

## Komparasi Naïve Bayes, C4.5 Dan SVM Untuk Klasifikasi Pendonor Darah Di PMI Kabupaten Grobogan

Puput Rismawati\*<sup>1</sup>, Eko Supriyadi<sup>2</sup>, Dhika Malita Puspita Arum<sup>3</sup>

<sup>1,2,3</sup> Program Studi Ilmu Komputer, Universitas An Nuur, Grobogan, Indonesia

e-mail: \*[puputrismawati29@gmail.com](mailto:puputrismawati29@gmail.com), [ekalaya56@gmail.com](mailto:ekalaya56@gmail.com),

[dhika.malita11@gmail.com](mailto:dhika.malita11@gmail.com)

### Abstrak

Kebutuhan darah di Indonesia masih tinggi dan sering tidak sebanding dengan jumlah pendonor yang tersedia. Untuk memastikan kualitas donor, PMI menerapkan kriteria kesehatan tertentu. Namun, proses seleksi masih dilakukan secara manual sehingga memerlukan metode analisis yang lebih efektif. Perkembangan teknik data mining memungkinkan penerapan algoritma klasifikasi untuk menentukan kelayakan pendonor darah secara akurat. Penelitian ini menggunakan dataset pendonor darah dari PMI Kabupaten Grobogan. Menggunakan tiga algoritma klasifikasi, yaitu Naïve Bayes, Decision Tree C4.5, dan Support Vector Machine (SVM), diimplementasikan menggunakan perangkat lunak RapidMiner. Hasil pengujian menunjukkan bahwa algoritma C4.5 memiliki performa terbaik dengan akurasi sebesar 96,77%. Naïve Bayes menghasilkan akurasi 75,47% dengan AUC, sedangkan SVM memperoleh akurasi 72,38%. Algoritma C4.5 paling efektif untuk klasifikasi kelayakan pendonor darah, karena akurat sekaligus mudah diinterpretasi. Ini bisa menjadi dasar pengembangan sistem pendukung keputusan berbasis data mining guna membantu proses seleksi pendonor secara lebih cepat dan objektif.

**Kata kunci**— C4.5, Donor Darah, Klasifikasi, Naïve Bayes, SVM

### Abstract

The demand for blood in Indonesia remains high and is often disproportionate to the number of available donors. To ensure donor quality, PMI applies specific health criteria; however, the selection process is still conducted manually, requiring a more effective analytical method. Advances in data mining techniques enable the application of classification algorithms to determine blood donor eligibility more accurately. This study utilized blood donor datasets from PMI Grobogan Regency. Three classification algorithms—Naïve Bayes, Decision Tree C4.5, and Support Vector Machine (SVM)—were implemented and tested using RapidMiner software. The experimental results showed that the C4.5 algorithm achieved the best performance with an accuracy of 96.77%. Naïve Bayes produced 75.47% accuracy with a high AUC, while SVM obtained 72.38% accuracy. The C4.5 algorithm proved to be the most effective for classifying blood donor eligibility due to its high accuracy and interpretability. This finding can serve as a foundation for developing data mining-based decision support systems to improve the donor selection process more efficiently and objectively.

**Keywords**— Blood Donor, C4.5, Classification, Naïve Bayes, C4.5, SVM

## 1. PENDAHULUAN

Darah merupakan komponen esensial dalam tubuh manusia yang berperan menjaga keberlangsungan organ. Kekurangan darah dapat mengganggu fungsi tubuh, sementara hingga kini belum tersedia alternatif darah buatan sehingga manusia tetap menjadi sumber alami [1]. Menurut WHO, kebutuhan global mencapai 180 juta kantong per tahun, dengan cadangan minimal 2% dari total populasi. Di Indonesia, jumlah tersebut setara dengan 4,5 juta kantong per tahun, namun distribusi darah masih timpang karena hampir 50% pasokan dunia berasal dari negara maju yang hanya mewakili 20% populasi [2]. Donor darah merupakan aktivitas sukarela guna menyumbangkan darah untuk kemudian disimpan di bank darah dan digunakan dalam transfusi bagi pasien yang membutuhkan [3].

Pelaksanaan donor darah di Indonesia berlandaskan kerangka hukum yang jelas, salah satunya melalui Peraturan Pemerintah Nomor 07 Tahun 2019 tentang Kepalaangmerahan. Regulasi tersebut menetapkan Palang Merah Indonesia (PMI) sebagai lembaga tunggal yang berwenang menyelenggarakan kegiatan kepalaangmerahan, termasuk pelayanan donor darah melalui Unit Donor Darah (UDD). Tugas ini mencakup seluruh tahapan manajemen darah, mulai dari perekrutan pendonor, pemeriksaan kelayakan, pemrosesan, penyimpanan, hingga distribusi darah ke fasilitas pelayanan kesehatan [4]. Kebijakan ini diperkuat oleh Undang-Undang Nomor 36 Tahun 2009 tentang Kesehatan serta Peraturan Pemerintah Nomor 7 Tahun 2011 tentang Pelayanan Darah, yang menegaskan tanggung jawab pemerintah dalam menjamin ketersediaan pelayanan darah yang aman, bermutu, mudah diakses, dan sesuai kebutuhan masyarakat. Tanggung jawab tersebut tidak hanya berada pada pemerintah pusat, tetapi juga melibatkan pemerintah daerah dalam aspek regulasi, pembinaan, pengawasan, hingga dukungan pendanaan. Pasal 90 UU No. 36/2009 dan Pasal 46 PP No. 7/2011 menegaskan bahwa dukungan pendanaan dapat diberikan melalui subsidi kepada UDD, baik bersumber dari Anggaran Pendapatan dan Belanja Negara (APBN), Anggaran Pendapatan dan Belanja Daerah (APBD), maupun bentuk bantuan lainnya [5]. Dengan demikian, regulasi ini diharapkan mampu memberikan dasar hukum yang kuat untuk menjamin ketersediaan darah yang mencukupi serta layanan transfusi darah yang berkualitas bagi masyarakat.

