



ANALISIS PERBANDINGAN ALGORITMA *SUPPORT VECTOR MACHINE*, NAIVE BAYES, DAN REGRESI LOGISTIK UNTUK MEMPREDIKSI DONOR DARAH

Hendriyana¹, Ichwanul Muslim Karo Karo², Sri Dewi³

¹ Rekayasa Perangkat Lunak, UPI Kampus Daerah Cibiru, Universitas Pendidikan Indonesia

^{2,3} Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Medan

Bandung, Jawa Barat, Indonesia 40614

Medan, Sumatera Utara, Indonesia 20371

hendriyana@upi.edu, ichwanul@unimed.ac.id, sridewi@unimed.ac.id

Abstract

Blood supplies and stocks are urgently needed. Regular donations from healthy volunteers are the only way to keep up with the blood supply. This research aims to develop and evaluate a machine-learning algorithm to predict whether a volunteer will donate or not. The machine learning algorithms are Naïve Bayes, Logistic Regression, and Support Vector Machine (SVM). This study also applies the process of normalizing data with a Z-score to standardize the dataset scale. The dataset is sourced from the Hsin-Chu City Blood Transfusion Service, Taiwan, and stored in the UCI repository. The evaluation methods are accuracy, precision, recall, and F-1 score. The research results with the Naïve Bayes algorithm were 89.90%, Logistic Regression 82.59%, and SVM 94.79%. The normalization process using the Z-Score method contributes positively to improving the performance of the classification model. Based on this performance, it provides predictive results for volunteers who will return to donate blood to offer blood reserves to those in need.

Keywords: Blood transfusion, Naive Bayes, Logistic Regression, SVM, Z-Score

Abstrak

Suplai dan stok darah sangat dibutuhkan di saat genting. Satu-satunya cara untuk memenuhi ketersediaan darah adalah sumbangan rutin dari sukarelawan yang sehat. Tujuan dari penelitian ini adalah untuk mengembangkan dan mengevaluasi algoritma *machine learning* untuk memprediksi seorang sukarelawan akan berdonor atau tidak. Algoritma *machine learning* yang digunakan adalah Naïve Bayes, Regresi Logistik dan *Support Vector Machine* (SVM). Penelitian ini juga menerapkan proses normalisasi data dengan *Z-score* untuk menstandarkan skala dataset. Adapun *dataset* bersumber dari Layanan Transfusi Darah Kota Hsin-Chu, Taiwan yang tersimpan di *UCI repository*. Metode evaluasi yang digunakan ialah akurasi, *precision*, *recall* dan *F-1 score*. Hasil penelitian dengan algoritma Naïve Bayes 89.90%, Regresi Logistik 82.59% dan SVM 94.79%. Proses normalisasi menggunakan metode *Z-Score* berkontribusi positif dalam meningkatkan performansi model klasifikasi. Berdasarkan performansi tersebut memberikan hasil prediksi bagi sukarelawan yang akan kembali mendonor darah untuk memberikan cadangan darah bagi yang membutuhkan.

Kata kunci: Donor Darah, Naive Bayes, Regresi *Logistic*, *SVM*, *Z-Score*

1. PENDAHULUAN

Keuntungan donor darah tidak hanya didapatkan bagi penerima darah namun juga bagi pendonor darah memiliki dampak positif yang banyak terutama bagi kesehatan. Terdapat 80% populasi dunia yang mendapatkan 20% darah yang sesuai dengan syarat kesehatan (pendonor dicek kelayakan kesehatan dan darah disimpan pada tempat benar)[1]. Menurut Palang Merah Amerika, hasil donor tidak dapat transfusi ke penerima (*recipient*) setelah 42

hari[1]. Hal ini memberikan dampak permasalahan yang sangat penting untuk menjaga kecukupan suplai darah, terutama saat terjadi bencana. Satu-satunya cara untuk memenuhi ketersediaan darah adalah dengan memberikan sumbangan rutin dari sukarelawan yang sehat.

