



Klasifikasi pendonor darah potensial menggunakan pendekatan algoritme pembelajaran mesin

Classification of potential blood donors using machine learning algorithms approach

Merinda Lestandy^{1*)}, Lailis Syafa'ah²⁾, Amrul Faruq²⁾

¹⁾Program Studi D3 Teknologi Elektronika, Fakultas Teknik, Universitas Muhammadiyah Malang
Jl. Raya Tlogomas No. 246, Tlogomas, Malang, Jawa Timur, Indonesia 65144

²⁾Program Studi Teknik Elektro, Fakultas Teknik, Universitas Muhammadiyah Malang
Jl. Raya Tlogomas No. 246, Tlogomas, Malang, Jawa Timur, Indonesia 65144

Cara sitasi: M. Lestandy, L. Syafa'ah, and A. Faruq, "Klasifikasi pendonor darah potensial menggunakan pendekatan algoritme pembelajaran mesin," *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 3, pp. 217-221, 2020. doi: [10.14710/jtsiskom.2020.13619](https://doi.org/10.14710/jtsiskom.2020.13619), [Online].

Abstract – Blood donation is the process of taking blood from someone used for blood transfusions. Blood type, sex, age, blood pressure, and hemoglobin are blood donor criteria that must be met and processed manually to classify blood donor eligibility. The manual process resulted in an irregular blood supply because blood donor candidates did not meet the criteria. This study implements machine learning algorithms includes kNN, naïve Bayes, and neural network methods to determine the eligibility of blood donors. This study used 600 training data divided into two classes, namely potential and non-potential donors. The test results show that the accuracy of the neural network is 84.3 %, higher than kNN and naïve Bayes, respectively of 75 % and 84.17 %. It indicates that the neural network method outperforms comparing with kNN and naïve Bayes.

Keywords – potential blood donor; kNN; Naïve Bayes; neural network; machine learning

Abstrak – Donor darah merupakan proses pengambilan darah dari seseorang dan digunakan untuk transfusi darah. Golongan darah, jenis kelamin, usia, tekanan darah, dan hemoglobin merupakan kriteria pendonor darah yang harus dipenuhi dan diproses secara manual untuk menentukan kelayakan pendonor darah. Proses manual tersebut mengakibatkan persediaan darah yang bersifat tidak tetap karena kandidat pendonor darah yang tidak sesuai dengan kriteria. Penelitian ini menggunakan metode kNN, naïve Bayes dan neural network untuk menentukan kelayakan pendonor darah. Data latih yang digunakan pada penelitian ini sejumlah 600 dataset yang dibagi menjadi dua kategori, yaitu potensial dan non potensial. Nilai akurasi neural network, yaitu sebesar 84,3 %, lebih

tinggi dibandingkan kNN dan naïve Bayes, yaitu sebesar 75 % dan 84,17 %. Hasil ini menunjukkan bahwa metode neural network terbukti mengungguli jika dibandingkan kNN dan naïve Bayes.

Kata kunci – pendonor darah potensial; kNN; naïve Bayes; neural network; pembelajaran mesin

I. PENDAHULUAN

Informasi yang diberikan oleh 164 negara ke standar *World Health Organization* (WHO) tentang keselamatan darah menunjukkan bahwa lebih dari 92 juta donor darah dikumpulkan setiap tahun di seluruh dunia. Dari jumlah tersebut, diperkirakan 1,6 juta unit dibuang karena terdapat darah yang terinfeksi. Di banyak negara, kriteria pemilihan donor masih didasarkan pada tradisi dan praktik adat daripada berdasarkan bukti-bukti keilmuan dan kriteria pendonor, serta tanpa mempertimbangkan profil populasi donor umum dan potensial [1], [2].

Donor darah merupakan proses pengambilan darah dari seseorang secara sukarela untuk disimpan di bank darah yang kemudian digunakan untuk transfusi darah bagi pasien yang membutuhkan [3]. Pemenuhan kebutuhan darah sangat penting untuk meningkatkan kualitas pelayanan kesehatan dan menyelamatkan nyawa seseorang. Berdasarkan pedoman WHO, darah dari pendonor tidak dapat digunakan setelah 42 hari [4]. Jika kondisi darah rusak ketika diterima pasien, maka dapat membahayakan pasien.

