

Klasifikasi Kelayakan Calon Pendoron Darah Di PMI menggunakan Decision Tree

Eligibility Classification of Blood Donor Candidates at PMI Using Decision Tree

**Arina Abrilian Pulung^{a,1}, Rebeka Mentari^{b,2}, Taufiq Denny Wicaksono^{c,3},
Enjelika Rosan^{d,4}, Ir. Aulia Agung Dermawan^{e,5}**

^{a,b}Institut Teknologi Batam, The Vitka City Complex, Tiban, Jl. Gajah Mada, Kota Batam, 29425,
Indonesia.

2212043@student.iteba¹.ac.id, [2212047@student.iteba.ac.id²](mailto:2212047@student.iteba.ac.id<sup>2</sup), [2212017@student.iteba.ac.id³](mailto:2212017@student.iteba.ac.id<sup>3</sup), [agung.derma.wan29@gmail.com⁴](mailto:agung.derma.wan29@gmail.com<sup>4</sup)

ABSTRAK

Data UDD PMI menunjukkan bahwa banyak pendonor yang wajib memenuhi persyaratan guna dijadikan pendonor darah. Selama ini, calon pendonor dipilih secara manual untuk mendonorkan darah. Sistem informasi sekarang belum menggali lebih banyak informasi dari banyak data guna menjadi pengetahuan. Wajib ada konsolidasi organisasi serta penilaian yang berkelanjutan mengenai kinerja PMI guna menangani permasalahan sosial kemanusiaan. Guna mengidentifikasi individu layak ataupun tidak guna mendonorkan darahnya, pengimplementasian data mining melalui metode klasifikasi menggunakan Algoritma Decision Tree C4.5 dapat mempergunakan hasil variable yang bernilai kontinu ataupun dikrit, misalnya umur, BB, serta tekanan darah (sistolik serta diastolik). Dengan mempergunakan data mining Algoritma Decision Tree C4.5, calon kelayakan pendonor darah bisa diklasifikasikan berdasar umur, BB, hemoglobin, serta tekanan darah. Variabel yang paling menentukan yaitu hemoglobin dengan nilai gain tertinggi (0.861212618).

Kata Kunci :

ABSTRACT

PMI UDD data shows that many donors must meet the requirements to become blood donors. So far, potential donors are selected manually to donate blood. Current information systems have not extracted more information from a lot of data to become knowledge. Organizational simplification and continual assessment of PMI's efficacy in addressing social and humanitarian issues are necessary. To determine whether someone is worthy or not to donate blood, The C4.5 Decision Tree Algorithm may be used to apply data mining to a classification approach for variable outcomes, such as age, weight, and blood pressure, that have continuous or discrete values. (both diastolic and systolic). The C4.5 Decision Tree Algorithm in data mining may classify possible blood donor eligibility by considering factors such as age, weight, hemoglobin levels, and blood pressure. Hemoglobin is the most influential variable, with the greatest gain value of 0.861212618.

Keywords :

Disubmit: 18 Maret 2024 Info Artikel : Diterima : 20 Mei 2024
Direview: 19 April 2024

Copyright © 2024 – Journal UPU. All rights reserved.

1. PENDAHULUAN

Unit Donor Darah PMI Kab. Kampar yaitu sebuah entitas ataupun organisasi yang bertugas menyediakan layanan donor darah. Organisasi ini bertanggung jawab mengatur, mengelola, dan menjalankan aktivitas donor darah guna memenuhi keperluan darah di wilayah Kabupaten Kampar. Mereka berperan penting dalam memastikan ketersediaan darah yang aman dan mencukupi bagi masyarakat yang membutuhkan, baik untuk keperluan medis darurat maupun untuk pasien yang memerlukan transfusi darah secara berkala. Tiap orang yang secara sukarela datang ke PMI dan menerima kantong darah dari masyarakat harus memenuhi syarat untuk dapat mendonorkan

