor-faktor lain seperti tekanan darah dan usia pasien menjadi lebih penting dalam menentukan jenis obat yang diperlukan.

- BP (Blood Pressure / Tekanan Darah)
 - BP = HIGH (Tinggi)
 - Age > 50.5: Pasien dengan tekanan darah tinggi dan usia lebih dari 50.5 tahun kemungkinan memiliki risiko komplikasi lebih tinggi, sehingga memerlukan obat DrugB. Obat ini dirancang untuk pasien usia lanjut dengan hipertensi.
 - Age ≤ 50.5: Pasien yang lebih muda dengan tekanan darah tinggi diprediksi akan menerima DrugA, yang mungkin lebih cocok untuk kelompok usia ini karena faktor metabolisme atau toleransi obat.
- BP = LOW (Rendah)
- Cholesterol = HIGH (Tinggi): Jika tekanan darah rendah tetapi kadar Kolesterol tinggi, obat DrugC diberikan. Obat ini dirancang untuk mengatasi kolesterol tinggi sambil mempertahankan tekanan darah pada tingkat yang aman.

- Cholesterol = NORMAL: Jika tekanan darah rendah dan kadar Kolesterol normal, obat DrugX dipilih. Ini menunjukkan bahwa obat ini aman untuk pasien tanpa kolesterol yang signifikan.
 - BP = NORMAL: Pasien dengan tekanan darah normal diprediksi menerima DrugX, yang mungkin merupakan obat generik atau standar untuk kondisi yang tidak memiliki tekanan darah sebagai faktor risiko utama
- Gambar hasil decision tree ditampilkan pada gambar 9 berikut.



Gambar 9. Desicion Tree

Model ini menunjukkan pola prediksi yang jelas dan logistik berdasarkan atribut kesehatan pasien. Root Node Na_to_K menunjukkan bahwa rasio natrium terhadap kalium memiliki pengaruh paling besar dalam menentukan jenis obat yang direkomendasikan. Pasien dengan rasio Na_to_K > 14.829 diprediksi menerima DrugY, menunjukkan bahwa rasio yang tinggi memerlukan penanganan spesifik dengan obat ini.

Pada cabang Na_to_K ≤ 14.829, model melanjutkan dengan pemeriksaan atribut tekanan darah (BP) dan usia pasien, yang memberikan informasi tambahan untuk menentukan jenis obat:

- Pasien dengan tekanan darah tinggi (BP = TINGGI) dipisahkan lebih lanjut berdasarkan usia, karena risiko komplikasi medis sering meningkat pada pasien yang lebih tua.
- Pasien dengan tekanan darah rendah (BP = LOW) dipengaruhi oleh kadar Kolesterol, menunjukkan bahwa Kolesterol berperan penting dalam pengelolaan tekanan darah rendah.
- Pasien dengan tekanan darah normal (BP = NORMAL) cenderung mendapatkan DrugX, menunjukkan obat ini cocok untuk kondisi tanpa risiko tekanan darah ekstrem.

Model ini menunjukkan tingkat presisi dan sensitivitas yang tinggi, terutama dalam memanfaatkan atribut yang saling berhubungan seperti Na_to_K, BP, Cholesterol, dan Age untuk menghasilkan prediksi yang akurat. Dengan tingkat kesalahan yang rendah dan pembagian yang logis, model ini sangat dapat diandalkan

untuk digunakan dalam konteks klinis untuk membantu dokter dalam membuat keputusan pengobatan yang lebih tepat dan cepat.

4.2. Hasil Confusion Matrix

Hasil penelitian menggunakan rapid minner menggunakan decision tree menghasilkan hasil confusion matrix seperti yang di jelaskan di Gambar tabel accuracy ditampilkan pada gambar 10 berikut.

accuracy: 99.00% +/- 3.00% (mikro: 99.00%)

	true drugY	true drugC	true drugX	true drugA	true drugB	class precision
pred. drugY	91	0	1	0	0	98.91%
pred. drugC	0	16	0	0	0	100.00%
pred. drugX	0	0	53	0	0	100.00%
pred. drugA	0	0	0	23	1	95.83%
pred. drugB	0	0	0	0	15	100.00%
class recall	100.00%	100.00%	98.15%	100.00%	93.75%	

Gambar 10. Tabel accuracy

- Accuracy : 99.00% +/-3.00% (mikro : 99.00%)

Menunjukkan model memiliki akurasi sebesar 99%, dengan kemungkinan variasi sebesar +/- 1.3%, yang menunjukkan bahwa Decision Tree sangat akurat dalam memprediksi pemilihan obat berdasarkan dataset yang diberikan. Confusion Matrix yang dihasilkan menunjukkan hasil sebagai berikut:

- Confusion Matrix :
 - DrugY : 91 prediksi benar, 1 prediksi salah (diprediksi sebagai drugX)
 - DrugC : 16 prediksi benar, 0 prediksi salah
 - DrugX : 53 prediksi benar, 0 prediksi salah
 - DrugA : 23 prediksi benar, 1 prediksi salah (diprediksi sebagai drugB)
 - DrugB : 15 prediksi benar, 0 prediksi salah
- Class Precision : Tingkat presisi masing – masing kelas :
 - DugY : 98.91%
 - DrugC : 100.00%
 - DrugX : 100.00%
 - DrugA : 95.83%
 - DrugB : 100.00%
- Class Recall : Tingkat recall masing - masing kelas :
 - DrugY : 100.00%
 - DrugC : 100.00%
 - DrugX : 98.15%
 - DrugA : 100.00%
 - DrugB : 93.75%

Model ini menunjukkan performa yang sangat baik dengan tingkat kesalahan rendah, sebagaimana terlihat pada Confusion Matrix , di mana sebagian besar prediksi

tepat untuk semua kelas obat. Tingkat presisi yang tinggi (98.91% untuk DrugY dan 100% untuk DrugC , DrugX , serta DrugB) menunjukkan bahwa model sangat akurat dalam memprediksi kelas tanpa menghasilkan prediksi positif palsu. Sementara itu, tingkat recall yang tinggi (100% untuk sebagian besar kelas) menunjukkan kemampuan model dalam mendeteksi semua kasus secara akurat, meskipun sedikit lebih rendah pada DrugB (93.75%). Kombinasi presisi dan sensitivitas yang tinggi memastikan bahwa model ini andal dan efektif untuk digunakan dalam konteks klinis, memberikan rekomendasi pengobatan yang tepat berdasarkan atribut pasien dengan tingkat keyakinan yang sangat tinggi. Gambar tabel kappa ditampilkan pada gambar 11 berikut.

kappa: 0.985 +/- 0.044 (mikro: 0.986)

	true drugY	true drugC	true drugX	true drugA	true drugB	class precision
pred. drugY	91	0	1	0	0	98.91%
pred. drugC	0	16	0	0	0	100.00%
pred. drugX	0	0	53	0	0	100.00%
pred. drugA	0	0	0	23	1	95.83%
pred. drugB	0	0	0	0	15	100.00%
class recall	100.00%	100.00%	98.15%	100.00%	93.75%	

Gambar 11. Tabel kappa

- kappa : 0.985 +/-0.044 (mikro : 0.986)

Kappa adalah metrik yang mengukur kesepakatan antara prediksi model dengan label asli, dengan memperhitungkan kesepakatan yang terjadi secara acak. Nilai 0.985 menunjukkan kesepakatan yang sangat tinggi antara prediksi dan label asli.

- Confusion Matrix:
 - DrugY : 91 prediksi benar, 1 prediksi salah (diprediksi sebagai drugX)
 - DrugC : 16 prediksi benar, 0 prediksi salah
 - DrugX : 53 prediksi benar, 0 prediksi salah
 - DrugA : 23 prediksi benar, 1 prediksi salah (diprediksi sebagai drugB)
 - DrugB : 15 prediksi benar, 0 prediksi salah
- Class Precision: Tingkat Presisi masing – masing kelas :
 - DrugY : 98.91%
 - DrugC : 100.00%
 - DrugX : 100.00%
 - DrugA : 95.83%
 - DrugB : 100.00%
- Class Recall: Tingkat recall masing – masing kelas :
 - DrugY : 100.00%
 - DrugC : 100.00%
 - DrugX : 98.15%
 - DrugA : 100.00%
 - DrugB : 93.75%

Nilai Kappa sebesar 0.985 ± 0.044 menunjukkan tingkat kesepakatan yang sangat tinggi antara model prediksi dan label asli, dengan perhitungan kemungkinan kesepakatan acak. Hal ini mencerminkan kualitas model yang andal dan akurat, sebagaimana diperkuat oleh hasil Confusion Matrix dengan prediksi benar yang dominan untuk semua kelas obat serta tingkat presisi dan recall yang tinggi (hingga 100% untuk sebagian besar kelas). Model ini terbukti efektif dan sesuai untuk mendukung pengambilan keputusan klinis.

Berdasarkan hasil penelitian kami dapat menghasilkan perbandingan dengan penelitian sebelumnya yang diteliti oleh Zaki, Mohammed J.Meira, Jr, Wagner [14] yang mendapat accuracy sebesar 0.93% dibanding penelitian kami dengan accuracy sebesar 99%. Yang menunjukkan penggunaan algoritma decision tree sudah tepat.

5. KESIMPULAN

Penelitian ini berhasil mengidentifikasi faktor-faktor kunci dalam pemilihan obat untuk pasien dengan tekanan darah tinggi dan kolesterol menggunakan algoritma decision tree. Dataset yang digunakan terdiri dari 200 data yang mencakup atribut seperti usia, jenis kelamin, tekanan darah, kolesterol, rasio natrium-kalium, dan jenis obat. Penelitian ini mencapai akurasi yang sangat tinggi, yaitu 99%, dibandingkan dengan akurasi 93% dari penelitian sebelumnya. Hal ini menunjukkan bahwa metode decision tree efektif dalam membantu pengambilan keputusan medis terkait pemilihan obat yang tepat.