Persediaan darah sering kali bersifat tidak tetap karena jumlah pendonor yang naik turun. Hal ini menjadi sebuah permasalahan ketika jumlah permintaan darah meningkat namun persediaan darah menipis. Hal tersebut disebabkan kandidat pendonor darah sering kali tidak sesuai dengan kriteria. Selama proses donor darah, darah diambil secara langsung dari pendonor dan disimpan dalam bank darah. Hal demikian harus diatur dan dikelola secara efektif untuk kebutuhan darurat di rumah sakit.

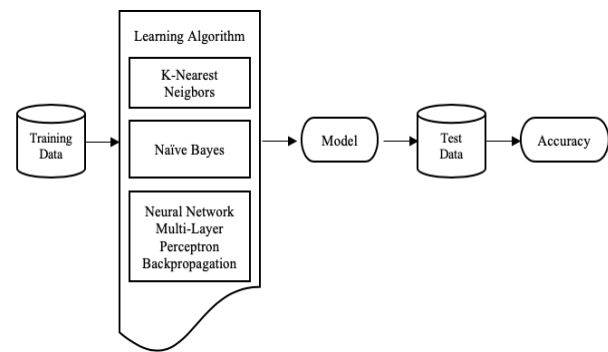
^{*)}Penulis korespondensi (Merinda Lestandy)
Email: merindalestandy@umm.ac.id

Pada kenyataannya, banyak pendonor darah yang gagal melakukan donor darah untuk kedua kalinya disebabkan beberapa alasan. Beberapa di antaranya adalah tidak memenuhi syarat sebagai pendonor, kondisi tubuh yang kurang sehat, tekanan darah yang tidak normal, dan beragam alasan kesehatan yang lain yang berkaitan dengan kualifikasi sebagai pendonor darah. Strategi untuk menemukan dan memahami alasan kegagalan donor darah menjadi sangat penting untuk meningkatkan pasokan darah untuk waktu berikutnya, salah satunya adalah dengan memprediksi perilaku pendonor darah yang potensial untuk mendonorkan darahnya kembali atau tidak mendonorkan menggunakan model *recency*, *frequency*, *monetary*, *time*, dan *churn probability* (RFMTC) dan metode urutan Bernoulli seperti dalam [5].

Kajian tentang penentuan pendonor darah yang potensial telah dilakukan dalam [3], [6], [7]. Analisis data menggunakan beragam metode untuk mengklasifikasikan data, di antaranya naïve Bayes (NB) dalam [3], [6], *decision tree* [3], *k-nearest network* (kNN) dan *neural network* (NN) dalam [6], dan pembelajaran mesin dalam [7]. Metode-metode tersebut cukup efektif dalam menghasilkan klasifikasi kelayakan pendonor darah dengan akurasi mencapai 81 % untuk metode naïve Bayes [3], sedangkan menggunakan NN dan kNN menghasilkan akurasi 80 % dan 71 % [6].

Algoritme kNN melakukan klasifikasi dengan menghitung jarak objek tak berlabel (tidak memiliki kelas) ke objek berlabel, menentukan label kelas *k* (jumlah tetangga terdekat), dan menggunakannya untuk menentukan label kelas objek tersebut seperti pada kajian [8]-[10]. Metode naïve Bayes memprediksi probabilitas keanggotaan suatu kelas berdasarkan data RFMTC dan membandingkan nilai probabilitas dengan uji. Klasifikasi ini menggunakan metode statistik untuk memprediksi kemungkinan keanggotaan berdasarkan label suatu data [11]. Dalam NN, neuron-neuron diwakili dengan node-node yang berisi nilai yang memiliki bobot tertentu dan disusun berlapis menggunakan banyak lapisan tersembunyi sehingga dapat melakukan klasifikasi berdasarkan data yang sudah dilatih ditambah penggunaan fungsi aktivasi [12].