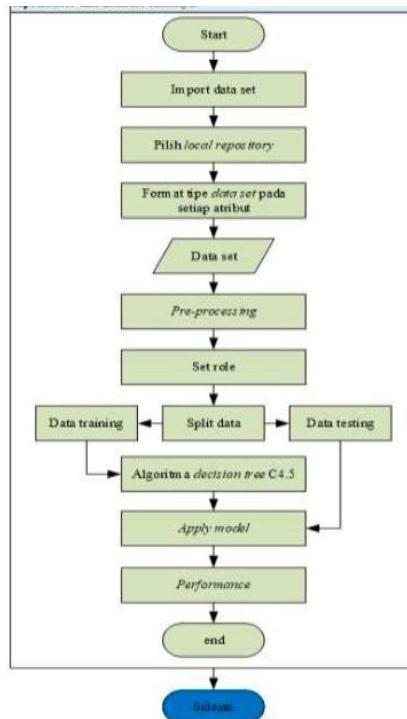
darahnya. Usia, BB, serta tekanan darah (baik sistolik maupun diastolik) akan dipergunakan sebagai variabel pendukung dalam pendekatan data mining penelitian ini, yang akan diklasifikasikan mempergunakan algoritma c4.5. Klasifikasi yaitu komponen penambangan data, di mana beberapa model diterapkan pada berbagai skenario dan menciptakan taraf akurasi yang berbeda.

Penambangan data yaitu proses menemukan pola baru dalam himpunan data begitu besar, terutama pola yang muncul oleh kemajuan kecerdasan buatan (AI). Algoritma C4.5 memanfaatkan rasio amplifikasi. Tindakan penting seperti mendonorkan darah dapat menyelamatkan banyak nyawa. Darah disuplai oleh Palang Merah Indonesia (PMI) untuk keperluan medis. Kebutuhan darah meningkat akhir-akhir ini karena pertumbuhan populasi dan layanan kesehatan yang lebih baik. Pasokan darah yang tepat waktu dan memadai yakni permasalahan bagi PMI. Analisis data dapat membantu memecahkan masalah ini. Donor darah dapat dikategorikan melalui analisis data berdasarkan sejumlah faktor, termasuk golongan darah, usia, dan riwayat kesehatan, serta seberapa sering mereka mendonor. Melalui mengelompokkan produk darah, PMI akan dapat lebih mempersiapkan diri dalam mengendalikan persediaan darah dan menginspirasi lebih banyak orang untuk mendonorkan darahnya. Menentukan karakteristik yang berdampak, antara lain usia, jenis kelamin, riwayat kesehatan, hasil pemeriksaan laboratorium, dll. Melalui mempergunakan metode Decision Tree C4.5 serta data calon donor tersedia, kembangkan model prediksi yang dapat dipercaya. membuat manajemen suplai darah lebih efisien dan tepat waktu. meningkatkan efektivitas kegiatan donor darah dengan tujuan yang lebih tepat sasaran. meningkatkan pemahaman dan kemahiran seseorang dalam analisis data, khususnya untuk kegunaan praktis di bidang medis. Hal ini akan menjadi referensi untuk penelitian masa depan mengenai penerapan data analitis dalam industri kesehatan. Verifikasi jumlah darah yang ada dan sesuai kebutuhan.

2. METODE

Pendekatan Pohon Keputusan, sebuah model ataupun fungsi yang dipergunakan untuk mendeskripsikan dan mengklasifikasikan konsep dan kelas data dengan tujuan memproyeksikan kelas untuk mengidentifikasi calon donor darah untuk mengukur kompatibilitas donor, dipergunakan oleh peneliti sebagai alat klasifikasi.