Darah yang sehat Darah yang sehat bersumber dari pendonor yang sehat dan disimpan dengan benar. Menurut penelitian[2] Kelayakan seorang pendonor dapat dilihat dari *Recency* (berapa bulan sejak terakhir kali mendonor),

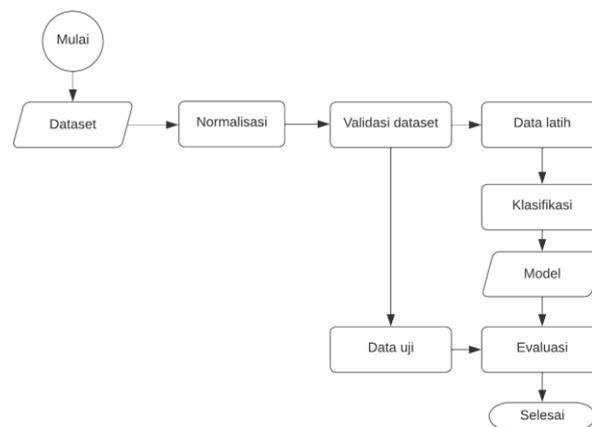
Frequency (berapa kali mendonasikan darah), *Monetary* (jumlah darah yang didonasikan), *Time* (berapa bulan sejak pertama kali mendonor). Selama ini proses identifikasi kelayakan pendonor dilakukan oleh para professional medis. Namun di masa yang akan datang, tidak menutup kemungkinan, identifikasi kondisi pendonor dilakukan oleh sebuah *machine*, sehingga dengan cepat mengidentifikasi dan mampu meningkatkan jumlah darah dari pendonor.

Beberapa penelitian yang mengidentifikasi pendonor darah dengan menggunakan pendekatan *machine learning*. Penelitian [3] mengklasifikasikan dan memprediksi jumlah pendonor darah menurut umur dan golongan darahnya. aplikasi *Weka* telah digunakan dan berhasil untuk menjalankan algoritma J48. Algoritma tersebut menghasilkan model klasifikasi dengan akurasi 80.88%. Penelitian [4] memprediksi kesehatan pendonor darah di *Blood Transfusion Organization* (BTO) dengan algoritma *Decision Tree C4.5*, *Naïve Bayes*, dan *Support Vector Machine* (SVM). Adapun dataset terdiri dari 11006 donor dan tujuh atribut (jenis kelamin, usia, pekerjaan, pendidikan, status perkawinan, jenis donor, hasil tes darah). Algoritma SVM memiliki akurasi tertinggi dibandingkan dengan dua algoritma lainnya. Hasil dari ketiga algoritma telah dibandingkan dan analisis biaya kesalahan telah dilakukan. Penelitian lainnya [5] mencoba mengidentifikasi apakah seorang pendonor akan kembali berdonor atau tidak dengan menggunakan algoritma *machine learning*. Adapun algoritma yang digunakan adalah *Random Forest*, SVM dan Regresi Logistik. Dataset diperoleh dari Bank darah Saudi dari 2017 hingga 2018, terdiri dari 9 variabel (*Id*, golongan darah, jenis kelamin, usia, jumlah berdonor, tanggal pertama kali berdonor darah, tanggal terakhir kali berdonor darah, kebangsaan, dan periode) serta dua kelas.

Tujuan penelitian ini adalah memprediksi apakah seorang sukarelawan akan mendonorkan darahnya atau tidak melalui pendekatan algoritma *machine learning* antara lain *Support Vector Machine*, *Naïve Bayes* dan Regresi Logistik. Adapun dataset yang diujikan adalah *database* donor dari Layanan Transfusi Darah Kota Hsin-Chu, Taiwan. *Dataset* dari *UCI repository* dipilih karena telah memiliki validasi data yang cocok untuk masalah klasifikasi.

2. METODE PENELITIAN

Penelitian ini terdiri dari beberapa proses (Gambar 1). Proses awal merupakan mengumpulkan *dataset*, dilanjutkan dengan proses normalisasi data, validasi data, dan proses klasifikasi. Adapun algoritma klasifikasi yang digunakan adalah *Support Vector Machine*, *Naïve Bayes* dan Regresi Logistik. Luaran proses klasifikasi merupakan sebuah model klasifikasi, dan Proses evaluasi menggunakan akurasi berdasarkan *confusion matrix*.



Gambar 1. Kerangka Kerja Penelitian

2.1 Dataset

Dataset donor darah berasal dari database donor Layanan Transfusi Darah Kota Hsin-Chu, Taiwan. *Dataset* tersebut tersimpan di *UCI repository*. *Dataset* terdiri dari 7480 *record*, Tabel 1 menjelaskan 4 variabel dari dataset; *Frequency* (berapa kali mendonasikan darah), *Monetary* (jumlah darah yang didonasikan), *Time* (berapa bulan sejak pertama kali mendonor) dan kelas (berdonasi atau tidak). *Dataset* ini juga telah digunakan pada penelitian sebelumnya[2].

Tabel 1. Deskripsi Variabel *Dataset*

No	Variabel	Deskripsi
1	<i>Recency</i>	berapa bulan sejak terakhir kali mendonor
2	<i>Frequency</i>	berapa kali mendonasikan darah.
3	<i>Monetary</i>	jumlah darah yang didonasikan
4	<i>Time</i>	berapa bulan sejak pertama kali mendonor
5	Kelas	0 berarti dapat berdonasi, 1 tidak dapat berdonasi.