Meskipun terdapat berbagai teknik serta metode klasifikasi dan prediksi dalam menentukan pendonor darah potensial, namun tingkat akurasi dan aplikasi metode untuk kasus secara umum masih menarik untuk dikaji, dievaluasi, dan ditingkatkan kemampuannya. Penelitian ini mengkaji penerapan kNN, naïve Bayes dan *neural network multi-layer perceptron backpropagation* (MLP-BP NN) sebagai algoritme pembelajaran mesin dalam menentukan calon pendonor potensial. Proses dan hasil dari metode kNN, naïve Bayes dan MLP-BP NN pada kajian ini dioptimalkan tingkat akurasi dan dianalisis untuk menentukan metode yang tepat dalam klasifikasi calon pendonor darah potensial dan mengevaluasi ketiga metode tersebut dengan menggunakan data RFMTC.



Gambar 1. Skematik metode pembelajaran mesin

II. METODE PENELITIAN

Penelitian eksperimen ini bertujuan untuk mencari algoritme paling tepat untuk mengklasifikasikan calon pendonor darah potensial. Dataset diperoleh dari *University of California, Irvine (UCI) machine learning datasets* dengan total sebanyak 600 data. Data tersebut dianalisis untuk menentukan algoritme yang tepat untuk digunakan dalam klasifikasi calon pendonor darah potensial.

Tahapan pada penelitian ini ditunjukkan pada **Gambar 1**. Pemodelan dilakukan dengan menggunakan *training data* (data latih). Data latih ini diperlukan sebagai data pembelajaran untuk melakukan klasifikasi. Setiap sampel dari data latih memiliki atribut dan kelas (label). Data dengan atribut dan label ini digunakan oleh algoritme pembelajaran untuk menghasilkan model pembelajaran. Model yang didapat dan sudah dilatih digunakan untuk melakukan klasifikasi dan prediksi dengan menggunakan *testing data* (data uji). Validasi dinyatakan dalam matriks konfusi untuk menghitung akurasi sistem secara keseluruhan.

Algoritme pembelajaran pada kajian ini melibatkan kNN, naïve Bayes dan MLP-BP NN [13]-[15]. Parameter *k* dalam kNN ditentukan bernilai 4. Penghitungan jarak Euclidean objek terhadap data latih yang diberikan dinyatakan dalam Persamaan 1.

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Penghitungan peluang hipotesis pada naïve Bayes dinyatakan dalam Persamaan 2. Parameter $P(H|X)$ menyatakan peluang hipotesis *H* berdasarkan kondisi *X* dengan *X* menyatakan data latih dengan label yang diketahui dan *H* data dengan label *C*. $P(H)$ menyatakan peluang dari hipotesis *X*, $P(X)$ peluang dari *X* yang diamati, dan $P(H|X)$ peluang *X* berdasarkan kondisi hipotesis *H*. Pada 600 RFMTC-UCI dataset diketahui bahwa status pendonor darah potensial dan nonpotensial secara berturut-turut adalah 164 dan 436. Dengan menggunakan Persamaan 2, nilai $P(H|X)$ untuk $P(\text{potensial})$ adalah 2,21 dan $P(\text{nonpotensial})$ adalah 1,15. Data latih diklasifikasikan ke dalam kelas

potensial karena kelas potensial memiliki nilai yang lebih besar daripada kelas non potensial.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (2)$$

A. Dataset

Dataset diperoleh dari UCI Machine Learning dataset dengan atribut RFMTC yang merupakan pengembangan dari metode *Recency*, *Frequency*, *Monetary* (RFM) [16]. Model pembelajaran mesin kNN, naïve Bayes, dan MLP-BP NN pada kajian ini menggunakan data awal 200. Penambahan data menjadi 600 dataset dilakukan karena jika dataset kecil, maka kemungkinan kesalahan berdampak besar pada hasil analisis dan pemrosesan data [17].