Dengan mempergunakan teknik ini, peneliti membuat rancangan dalam flowchart seperti dibawah ini:



Gambar 1 Menunjukkan Urutan Langkah Penelitian

A. Analitik Data

Pembersihan Transformasi Analisis Data Data dikumpulkan, dianalisis, dan diinterpretasikan untuk memberikan data analitis—informasi yang berguna. Data mentah diubah menjadi wawasan melalui analisis data, yang kemudian dapat diterapkan pada pengambilan keputusan, peningkatan efisiensi, dan penciptaan barang dan jasa baru. Berbagai instrumen, sistem, dan prosedur dipergunakan dalam analisis data analitis untuk mengidentifikasi pola dan menyelesaikan masalah. Selain itu, analisis data dapat memengaruhi prosedur bisnis, meningkatkan penilaian, dan memacu ekspansi. Bisnis dapat lebih memahami dan melihat operasi dan penawaran mereka berkat analisis data. Bisnis dapat memperoleh wawasan tentang permintaan dan pengalaman pelanggan mereka dengan mempergunakan data analitis ini. Tindakan dan wawasan dihubungkan oleh perubahan paradigma dari data ke tindakan.

B. Data Mining

Tujuan dari penambangan data yaitu untuk mengekstrak informasi penting dari data mempergunakan metode pengumpulan dan pemrosesan data. Angka-angka ataupun informasi yang memiliki beberapa kegunaan dapat menjadi salah satu bentuk data yang dikumpulkan. Data ini dapat dikumpulkan dan diproses dengan bantuan perangkat lunak yang dirancang untuk mengolah statistik dan angka. Perangkat lunak ini memungkinkan pengumpulan data yang lebih efisien dan akurat, serta pemrosesan yang cepat untuk menghasilkan informasi yang berguna. Dengan

teknologi terkini, perangkat lunak ini juga dapat menangani berbagai jenis data dan melakukan analisis yang kompleks, sehingga mempermudah pengambilan keputusan berbasis data.

C. *Algoritma decision tree*

Pendekatan klasifikasi yang biasanya dipergunakan yaitu algoritma Decision Tree C4.5. Teknik ini mengambil masukan dalam jumlah besar yang disebut atribut guna melakukan klasifikasi berdasar pola pada data. Karena Pohon Keputusan, yang menciptakan prediksi yang bisa diamati, memberikan keuntungan visualisasi, pendekatan ini begitu mudah beradaptasi. Metode C4.5 disusun seperti pohon, melalui daun yang menunjukkan kelas, cabang yang mewakili hasil kualitas yang diuji, serta simpul inti yang mewakili atribut. Untuk memastikan tidak ada cabang baru yang terbentuk, metode perhitungan entropi diterapkan secara rekursif untuk memilih atribut terbaik.

Entropi dapat dihitung mempergunakan rumus berikut. $\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n p_i (-\log_2 p_i)$

Ket:

S = Kumpulan Kasus n = Total partisi S p_i =
kemungkinan diperoleh oleh total kelas yang
dibagi.

Gain dipergunakan guna menguji kelengkapan yang nantinya ditentukan dalam proses pemecahan node dalam algoritma pembelajaran mesin ataupun pohon keputusan. Dalam konteks ini, gain mengukur seberapa baik suatu atribut memisahkan data berdasarkan target klasifikasi, dengan atribut yang memiliki gain tertinggi dipilih untuk membagi node. Hal ini membantu dalam menentukan atribut mana yang paling efektif untuk dipergunakan dalam pemecahan node, sehingga pohon keputusan yang dihasilkan lebih akurat dan efisien dalam melakukan klasifikasi. Persamaan ini dipergunakan guna menghitung gain atribut ke-A: $\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n (s^i/s^2) * \text{Entropy}(S^i)$

Ket:

S =
Kumpulan
kasus A =
Atribut n =
Total
atribut

$|S_i|$ = Total partisi ke -i $|S|$ = total kasus pada S

D. *Rapid Miner*

RapidMiner yaitu alat pemrosesan data mining sumber terbuka berbasis Java yang bekerja dengan sejumlah sistem operasi berbeda. RapidMiner berfokus pada analisis teks dalam penambangan teks, yang memerlukan ekstraksi pola besar-besaran dari kumpulan data dan integrasi dengan database, kecerdasan buatan, dan teknik statistik.