Dalam praktiknya, darah memiliki peran yang sangat vital dalam penyelamatan nyawa manusia, terutama bagi pasien yang membutuhkan transfusi akibat kondisi medis tertentu seperti kecelakaan, operasi besar, komplikasi persalinan, maupun penyakit kronis seperti thalassemia dan kanker. Oleh karena itu, PMI secara berkesinambungan melaksanakan berbagai program kampanye untuk meningkatkan kesadaran masyarakat mengenai pentingnya donor darah. Kampanye tersebut tidak semata menekankan donor darah sebagai wujud solidaritas sosial, melainkan juga mendorong masyarakat untuk mengintegrasikan aktivitas donor darah ke dalam pola hidup sehat (lifestyle) [6]. Meski demikian, terdapat tantangan besar dalam menjamin kualitas dan keamanan darah. Hal ini karena darah juga dapat menjadi salah satu media penularan berbagai penyakit menular, seperti HIV/AIDS, hepatitis B, hepatitis C, maupun sifilis. Oleh sebab itu, darah transfusi yang digunakan di fasilitas kesehatan sebaiknya berasal dari pendonor dengan risiko rendah guna meminimalisasi potensi penularan [7].

Sejalan dengan hal tersebut, World Health Organization (WHO) merekomendasikan bahwa darah yang aman, berkualitas, dan berisiko rendah sebaiknya diperoleh dari pendonor sukarela yang mendonorkan darahnya tanpa ditujukan kepada penerima tertentu, bukan dari donor pengganti yang biasanya memberikan darah hanya ketika ada anggota keluarga atau kerabat yang membutuhkan transfuse [8]. Donor sukarela dianggap lebih aman karena umumnya dilakukan secara rutin dan atas dasar kesadaran pribadi, sehingga lebih mudah diseleksi serta dipantau kesehatannya. Sebaliknya, donor pengganti sering kali bersifat insidental dan tidak menutup kemungkinan memiliki risiko kesehatan yang tidak teridentifikasi sebelumnya [9]. Namun, fakta di lapangan menunjukkan bahwa minimnya kesadaran masyarakat untuk menjadi pendonor sukarela masih menjadi permasalahan serius yang mengakibatkan rendahnya ketersediaan darah di Unit Donor Darah (UDD). Kondisi ini berimplikasi pada belum optimalnya pemenuhan kebutuhan darah nasional, sehingga berbagai

upaya strategis perlu dilakukan, baik oleh pemerintah, PMI, maupun masyarakat, untuk meningkatkan partisipasi donor sukarela secara berkelanjutan [10].

Perkembangan teknologi data mining membuka peluang dalam klasifikasi kelayakan pendonor darah. Data mining, yang dikenal pula sebagai Knowledge Discovery in Database (KDD), adalah kegiatan multidisipliner untuk mengekstraksi pola dan pengetahuan baru dari basis data berukuran besar yang memanfaatkan metode statistika, komputasi, kecerdasan buatan, dan pembelajaran mesin [11]. Salah satu teknik yang umum diterapkan dalam data mining adalah klasifikasi, yakni proses pengelompokan objek ke dalam kategori tertentu berdasarkan atribut-atribut yang dianggap relevan [12]. Tiga algoritma yang sering digunakan adalah Naïve Bayes, C4.5, serta Support Vector Machine (SVM). Naïve Bayes mengklasifikasi berdasarkan prinsip probabilitas [13], C4.5 menciptakan pohon keputusan dengan entropy dan information gain [14], sedangkan SVM menemukan hyperplane optimal guna memisahkan kelas data, termasuk data non-linear melalui fungsi kernel [15].