Atribut RFMTC ini meliputi jarak atau jumlah sejak bulan terakhir kali menyumbang darah (*recency*), total berapa kali mendonor (*frequency*), total sumbangan darah dari pendonor dalam cc (*monetary*), total bulan menyumbangkan darah sejak pertama kali (*time*), dan *churn probability* untuk mempresentasikan pendonor akan mendonorkan darahnya kembali atau tidak. *Churn probability* bernilai 1 untuk menyatakan donor darah (potensial) yang jumlahnya sebesar 164 dan 0 untuk menyatakan tidak menyumbang darah (nonpotensial) sebesar 436. Atribut RFMTC ini ditunjukkan pada Tabel 1.

B. Pengujian

Ukuran kinerja yang dikaji adalah akurasi, sensitifitas, spesifikasi, dan *area under the curve* (AUC). Ketiga perhitungan tersebut dinyatakan dengan menggunakan matriks konfusi yang terdiri dari empat aspek, yaitu *True Positive* (TP), *False Positive* (FP), *False Negative* (FN) dan *True Negative* (TN).

Akurasi menunjukkan seberapa baik klasifikasi pendonor darah potensial dan nonpotensial. Sensitifitas menunjukkan seberapa baik metode KNN dan Naïve Bayes dapat memprediksi dengan benar pendonor darah sebenarnya yang telah menyumbang. Spesifikasi memungkinkan untuk mengukur seberapa baik ketiga metode pembelajaran mesin tersebut dapat memprediksi nonpotensial di antara yang tidak menyumbang darah. Klasifikasi akurasi menggunakan AUC dinyatakan dalam lima kategori, yaitu *excellent* (0,90–1,00), *good* (0,80–0,90), *fair* (0,70–0,80), *poor* (0,60–0,70), dan *failure* (0,50–0,60) [18].

III. HASIL DAN PEMBAHASAN

Kajian ini melakukan klasifikasi calon pendonor darah potensial dan melihat akurasi metode kNN, naïve Bayes dan MLP-BP NN dengan menggunakan dataset RFMTC. Hasil klasifikasi dengan ketiga metode tersebut dibandingkan untuk memperoleh metode yang tepat digunakan untuk melakukan klasifikasi calon pendonor darah potensial.

Tabel 1. Atribut RFMTC

Recency (months)	Frequency (times)	Monetary (cc. blood)	Time (months)	Donated
2	50	12500	98	1
0	13	3250	28	1
1	16	4000	35	1
2	20	5000	45	1
1	24	6000	77	0

Tabel 2. Perhitungan kNN

R	F	M	T	d(x,y)	Jarak	Label
2	50	12500	98	12000,47	5	0
0	13	3250	28	2750,13	1	0
1	16	4000	35	3500,18	2	1
2	20	5000	45	4500,23	3	1
1	24	6000	77	550,58	4	1

Tabel 3. Kinerja kNN untuk 200 data

Nilai Uji	Presisi	Sensitifitas	Skor f1	Akurasi
0,2	0,70	0,70	0,70	0,70
0,3	0,71	0,72	0,71	0,72
0,4	0,67	0,68	0,67	0,70

Tabel 4. Kinerja kNN untuk 600 data

Nilai Uji	Presisi	Sensitifitas	Skor f1	Akurasi
0,2	0,71	0,72	0,71	0,72
0,3	0,73	0,71	0,71	0,71
0,4	0,69	0,69	0,69	0,69

Komposisi data latih dan data uji pada algoritme KNN dan NB dilakukan sebanyak 3 kali percobaan, yaitu dengan menentukan nilai uji 0,2, 0,3 dan 0,4. Komposisi tersebut masing-masing berturut turut adalah 80 % data latih dan 20 % data uji, kemudian 70 % data latih dan 30 % data uji, serta 60 % data latih dan 40 % data uji.

Perhitungan klasifikasi pendonor darah potensial dengan metode kNN dan k bernilai 4 ditunjukkan pada Tabel 2. Data dapat diklasifikasikan ke dalam kelas potensial karena jumlah kategori potensial (label=1) lebih besar dibandingkan nonpotensial (label=0). Hasil eksperimen berdasarkan beragam komposisi nilai uji untuk data latih dan data uji dapat dilihat pada Tabel 3 untuk 200 data dan Tabel 4 untuk 600 data. Klasifikasi dengan metode kNN memperoleh akurasi yang hampir sama pada dataset yang berbeda. Hal ini disebabkan kNN ini memiliki kelemahan dalam cara untuk meningkatkan akurasinya terhadap nilai k yang terbaik [8].