3. HASIL DAN PEMBAHASAN

A. *Analisis Data*

Sumber data dipergunakan pada studi didapat oleh PMI, dataset tersedia memiliki 4 atribut yang berbeda yaitu Nama Pendoron, Tekanan Darah, Berat Badan, Dan Umur. Setelah pengumpulan data dapat mengalami proses pengolahan awal untuk membersihkan data dari duplikasi, kesalahan, ataupun nilai yang hilang agar dapat dipergunakan dalam analisis lebih lanjut.

Tabel 1 Memperllihatkan Dataset Calon Pendoron Darah

No	Nama Pendoron	Tekanan Darah	Berat Badan	Umur
1	Budi Setiawan	Normal	70	32

2	SitiRahmawati	Rendah	65	28
3	Joko Susilo	Normal	68	30
4	Dewi Anggraini	Normal	63	26
5	Bambang Santoso	Rendah	75	35
6	Rini Sari	Normal	60	27
7	Adi Prasetyo	Normal	60	31
8	Dian Fitriani	Normal	67	29
9	Eko Wijaya	Rendah	69	33
10	Yuni Hartati	Normal	66	28
...				
91	Rini Sari	Normal	60	27
92	Adi Prasetyo	Normal	72	31
93	Dian Fitriani	Normal	67	29
94	Eko Wijaya	Rendah	69	33
95	Yuni hartati	Normal	66	28
96	Bambang Supriyanto	Normal	70	32
97	Siti Hidayah	Normal	64	27
98	Heru Setiawan	Rendah	68	30
99	Rina Permata	Normal	67	29

B. Pengolahan data

Proses penerapan algoritma Decision Tree C4.5 guna mengevaluasi taraf kecocokan calon pendonor. Awalnya, data calon pendonor diolah mempergunakan Tools RapidMiner. Pasca algoritma C4.5 diterapkan, hasilnya berupa pohon keputusan yang memvisualisasikan data serta aturan-aturan yang terbentuk dari analisis tersebut. Dengan kata lain, algoritma ini membantu mengidentifikasi serta memvisualisasikan pola dalam data guna menentukan kecocokan calon pendonor berdasar kriteria yang telah ditentukan.

a. Format tipe *data set* pendonor

Pada tahap ini, kumpulan data Pendonor akan diubah menjadi informasi dengan mengubah tipe data ataupun lebih sederhana dengan melakukan konversi ataupun

pengkodean untuk mengubah data menjadi format tertentu yang dapat dipergunakan, seperti berikut :

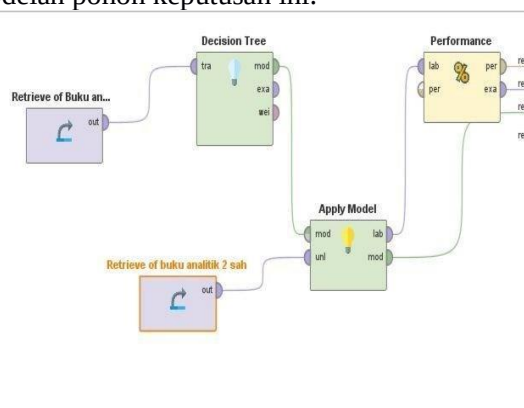
- 1) Nama Pendonor = Id
- 2) Tekanan Darah = Label
- 3) Berat Badan = Integer
- 4) Umur = Integer

Id	Nama	Umur	Berat Badan	Tekanan Darah
1	Andi	25	70	120
2	Budi	30	80	130
3	Cici	28	65	110
4	Dani	35	90	140
5	Evi	22	55	100
6	Fani	32	75	125
7	Gani	27	60	115
8	Hani	33	85	135
9	Iani	29	70	120
10	Jani	31	80	130
11	Kani	26	65	110
12	Lani	34	95	145
13	Mani	24	50	95
14	Nani	36	100	150
15	Oani	23	58	105
16	Pani	37	105	155
17	Qani	21	52	98
18	Rani	38	110	160
19	Sani	20	48	92
20	Tani	39	115	165