Sejumlah penelitian terdahulu yang menggunakan ide klasifikasi dengan memanfaatkan algoritma Naïve Bayes, C4.5 dan SVM antara lain : Komparasi Algoritma C4.5 dan Naïve Bayes dalam Penentuan Status Kelayakan Donor Darah [16], Perbandingan Klasifikasi Data Calon Pendonor Darah dengan Algoritma K-Nearest Neighbor dan Support Vector Machine [17]. Sesuai dengan penelitian yang telah dilakukan tersebut, dapat disimpulkan bahwa setiap algoritma klasifikasi mempunyai kelebihan dan kekurangan dalam menentukan kelayakan maupun status calon pendonor darah. Naïve Bayes cenderung unggul dalam proses komputasi yang sederhana dan cepat, namun pada beberapa kasus seperti penentuan kelayakan donor darah hasil akurasi masih perlu dibandingkan dengan algoritma lain [16]. Algoritma C4.5 mampu memberikan interpretasi yang lebih jelas melalui pohon keputusan sehingga memudahkan analisis, meskipun terkadang kinerjanya dipengaruhi oleh kompleksitas data. Sementara itu, metode SVM memberikan hasil yang lebih baik dalam menangani data dengan dimensi tinggi dan menghasilkan akurasi yang cukup tinggi dibandingkan metode lain [17].

Analisis klasifikasi dilakukan menggunakan perangkat lunak RapidMiner yang menyediakan berbagai operator untuk pre-processing, pemodelan, serta evaluasi [18]. Dalam penggunaannya, dataset terlebih dahulu diimpor ke dalam aplikasi. Selanjutnya, pengguna dapat menyusun operator sesuai kebutuhan, misalnya untuk menangani data yang hilang. Setelah data diproses dengan lengkap, tahap pemodelan atau analisis dapat dilakukan. Proses kerja dilakukan melalui design view, yang menampilkan alur operator dan dataset, sedangkan hasil pemrosesan dapat diperiksa melalui result view.

Oleh karena itu, penelitian ini bertujuan mengklasifikasikan calon pendonor darah di UDD PMI Kabupaten Grobogan ke dalam kategori “Layak” dan “Tidak Layak” menggunakan algoritma Naïve Bayes, C4.5, dan SVM. Selain itu, penelitian ini membandingkan tingkat akurasi ketiga algoritma dengan parameter usia, tekanan darah, kadar hemoglobin, dan berat badan, serta melakukan validasi hasil prediksi menggunakan confusion matrix.

## 2. METODE PENELITIAN

Studi ini dilakukan melalui pendekatan kuantitatif dengan metode eksperimen sekaligus komparatif. Tujuan utamanya adalah menguji serta membandingkan kinerja tiga algoritma klasifikasi, yakni Naïve Bayes, C4.5, dan SVM, dalam menentukan kelayakan calon pendonor darah. Ketiga algoritma ini dipilih karena memiliki karakteristik yang berbeda. Naïve Bayes bekerja dengan prinsip probabilitas sederhana, C4.5 menghasilkan model berupa pohon keputusan yang mudah dipahami, sedangkan SVM berupaya mencari garis pemisah terbaik untuk membedakan data antar kelas.

Data yang digunakan dalam penelitian ini merupakan sampel data pendonor darah di UDD PMI Kabupaten Grobogan periode 2021–2024 yang berjumlah 3013. Data tersebut dipilih karena mencakup informasi aktual yang relevan dengan proses donor darah. Atribut yang dianalisis meliputi umur, berat badan, tekanan darah (sistolik dan diastolik), serta kadar

hemoglobin. Seluruh atribut ini sesuai dengan syarat medis yang ditetapkan PMI untuk memastikan keamanan dan kelayakan proses donor darah.

Untuk menganalisis data, penelitian ini memanfaatkan aplikasi RapidMiner AI Studio versi 2025.1.0. Perangkat ini dipilih karena menyediakan fitur lengkap untuk mengolah data, melatih model, dan mengukur kinerja algoritma dengan berbagai indikator seperti akurasi, presisi, dan recall. Dengan bantuan perangkat lunak tersebut, tahapan penelitian dapat dilakukan secara lebih terstruktur, mulai dari pengolahan data awal hingga evaluasi hasil klasifikasi.

### 2.1 Deskripsi Operasional Variabel

Pada tahap ini peneliti menyusun dataset akhir dengan menyeleksi data mentah berdasarkan satu atribut sebagai ID, enam atribut prediktor dan satu atribut hasil. Dataset tersebut digunakan dalam tahap pemodelan untuk menentukan status pendonor darah, apakah dinyatakan berhasil atau gagal, sebagaimana ditunjukkan pada tabel berikut :

Tabel 1 Jenis Atribut

No	Nama Atribut	Tipe Data	Keterangan
1	Nama	Polinomial	Identitas Pendonor
2	Tanggal Donor	Date	Tanggal donor dilakukan
3	Jenis Kelamin	Polinomial	Laki-Laki/ Perempuan
4	Golongan Darah	Integer	A+/B+/O+/AB+
5	Umur	Integer	Umur dalam tahun
6	Berat Badan	Integer	Berat dalam Kilogram
7	Hemoglobin	Integer	Kadar HB dalam g/dL
8	Tensi_Sistolik	Integer	Tensi darah atas
9	Tensi_Diastolik	Integer	Tensi darah bawah
10	Status	Binomial	Layak/ Tidak

Setelah atribut dan kategori ditetapkan, dataset divalidasi berdasarkan atribut yang tersedia untuk kemudian digunakan dalam eksperimen dengan algoritma Naïve Bayes, C4.5, dan SVM. Hasil penelitian mencakup temuan kualitatif dan kuantitatif yang diperoleh melalui proses perhitungan menggunakan model yang diusulkan. Sebagai bentuk validasi awal, data disajikan dalam tabel sampel dataset berikut.