Dengan menggunakan data yang sama pada Tabel 1, perhitungan klasifikasi pendonor darah potensial menggunakan metode naïve Bayes dengan dengan 600 dataset ditunjukkan dalam Tabel 5. Pada tahap akhir, perbandingan antara nilai $P(X|C_i)$ potensial dan nonpotensial dilakukan. Data latih diklasifikasikan ke dalam kelas potensial karena kelas potensial memiliki

nilai yang lebih besar daripada kelas nonpotensial. Hasil eksperimen nilai tes ditunjukkan Tabel 6 untuk 200 data dan Tabel 7 untuk 600 data.

Model klasifikasi donor darah dengan menggunakan MLP-BP NN dievaluasi dengan laju pembelajaran sebesar 0,1. Klasifikasi model menghasilkan presisi, sensitifitas, skor f1 dan akurasi seperti ditunjukkan pada Tabel 8. MLP-BP NN mampu mengklasifikasikan dengan nilai presisi 0,53, sensitifitas 0,41, skor f1 0,46 dan nilai akurasi sebesar 0,47 dengan 200 data. Dengan menambahkan data menjadi 600 data, presisi menjadi lebih besar, yaitu 0,81, nilai sensitifitas 0,88, skor f1 0,83 dan nilai akurasi sebesar 0,84 sesuai dengan [17].

Perbandingan kinerja ketiga metode dengan menggunakan 200 dan 600 data ditunjukkan pada Tabel 9. Jika dibandingkan dengan [3] dan [6], akurasi kNN, naïve Bayes dan MLP-BP NN untuk prediksi pendonor potensial ditunjukkan pada Tabel 10. Semua metode tersebut menunjukkan hasil yang bagus dan tingkat akurasi yang relatif tidak jauh berbeda. Namun, dengan metode kNN pada kajian ini untuk k=4 menghasilkan akurasi yang lebih tinggi sebesar 72,5 % daripada [6] sebesar 71 % dengan jumlah data yang lebih sedikit.

Selain itu, dengan data yang sama dan jumlah yang lebih sedikit pada metode naïve Bayes, kajian ini juga menghasilkan akurasi yang lebih tinggi sebesar 84,17 % dibandingkan pada [3], [6], yaitu berturut-turut dengan nilai akurasi 81,5 % dan 71 %. Metode MLP-BP NN juga menghasilkan akurasi yang lebih tinggi yaitu sebesar 84,3 % dibandingkan [6], yaitu sebesar 80 %.

Dari hasil simulasi dan dengan pemilihan nilai parameter yang tepat, implementasi algoritme kNN, naïve Bayes dan MLP-BP NN untuk penentuan pendonor potensial mengalami perbaikan yang ditunjukkan dengan nilai akurasi yang lebih tinggi. Dengan hasil akurasi yang lebih baik, kajian ini memperkuat hasil kajian [6] yang menyatakan bahwa teknik algoritme NN mampu mengungguli naïve Bayes, kNN pada klasifikasi penentuan pendonor darah potensial.

IV. KESIMPULAN

Klasifikasi pendonor darah potensial dapat dilakukan menggunakan metode kNN, naïve Bayes dan NN dengan model RFMTC dari 600 dataset UCI. Klasifikasi calon pendonor menggunakan NN menunjukkan akurasi yang lebih baik, yaitu sebesar 84,3 %, dibandingkan naïve Bayes dan kNN dengan akurasi masing-masing 84,167 % dan 75 %.

UCAPAN TERIMA KASIH

Tim peneliti mengucapkan terima kasih dan penghargaan kepada Fakultas Teknik, Universitas Muhammadiyah Malang atas dukungan terselenggaranya pekerjaan ini melalui skema Pusat Kajian dan Rekayasa Teknik PUSKAREKA 2019.