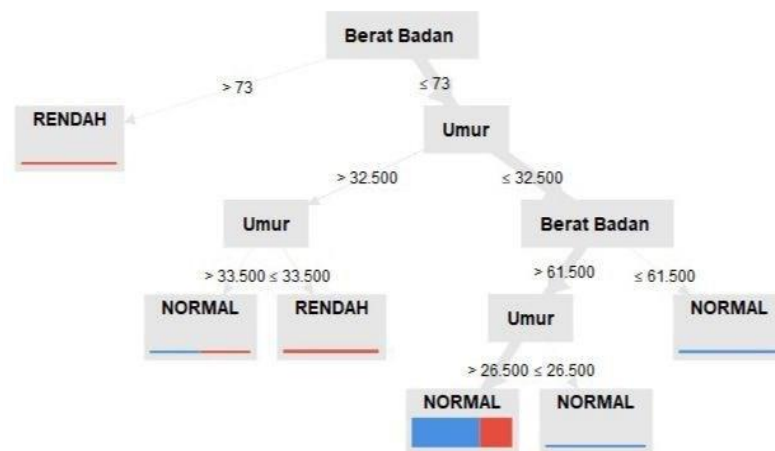
Gambar 2 format tipe *DataSet* setiap Atribut

b. Permodelan Algoritma *Decision Tree Series C4.5*

Pada tahap pemodelan algoritma decision tree C4.5, rasio kriteria pendapatan dipergunakan untuk menganalisis data set yang berkaitan dengan kesehatan jantung. Data set ini mencakup empat karakteristik dari pendonor darah serta satu label yang berfungsi sebagai hasil akhir prediksi. Proses ini bertujuan untuk mengembangkan model prediksi yang dapat membantu dalam menentukan kondisi jantung berdasarkan karakteristik pendonor darah yang telah ditentukan. Data pelatihan dipergunakan untuk menentukan pemodelan pohon keputusan ini.



Gambar 3 Memperlihatkan Tampilan menghubungkan *Dataset*, *Decision tree*, *Aply Model*, serta *Performance*



Gambar 4 Memperlihatkan *Decision Tree*

4. KESIMPULAN

Temuan penelitian yang meliputi tipe data nominal dan numerik serta didasarkan pada 100 titik data dari dataset yang tersedia, dapat disimpulkan dari penjelasan sebelumnya. Tiga atribut tersebut terdiri dari satu atribut label dan empat atribut fitur. Data ini diproses melalui suatu algoritma yang menjalankan prosedur klasifikasi. Pohon keputusan dari seri C4.5. untuk mendapatkan perhitungan yang tepat secara real time saat mempergunakan program Rapidminer. Pemrosesan awal, penetapan peran (data terpisah = data pelatihan : data pengujian), pemodelan metode rangkaian pohon keputusan C4.5 pada data pelatihan, penerapan model pada data pengujian, dan evaluasi kinerja model sebagai pengujian untuk memastikan kemanjuran model semuanya sudah selesai. keakuratan hasil tes. Kemudian, dengan mempergunakan proses ini, kami dapat menentukan apa yang menarik calon kontributor. Dijadikan atribut utama yang dipergunakan menjadi tahap pengambilan putusan ataupun tahap klasifikasi dari kecocokan calon pendonor, dimana atributnya dari berat badan dan jika berat badan jika berat badannya lebih besar dari 73 maka tekanan darahnya rendah dan jika Jika angka kesialannya kurang dari 73, maka orang tersebut akan kembali ke umur semula lagi. Jika angka keberuntungannya lebih dari 33, maka darah akan berjalan normal; jika angka keberuntungannya kurang dari 33, maka darahnya akan semakin parah. Jika angka keberuntungannya kurang dari 61, maka orang tersebut akan kembali ke usia semula. Jika angka keberuntungan lebih dari 26 maka darah akan berjalan normal.