No	Nama	Tanggal Donor	JK	Golda	Umur	BB	HB	Tensi	Status
1	Aseh Dwi Jayant	6/11/2022	P	O+	24	68	13.4	136/85	Layak
2	Aseh Dwi Jayant	9/8/2022	P	O+	24	68	13.2	131/88	Layak
3	Aseh Dwi Jayant	12/1/2022	P	O+	24	69	13.2	124/86	Layak
4	Aseh Dwi Jayant	3/14/2023	P	O+	25	69	12.7	113/74	Layak
5	Aseh Dwi Jayant	7/8/2023	P	O+	25	69	13.1	119/79	Layak
6	Aseh Dwi Jayant	11/10/2023	P	O+	25	70	12.9	104/78	Layak
7	Aseh Dwi Jayant	2/20/2024	P	O+	26	70	15.9	110/80	Layak
8	Aseh Dwi Jayant	4/23/2024	P	O+	26	70	13.5	128/86	Layak
9	Aseh Dwi Jayant	7/16/2024	P	O+	26	69	12.9	122/87	Layak
10	Diky Putra	6/30/2024	L	B+	29	97	13.6	132/78	Layak
...	...	...	...	...	...	...	...	...	...
3003	Zuvita Sasa Dila	8/17/2021	P	O+	34	60	13.2	119/76	Layak
3004	Zuvita Sasa Dila	3/17/2022	P	O+	35	59	14.6	116/70	Layak
3005	Zuvita Sasa Dila	6/18/2022	P	O+	35	58	13.6	112/82	Layak
3006	Zuvita Sasa Dila	8/20/2022	P	O+	35	56	14.6	110/83	Layak
3007	Zuvita Sasa Dila	11/26/2022	P	O+	35				Layak
3008	Zuvita Sasa Dila	2/19/2023	P	O+	35	59	12.6	111/72	Layak
3009	Zuvita Sasa Dila	3/16/2023	P	O+	36	60	13.5	118/72	Layak
3010	Zuvita Sasa Dila	6/18/2023	P	O+	36	60	14.2	128/80	Layak
3011	Zuvita Sasa Dila	9/4/2023	P	O+	36	60	12.9	129/76	Layak
3012	Zuvita Sasa Dila	12/8/2023	p	O+	36	60	15.5	130/90	Layak
3013	Zuvita Sasa Dila	3/7/2024	P	O+	37	60	15.2	122/74	Layak

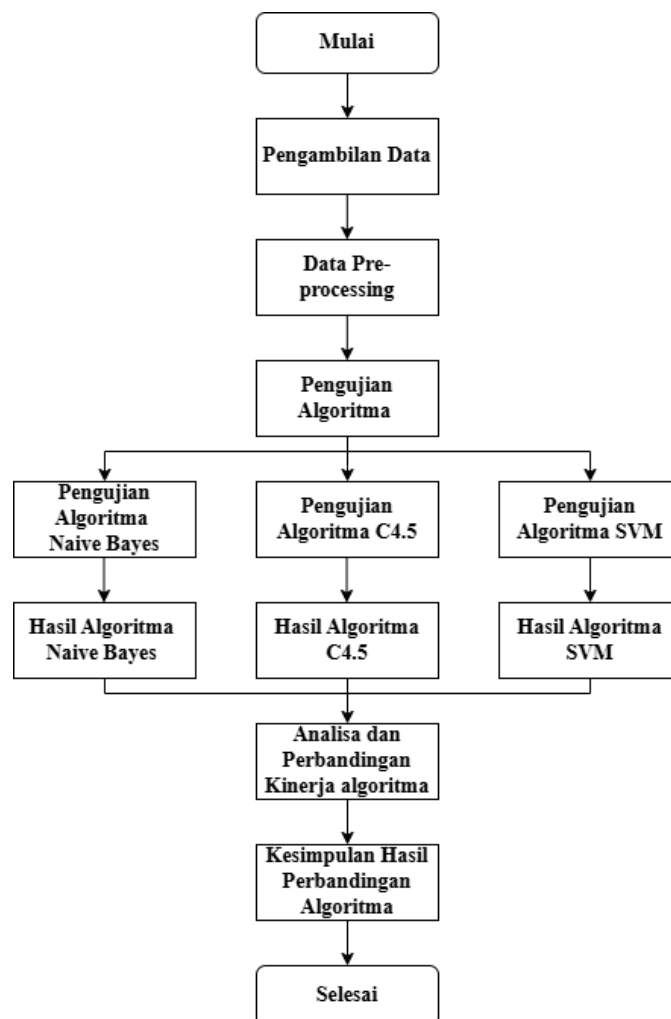
Gambar 1 Dataset PMI Kabupaten Grobogan

### 2.2 Teknik Analisis

Proses analisis dalam penelitian ini dilaksanakan melalui beberapa tahapan, antara lain:

- a. Preprocessing Data :
  - 1) Data cleaning untuk menghapus data tidak lengkap atau kosong.
  - 2) Data reduction dengan menghapus atribut yang tidak relevan.
  - 3) Data transformation dengan mengubah format file dari Excel ke CSV.
- b. Pengujian Algoritma :
  - 1) Algoritma Naïve Bayes untuk memprediksi kelayakan pendonor berdasarkan probabilitas dari atribut-atributnya.
  - 2) Algoritma C4.5 menciptakan pohon keputusan dengan mempertimbangkan atribut yang paling berpengaruh untuk menentukan kelayakan pendonor darah.
  - 3) Algoritma SVM menemukan hyperplane berdimensi (p-1) guna memisahkan data pelatihan berdasarkan kategorinya.