Tabel 5. Perhitungan naïve Bayes

Label	R=4	F=16	M=1250	T=16	P(X C _i)	
1	164	0,335	0,024	0,116	0,085	2,211
0	436	0,183	0,016	0,078	0,069	1,149

Tabel 6. Kinerja naïve Bayes untuk 200 data

Nilai Uji	Presisi	Sensitifitas	Skor f1	Akurasi
0,2	0,64	0,65	0,62	0,65
0,3	0,59	0,60	0,53	0,60
0,4	0,55	0,57	0,49	0,58

Tabel 7. Kinerja naïve Bayes untuk 600 data

Nilai Uji	Presisi	Sensitifitas	Skor f1	Akurasi
0,2	0,81	0,82	0,79	0,82
0,3	0,77	0,79	0,76	0,79
0,4	0,75	0,78	0,75	0,78

Tabel 8. Kinerja MLP-BP NN

Jumlah Data	Presisi	Sensitifitas	Skor f1	Akurasi
200	0,53	0,41	0,46	0,47
600	0,81	0,88	0,83	0,84

Tabel 9. Hasil perbandingan kinerja antar metode

Kinerja	KNN		Naïve Bayes		MLP-BP NN	
	200	600	200	600	200	600
Akurasi	0,65	0,65	0,75	0,82	0,47	0,84
Presisi	0,64	0,64	0,73	0,81	0,53	0,81
Sensitifitas	0,65	0,65	0,75	0,82	0,41	0,88
Skor f1	0,65	0,62	0,74	0,79	0,46	0,83

Tabel 10. Perbandingan akurasi antar metode

Metode	Referensi/ Jumlah Data	Akurasi (%)
Decision Tree	[3]/500 data	78,5
naïve Bayes		81,5
Neural Network	[6]/800 data	80
kNN		71
naïve Bayes		71
NN		80
kNN	200 data	65
	600 data	75
naïve Bayes	200 data	72,5
	600 data	84,17
MLP-BP NN	200 data	47
	600 data	84,3

DAFTAR PUSTAKA

- [1] F. Boulton, "Evidence-based criteria for the care and selection of blood donors, with some comments on the relationship to blood supply, and emphasis on