REFERENSI

- [1] Gunawan, I. M., Saraswati, I. D., Agung, I. D., & Putra, I. P. (2023). Klasifikasi penyakit jantung mempergunakan algoritma Decision Tree C4.5 dengan RapidMiner. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 73-83.
- [2] Hutrim, H. P., & Nurahman. (2020). Pengimplementasian data mining guna penentuan kelayakan calon pendonor darah mempergunakan teknik Naïve Bayes (Studi kasus PMI Kabupaten Kotawaringin Timur). *JUISI*.
- [3] Irawan, Y. (2021). Pengimplementasian algoritma Decision Tree C4.5 bagi prediksi kelayakan calon pendonor darah melalui klasifikasi data mining. *Jurnal Teknologi Informasi dan Multimedia*, 181-189.
- [4] Pascalina, D., Widhiastono, R., & Juliane, C. (2023). Pengujian kesiapan transformasi digital Smart City mempergunakan aplikasi RapidMiner. *Technomedia Journal (TMJ)*.

- [5] Rofiani, R., Oktaviani, L., D. V., & T. H. (n.d.). Pengimplementasian teknik klasifikasi Decision Tree bagi prediksi kanker paru-paru mempergunakan algoritma C4.5. *Jurnal TEKNO KOMPAK*, 126-139.
- [6] Yulisman, Y., Wahyuni, R., & Irawan, Y. (2020). Aplikasi pengarsipan surat masuk serta surat keluar berbasis web pada SMP Negeri 32 Pekanbaru. *Jurnal Teknologi Sistem Informasi dan Aplikasi*, 3(4), 252. <https://doi.org/10.32493/jtsi.v3i4.7345>
- [7] Mustika, B. (2020). pemakaian text mining pada Twitter melalui algoritma K-Means Clustering sebagai dasar kebijakan marketing biro perjalanan wisata. *Jurnal Ilmu Komputer*, 9(2), 138–147. <https://doi.org/10.33060/JIK/2020/Vol9.Iss2.180>
- [8] Irawan, Y. (2019). Pengimplementasian data mining bagi evaluasi data penjualan mempergunakan teknik clustering serta algoritma hirarki divisive. *Jurnal Teknologi Informasi Universitas Lampung*, 3(1), 13–20.
- [9] Magdalena, L., & Mulyasari, H. (2018). Rancangan SI PMI melalui menyatukan data pendonor serta stok darah antar cabang PMI di Wilayah III Cirebon. Retrieved from <https://doi.org/10.1002/9781118874059>
- [10] Al, F., & Heru, T. (2015). *MSDM sebagai dasar meraih keunggulan bersaing*. UPP STIM YKPN.
- [11] Yogyakarta Andini, P. (2010). Evaluasi penganggaran biaya rumah sakit (Studi kasus Rumah Sakit Umum Palang Merah Indonesia (RSU PMI) Bogor). *Jurnal Akuntansi Manajemen*.
- [12] Achsan, M. (2014). Insidensi infeksi menular lewat transfusi darah (IMLTD) di Unit Donor Darah PMI Kota Semarang. *Medica Hospitalia-Journal of Clinical Medicine*, 2(2).
- [13] Badan Pengawas Obat serta Makanan Republik Indonesia. (2016). Draft pedoman CPOB untuk unit penyedia darah. BPOM RI. Jakarta.
- [14] Custer, B., Janssen, M. P., & Alliance of Blood Operators Risk-Based DecisionMaking (RBDM) Initiative. (2015). Health economics and outcomes methods in riskbased decision-making for blood safety. *Transfusion*, 55(8), 2039-2047.
- [15] Agustina, H. (2015). *Kualitas pelayanan Palang Merah Indonesia (PMI) Provinsi Lampung (Studi pada Unit Donor Darah Cabang Pembina Provinsi Lampung)*. Doctoral dissertation, Fakultas Ilmu Sosial dan Ilmu Politik.