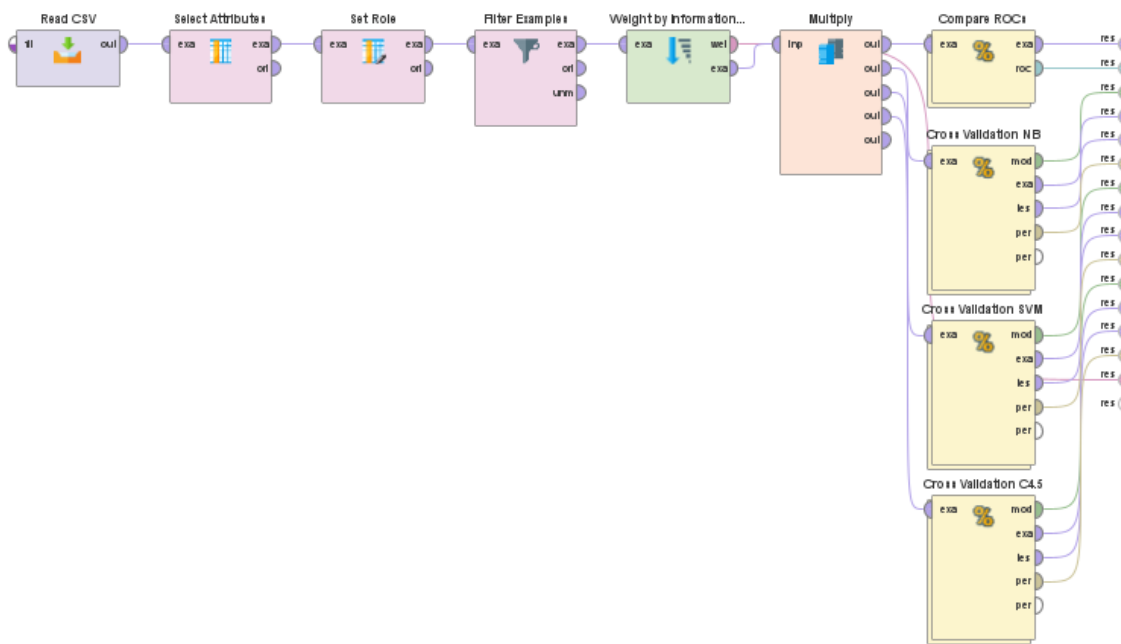
### 2.3 Diagram Alur Penelitian



Gambar 2 Diagram Alur Penelitian

### 2.4 Implementasi di RapidMiner

Studi ini mengimplementasikan tiga algoritma klasifikasi, meliputi Naïve Bayes, C4.5, dan SVM, yang digabungkan dalam satu skema pemrosesan pada RapidMiner AI Studio 2025.1.0. Skema pemodelan tersebut dapat dilihat pada gambar berikut:



Gambar 3 Skema Implementasi di RapidMiner

### 3. HASIL DAN PEMBAHASAN

Setelah seluruh operator disusun dalam alur proses, langkah berikutnya adalah menjalankan pengujian untuk mengevaluasi performa model yang menghasilkan Confusion Matrix, berikut ini penjelasannya :

#### 3.1 Hasil Weight by Information Gain

attribute	weight
BB	0.001
Umur	0.002
Tensi_Si...	0.019
Tensi_Di...	0.023
Tanggal ...	0.052
HB	0.193

Gambar 4 Hasil Weight by Infromation Gain

Tabel hasil pembobotan atribut menunjukkan bahwa setiap variabel memiliki tingkat kontribusi yang berbeda terhadap klasifikasi kelayakan pendonor darah. Atribut berat badan (0,001) dan umur (0,002) memiliki bobot paling rendah sehingga pengaruhnya relatif kecil dalam menentukan kelayakan. Sementara itu, atribut tekanan darah, baik sistolik (0,019) maupun diastolik (0,023), memberikan kontribusi yang lebih besar dibandingkan berat badan

dan umur, yang menandakan bahwa kondisi tekanan darah cukup penting dalam proses klasifikasi. Di sisi lain, atribut hemoglobin (0,193) memiliki bobot tertinggi dibandingkan variabel lainnya, sehingga kadar hemoglobin merupakan faktor paling dominan dalam menentukan kelayakan pendonor darah sesuai dengan standar medis.