6. SARAN

Saran Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi algoritma lain seperti Random Forest, Gradient Boosting, atau Support Vector Machines (SVM) guna membandingkan kinerja masing-masing dalam klasifikasi pemilihan obat, sehingga dapat ditemukan metode yang lebih optimal.

DAFTAR PUSTAKA

- [1] K. Abdul Khalim, U. Hayati, and A. Bahtiar, "Perbandingan Prediksi Penyakit Hipertensi Menggunakan Metode Random Forest Dan Naïve Bayes," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 498–504, 2023, doi: 10.36040/jati.v7i1.6376.
- [2] M. Mumpuni, I. Kusumastuti, and S. Manurung, "Hubungan Tingkat Pengetahuan Gizi Dan Kepatuhan Diet Terhadap Kadar Kolesterol Darah Penderita Penyakit Jantung Koroner," *J. Med. (Media Inf. Kesehatan)*, vol. 10, no. 2, pp. 279–294, 2023, doi: 10.36743/medikes.v10i2.538.
- [3] T. B. Pamungkas, S. Maesaroh, and P. Ardiansyah, "Implementasi Data Mining Pada Stok Penggunaan Barang Di Gmf Aeroasia Menggunakan Algoritma K-Means Clustering," *J. Ilm. Sains dan Teknol.*, vol. 7, no. 2, pp. 112–123, 2023, doi: 10.47080/saintek.v7i2.2697.
- [4] M. A. Fais *et al.*, "Implementasi Algoritma Decision Tree untuk Klasifikasi Serangan Jantung," vol. 1, no. 4, pp. 207–212, 2023, [Online]. Available: <https://doi.org/10.59581/jusiik-widyakarya.v1i4.1895>

- [5] M. F. R. Aditya, N. Lutvi, and U. Indahyanti, "Prediksi Penyakit Hipertensi Menggunakan Metode Decison Tree dan Random Forest," *J. Ilm. Komputasi*, vol. 23, no. 1, pp. 9–16, 2024, doi: 10.32409/jikstik.23.1.3503.
- [6] R. Antika, A. Rifa'l, F. Dikananda, D. Indriya Efendi, and R. Narasati, "Penerapan Algoritma Decision Tree Berbasis Pohon Keputusan Dalam Klasifikasi Penyakit Jantung," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 7, no. 6, pp. 3688–3692, 2024, doi: 10.36040/jati.v7i6.8264.
- [7] N. Swastini, "Efektivitas Daun Sirsak (*Annona muricata* Linn) terhadap Penurunan Tekanan Darah Pada Hipertensi," *J. Ilm. Kesehat. Sandi Husada*, vol. 10, no. 2, pp. 413–415, 2021, doi: 10.35816/jiskh.v10i2.618.
- [8] H. S. Ardytama Kusuma Yudha, "STUDI KORELASI POLA MAKAN DENGAN KADAR KOLESTEROL PADA PASIEN STROKE," *Nucl. Phys.*, vol. 13, no. 1, pp. 104–116, 2023.
- [9] N. Abdillah, H. Susilo, and M. Ihksan, "Sosialisasi Pemanfaatan Teknologi Data Mining Untuk Analisis Data Kesehatan Di Klinik Amanah," *J. Abdimas Saintika*, vol. 5, no. 1, pp. 181–186, 2023, [Online]. Available: <https://jurnal.syedzasaintika.ac.id/index.php/abdimas/article/view/1940/1354%0Ahttps://jurnal.syedzasaintika.ac.id/index.php/abdimas/article/view/1940>
- [10] R. Nursyahfitri, A. N. Maharadja, R. A. Farissa, and Y. Umidah, "Klasifikasi Penentuan Jenis Obat Menggunakan Algoritma Decision Tree," *J. Inform. Polinema*, vol. 7, no. 3, pp. 53–60, 2021, doi: 10.33795/jip.v7i3.629.
- [11] M. M. Santoni, N. Chamidah, and N. Matondang, "Prediksi Hipertensi menggunakan Decision Tree, Naïve Bayes dan Artificial Neural Network pada software KNIME," *Techno.Com*, vol. 19, no. 4, pp. 353–363, 2020, doi: 10.33633/tc.v19i4.3872.
- [12] Aditya, F. M. Rizki, and A. Nuril Lutvi, "Prediction of Hypertension Disease Using Decision Tree and Random Forest Methods [Prediksi Penyakit Hipertensi Menggunakan metode Decision Tree dan Random Forest]," pp. 1–9, 2023.
- [13] S. A. Pratiwi, A. Fauzi, S. Arum, P. Lestari, and Y. Cahyana, "KLIK: Kajian Ilmiah Informatika dan Komputer Prediksi Persediaan Obat Pada Apotek Menggunakan Algoritma Decision Tree," *Media Online*, vol. 4, no. 4, pp. 2381–2388, 2024, doi: 10.30865/klik.v4i4.1681.
- [14] M. J. Zaki and W. Meira, Jr, "Decision Tree Classifier," *Data Min. Anal.*, pp. 481–497, 2018, doi: 10.1017/cbo9780511810114.020.