- the management of donation-induced iron depletion," *Transfusion Medicine*, vol. 18, no. 1, pp. 13-27, 2008. doi: [10.1111/j.1365-3148.2007.00818.x](https://doi.org/10.1111/j.1365-3148.2007.00818.x)
- [2] A. Eder, M. Goldman, S. Rossmann, D. Waxman, and C. Bianco, "Selection criteria to protect the blood donor in North America and Europe: past (dogma), present (evidence), and future (hemovigilance)," *Transfusion Medicine Reviews*, vol. 23, no. 3, pp. 205-220, 2009. doi: [10.1016/j.tmr.2009.03.003](https://doi.org/10.1016/j.tmr.2009.03.003)
- [3] W. B. Zulfikar, Y. A. Gerhana, and A. F. Rahmania, "An approach to classify eligibility blood donors using decision tree and naive bayes classifier," in *6th International Conference Cyber IT Service Management (CITSM)*, Parapat, Indonesia, Aug. 2018, pp. 1-5. doi: [10.1109/CITSM.2018.8674353](https://doi.org/10.1109/CITSM.2018.8674353)
- [4] *Blood donor selection: guidelines on assessing donor suitability for blood donor donation*. World Health Organization, 2012.
- [5] I.-C. Yeh, K.-J. Yang, and T.-M. Ting, "Knowledge discovery on RFM model using Bernoulli sequence," *Expert System with Applications*, vol. 36, no. 3, part 2, pp. 5866-5871, 2009. doi: [10.1016/j.eswa.2008.07.018](https://doi.org/10.1016/j.eswa.2008.07.018)
- [6] W. E. Susanto and D. Riana, "Komparasi algoritma neural network, k-nearest network dan naive bayes untuk memprediksi pendonor darah potensial," *Jurnal Speed*, vol. 8, no. 3, pp. 18-27, 2016.
- [7] B. M. Shashikala, M. P. Pushpalatha, and B. Vijaya, "Machine learning approaches for potential blood donors prediction," in *Emerging Research in Electronics, Computer Science and Technology*, vol. 545, 2019, pp. 483-491. doi: [10.1007/978-981-13-5802-9_44](https://doi.org/10.1007/978-981-13-5802-9_44)
- [8] M. A. jabbar, B. L. Deekshatulu, and P. Chandra, "Classification of heart disease using k-nearest neighbor and genetic algorithm," *Procedia Technology*, vol. 10, pp. 85-94, 2013. doi: [10.1016/j.protcy.2013.12.340](https://doi.org/10.1016/j.protcy.2013.12.340)
- [9] H. Yigit, "A weighting approach for KNN classifier," in *International Conference on Electronics, Computer and Computation*, Ankara, Turkey, Nov. 2013, pp. 228-231. doi: [10.1109/ICECCO.2013.6718270](https://doi.org/10.1109/ICECCO.2013.6718270)
- [10] X. Wu et al., "Top 10 algorithms in data mining," *Knowledge and Information System*, vol. 14, no. 1, pp. 1-37, 2008. doi: [10.1007/s10115-007-0114-2](https://doi.org/10.1007/s10115-007-0114-2)
- [11] Y. Kumar and G. Sahoo, "Prediction of different types of liver diseases using rule based classification model," *Technology and Health Care*, vol. 21, no. 5, pp. 417-432, 2013. doi: [10.3233/THC-130742](https://doi.org/10.3233/THC-130742)
- [12] M. Mishra and M. Srivastava, "A view of artificial neural network," in *International Conference on Advances in Engineering & Technology Research*, Unnao, India, Aug. 2014, pp. 1-3. doi: [10.1109/ICAETR.2014.7012785](https://doi.org/10.1109/ICAETR.2014.7012785)
- [13] W. B. Zulfikar, M. Irfan, C. N. Alam, and M. Indra, "The comparison of text mining with Naive Bayes classifier, nearest neighbor, and decision tree to detect Indonesian swear words on Twitter," in *5th International Conference Cyber IT Service Management (CITSM)*, Denpasar, Indonesia, Aug. 2017, pp. 1-5. doi: [10.1109/CITSM.2017.8089231](https://doi.org/10.1109/CITSM.2017.8089231)
- [14] W. B. Zulfikar and N. Lukman, "Perbandingan Naive Bayes classifier dengan Nearest Neighbor untuk identifikasi penyakit mata," *Journal Online Informatika*, vol. 1, no. 2, pp. 82-86, 2016, doi: [10.15575/join.v1i2.33](https://doi.org/10.15575/join.v1i2.33)
- [15] J. Yang, Z. Ye, X. Zhang, W. Liu, and H. Jin, "Attribute weighted Naive Bayes for remote sensing image classification based on cuckoo search algorithm," in *2017 International Conference Security Pattern Analylis, and Cybernetics (SPAC)*, Shenzhen, China, Dec. 2017, pp. 169-174. doi: [10.1109/SPAC.2017.8304270](https://doi.org/10.1109/SPAC.2017.8304270)
- [16] M. Darwiche, M. Feuilloy, G. Bousaleh, and D. Schang, "Prediction of blood transfusion donation," in *4th International Conference on Research Challenges in Information Science (RCIS)*, Nice, France, May 2010, pp. 51-56. doi: [10.1109/RCIS.2010.5507363](https://doi.org/10.1109/RCIS.2010.5507363)
- [17] B. Gabrys and L. Petrakieva, "Combining labelled and unlabelled data in the design of pattern classification systems," *Int. J. Approx. Reason.*, vol. 35, no. 3, pp. 251-273, 2004. doi: [10.1016/j.ijar.2003.08.005](https://doi.org/10.1016/j.ijar.2003.08.005)
- [18] S. Agarwal, "Data mining: Data mining concepts and techniques," in *International Conference on Machine Intelligence and Research Advancement*, Katra, India, Dec. 2013, pp. 203-207. doi: [10.1109/ICMIRA.2013.45](https://doi.org/10.1109/ICMIRA.2013.45)