### 3.2 Hasil Implementasi Ketiga Metode

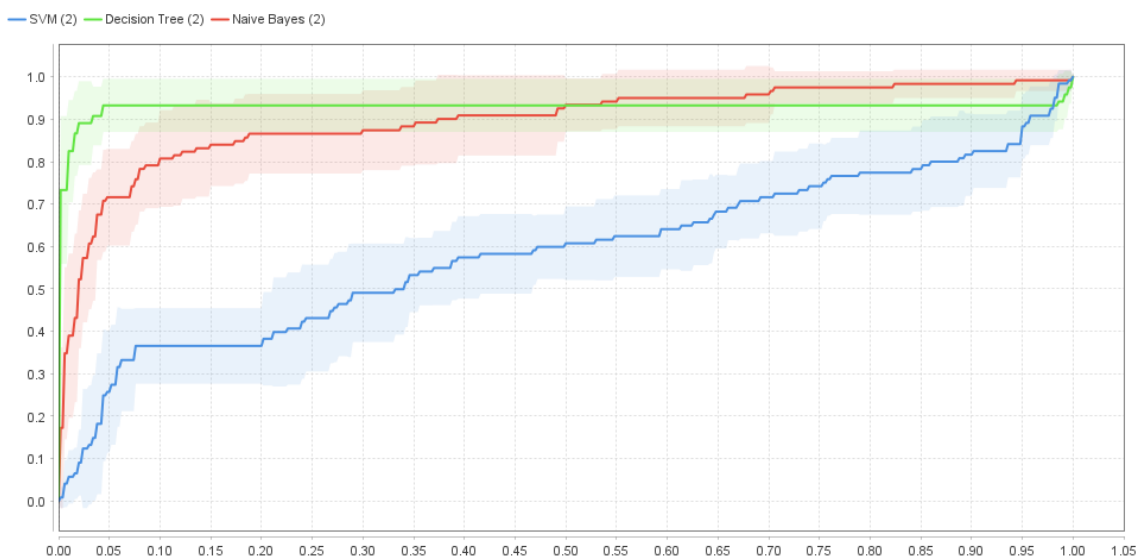
Dalam perbandingan yang dilakukan dari performa ketiga metode klasifikasi ini, peneliti akan membandingkan beberapa algoritma untuk menentukan yang mana metode yang menghasilkan hasil terbaik yaitu Naïve Bayes, SVM atau C4.5. Perbandingan ini akan dilakukan dengan membandingkan hasil dari performa ketiga metode untuk mengetahui hasil metrik yang dihasilkan seperti Accuracy, Precision, dan Recall. Berikut hasil dari perbandingan ketiga metode :

Tabel 2 hasil Implementasi Ketiga Algoritma

<b>Algoritma</b>	<b>Akurasi</b>	<b>Presisi</b>	<b>Recall</b>	<b>AUC</b>
Naïve Bayes	75.47%	99.32%	74.61%	0.917
C4.5	96.77%	99.86%	96.73%	0.971
SVM	72.38%	99.22%	71.39%	0.904

Berdasarkan hasil pengujian yang ditunjukkan pada tabel, terlihat bahwa setiap algoritma memiliki performa yang berbeda dalam mengklasifikasikan kelayakan pendonor darah. Algoritma Naïve Bayes menghasilkan akurasi sebesar 75,47% dengan tingkat presisi yang sangat tinggi, yaitu 99,32%, namun nilai recall masih relatif rendah sebesar 74,61%. Temuan ini mengindikasikan bahwa meskipun model mampu menghasilkan prediksi positif dengan tingkat akurasi yang baik, masih terdapat kasus positif yang tidak terdeteksi. Selanjutnya, algoritma Decision Tree C4.5 memperoleh performa paling unggul dengan akurasi 96,77%, presisi 99,86%, recall 96,73%, serta nilai AUC 0,971. Nilai tersebut menunjukkan bahwa C4.5 memiliki kemampuan yang lebih stabil dan seimbang dalam mengklasifikasikan data dibandingkan algoritma lainnya. Sementara itu, algoritma SVM menunjukkan akurasi terendah yaitu 72,38% dengan presisi 99,22% dan recall 71,39%. Meskipun tingkat presisi yang diperoleh cukup tinggi, rendahnya akurasi dan recall mengindikasikan keterbatasan algoritma ini dalam mengenali seluruh data positif secara konsisten. Hasil analisis menunjukkan bahwa algoritma C4.5 memiliki kinerja terbaik di antara ketiga algoritma yang diuji, ditandai dengan kemampuan menjaga keseimbangan antara akurasi, presisi, recall, serta nilai AUC yang relatif tinggi.

### 3.2 Hasil Compare ROCs



Gambar 5 Hasil Compare ROCs

Gambar tersebut memperlihatkan perbandingan kurva Receiver Operating Characteristic (ROC) dari tiga algoritma klasifikasi, yakni SVM yang direpresentasikan dengan garis biru, C4.5 dengan garis hijau, dan Naïve Bayes dengan garis merah. Kurva ROC menunjukkan hubungan antara sensitivitas dengan berbagai nilai ambang batas (threshold) dalam proses klasifikasi. Secara umum, semakin mendekati sudut kiri atas grafik, kurva ROC mencerminkan kinerja algoritma yang lebih baik dalam membedakan kelas positif dan negatif secara konsisten.

Berdasarkan visualisasi tersebut, algoritma C4.5 menunjukkan performa paling unggul dengan kurva ROC yang berada pada posisi tertinggi dan cenderung stabil, menegaskan kemampuannya dalam membedakan kelas “Layak” dan “Tidak Layak” dengan tingkat akurasi yang tinggi. Algoritma Naïve Bayes juga memperlihatkan performa yang cukup baik, terlihat dari posisi kurvanya yang relatif dekat dengan C4.5, meskipun tidak seoptimal algoritma tersebut. Di sisi lain, SVM menempati posisi paling rendah dibandingkan dua algoritma lainnya, menunjukkan bahwa meskipun tingkat presisinya cukup tinggi, model ini masih memiliki keterbatasan dalam mendeteksi seluruh kasus positif secara konsisten.

Secara keseluruhan, hasil perbandingan kurva ROC mendukung temuan sebelumnya bahwa algoritma C4.5 merupakan metode paling optimal dalam klasifikasi kelayakan pendonor darah. Algoritma Naïve Bayes masih dapat dijadikan alternatif dengan performa yang cukup baik, sedangkan SVM, meskipun layak digunakan, menunjukkan kinerja yang relatif lebih rendah dibandingkan keduanya.

## 4. KESIMPULAN

Penelitian mengenai klasifikasi kelayakan pendonor darah di PMI Kabupaten Grobogan dengan algoritma Naïve Bayes, C4.5, dan Support Vector Machine (SVM) menunjukkan bahwa C4.5 memberikan performa terbaik dengan akurasi 96,77%, presisi 99,86%, dan recall 96,73%. Hasil ini jauh lebih tinggi dibandingkan Naïve Bayes (akurasi 75,47%) maupun SVM (akurasi 72,38%), sehingga menegaskan keunggulan C4.5 dalam menghasilkan klasifikasi yang akurat, konsisten, dan stabil. Selain keunggulan teknis, C4.5 juga menghasilkan model pohon keputusan yang transparan dan mudah diinterpretasikan, menjadikannya relevan untuk aplikasi di bidang kesehatan. Analisis atribut menunjukkan bahwa kadar hemoglobin merupakan faktor

utama dalam penentuan kelayakan donor, diikuti tekanan darah, usia, dan berat badan sebagai indikator pendukung. Secara keseluruhan, penelitian ini menyimpulkan bahwa C4.5 adalah metode paling optimal dalam klasifikasi kelayakan pendonor darah, sedangkan Naïve Bayes masih layak dipertimbangkan sebagai alternatif, dan SVM relevan dalam kondisi tertentu meskipun kurang kompetitif. Temuan ini membuka peluang penelitian lanjutan melalui pengembangan model hibrida atau optimasi parameter untuk meningkatkan akurasi dan reliabilitas sistem klasifikasi secara lebih komprehensif.

## 5. SARAN

Berdasarkan hasil penelitian dan evaluasi performa, terdapat beberapa poin penting yang dapat dijadikan pertimbangan. Pertama, algoritma Decision Tree C4.5 terbukti sebagai metode paling optimal dalam klasifikasi kelayakan pendonor darah di PMI Kabupaten Grobogan, ditunjukkan melalui kinerja terbaik pada seluruh metrik evaluasi (akurasi, presisi, recall, dan AUC) serta keunggulannya dalam menghasilkan pohon keputusan yang transparan dan mudah diinterpretasikan. Kedua, penelitian lanjutan disarankan untuk mengeksplorasi algoritma lain, seperti Random Forest atau Gradient Boosting, guna memperoleh hasil klasifikasi yang lebih akurat dan stabil. Ketiga, keterbatasan penelitian ini terletak pada penggunaan lima atribut utama (umur, berat badan, hemoglobin, tekanan darah sistolik, dan diastolik), sehingga pada studi berikutnya perlu dipertimbangkan penambahan atribut lain, termasuk riwayat penyakit, gaya hidup, frekuensi donor, serta hasil pemeriksaan laboratorium, agar model yang dihasilkan lebih komprehensif dan reliabel.

## DAFTAR PUSTAKA

- [1] S. S. Elfaretta, A. A. Arifiyanti, and A. S. Fitri, "KLASIFIKASI CALON PENDONOR DARAH POTENSIAL MENGGUNAKAN ALGORITMA DECISION TREE DI UTD PMI KOTA SURABAYA," *JITET (Jurnal Inform. dan Tek. Elektro Ter.*, vol. 12, no. 3, 2024.
- [2] M. R. Firdaus *et al.*, "KLASIFIKASI KELAYAKAN CALON PENDONOR DARAH MENGGUNAKAN NEURAL NETWORK," *Sist. J. Sist. Inf.*, vol. 9, pp. 362–371, 2020.
- [3] I. T. Utami, Fadryani, and D. Daniaty, "PERBANDINGAN KLASIFIKASI STATUS PENDONOR DARAH DENGAN MENGGUNAKAN REGRESI LOGISTIK DAN K-NEAREST," *J. Apl. Stat. Komputasi Stat.*, vol. 12, 2020.
- [4] N. P. P. Pratama, T. Sukmono, B. I. Putra, and Sulistiyowati, "Forecasting the Amount of Blood Storage Using the Support Vector Machine (Svm) Method," *Procedia Eng. Life Sci.*, vol. 3, no. 07, 2023, doi: 10.21070/pels.v3i0.1359.
- [5] Y. Irawan, "Penerapan Algoritma Decision Tree C4 . 5 Untuk Prediksi Kelayakan Calon Pendonor Darah Dengan Klasifikasi Data Mining ( Application of the C4 . 5 Decision Tree Algorithm to Predict the Eligibility of Prospective Blood Donors by Classification Data Mining)," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 181–189, 2021.
- [6] N. Nur, N. Syahra, and Asmawati, "Perbandingan Metode k-NN dan Naïve Bayes dalam Klasifikasi Penentuan Calon Pendonor Darah," *JCIS (Journal Comput. Inf. Syst.*, vol. 1, no. 1, pp. 21–28, 2021, doi: 10.31605/jcis.v1i1.875.
- [7] G. Wahyu, N. Wibowo, A. Sucipto, and H. Kristanto, "OPTIMASI PSO PADA NAÏVE BAYES UNTUK PREDIKSI CALON PENDONOR DARAH TETAP ( STUDI KASUS PMI KABUPATEN KUDUS )," *JTINFO (Jurnal Tek. Inform.*, vol. 2, no. 2, 2023.
- [8] N. K. M. Astuti, N. W. Utami, and I. G. P. K. Juliharta, "Classification of Blood Donor Data Using C4.5 and K-Nearest Neighbor Methods (Case Study: Utd Pmi Bali Province)," *J. Pilar Nusa Mandiri*, vol. 18, no. 1, pp. 9–16, 2022, doi: 10.33480/pilar.v18i1.2790.
- [9] H. Wahono and D. Riana, "Prediksi Calon Pendonor Darah Potensial Dengan Algoritma

- Naïve Bayes, K-Nearest Neighbors dan Decision Tree C4.5,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 7, 2020, doi: 10.30865/jurikom.v7i1.1953.
- [10] M. Lestandy, A. Faruq, and A. Faruq, “Klasifikasi pendonor darah potensial menggunakan pendekatan algoritme pembelajaran mesin,” *J. Teknol. dan Sist. Komput.*, vol. 8, no. April, pp. 217–221, 2020, doi: 10.14710/jtsiskom.2020.13619.
- [11] U. Suriani, “Penerapan Data Mining untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree C4.5,” *Journalcisa*, vol. 3, no. 2, pp. 55–66, 2023, [Online]. Available: <http://jesik.web.id/index.php/jesik/article/view/91>
- [12] H. P. Hutrim and Nurahman, “PENERAPAN DATA MINING UNTUK PENENTUAN ( STUDI KASUS PMI KABUPATEN KOTAWARINGIN TIMUR ),” *JUISI*, vol. 06, no. 02, 2020.
- [13] A. Rudiantoro, “PENERAPAN ALGORITMA NAÏVE BAYES CLASSIFIER UNTUK KLASIFIKASI DONOR DARAH BERBASIS WEB ( Studi Kasus Palang Merah Indonesia Kota Kebumen ),” Universitas Teknologi Yogyakarta, 2019.
- [14] M. Zainuri, “Komparasi Metode Klasifikasi Data Mining Algoritma C5.0 Dan Naïve Bayes Untuk Memprediksi Jurusan Siswa (Studi Kasus: Sman 1 Gondanglegi),” Universitas Islam Raden Rahmat, 2022.
- [15] Hendriyana, I. M. Karo Karo, and S. Dewi, “Analisis perbandingan Algoritma Support Vector Machine, Naive Bayes dan Regresi Logistik untuk Memprediksi Donor Darah,” *J. Teknol. Terpadu*, vol. 8, no. 2, pp. 121–126, 2022, doi: 10.54914/jtt.v8i2.581.
- [16] K. Handayani, Lisnawanty, A. Latif, M. R. Firdaus, and F. N. Hasan, “Komparasi Algoritma C4 . 5 dan Naïve Bayes dalam Penentuan Status Kelayakan Donor Darah,” *Sist. J. Sist. Inf.*, vol. 10, pp. 676–687, 2021.
- [17] R. Ramadhan and A. Triyudi, “Perbandingan Klasifikasi Data Calon Pendonor Darah dengan Algoritma K-Nearest Neighbor dan Support Vector Machine,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 2, pp. 1250–1260, 2022, doi: 10.35957/jatisi.v9i2.2018.
- [18] A. Widyanto, Kusrini, and Kusnawi, “Pengaruh Keseimbangan Data terhadap Akurasi Model Support Vector Machine pada Data Set Donor Darah,” *J. Teknol. Terpadu*, vol. 9, no. 2, pp. 79–88, 2023, doi: 10.54914/jtt.v9i2.771.