Lebih lanjut, Tabel 2 menyajikan contoh *dataset* yang diperoleh dari *UCI repository*. Keseluruhan variabel *dataset* bertipe numerik

Tabel 2. Contoh *Dataset*

<i>Recency</i>	<i>Frequency</i>	<i>Monetary</i>	<i>Time</i>	Kelas
2	50	12500	98	0
0	13	3250	28	0
1	16	4000	35	0
2	20	5000	45	0
1	24	6000	77	1

2.2 Algoritma *Support Vector Machine*

Algoritma *machine learning* lainnya yang umum digunakan dan handal adalah *Support Vector Machine* (SVM)[12]. *Support Vector Machine* (SVM) merupakan salah satu algoritma pembelajaran mesin yang efektif untuk

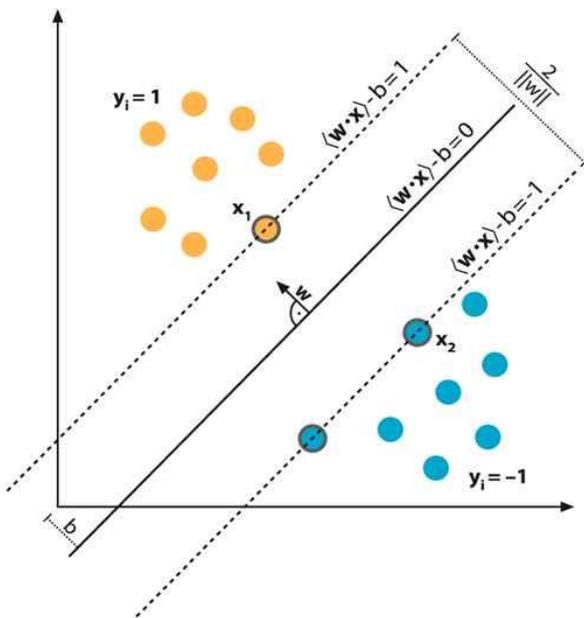
mengakomodasi banyak variabel dan banyak kelas[11]. Ide Mesin vektor pendukung juga kadang-kadang disebut "pengklasifikasi margin besar" karena ketika mendefinisikan batas keputusan antara dua kelas, ia mencoba untuk memaksimalkan margin antara setiap kelas dan batas.

Ide dari *Support Vector Machine* adalah membuat *hyperplane* (Gambar 2). Tujuannya untuk memisahkan satu set pelatihan sampel l dengan vektor data x_i dan label kelas yang sesuai y_i ($x_{i_1}, y_{i_1}, \dots, (x_{i_n}, y_{i_n}) \in \mathbb{R}^n \times \{-1, 1\}$) dengan mencari vektor bobot $w \in \mathbb{R}^n$ dan offset $b \in \mathbb{R}$ dari sedemikian hingga memenuhi persamaan (1).

$$H: \mathbb{R}^n \rightarrow \{-1, 1\} \tag{1}$$

$$x \rightarrow \text{sign}(w \cdot x + b)$$

Persamaan (1) digambarkan melalui Gambar 2.



Gambar 2. Ilustrasi Hyperplane SVM

2.3 Algoritma Naïve Bayes

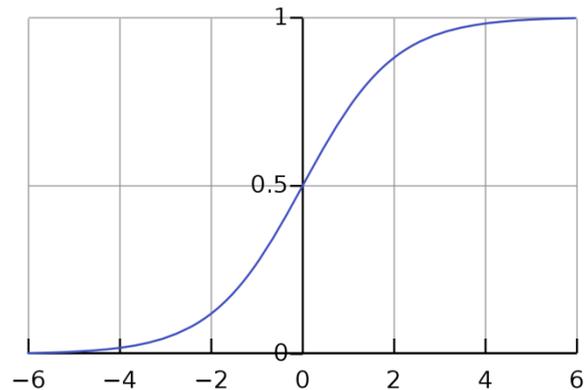
Naive Bayes adalah algoritma klasifikasi sederhana dengan pendekatan probabilitas[6]. Algoritma ini merupakan model klasifikasi berdasarkan teorema Bayes [7]. Algoritma Naïve Bayes mensyaratkan bahwa semua fitur tidak berhubungan satu sama lain dan tidak saling mempengaruhi[6]. Teorema Bayes (persamaan (2)) menghitung probabilitas posterior $P(C|X)$ dari probabilitas sebelumnya $P(C)$, $P(X|C)$ yang merupakan probabilitas sebelumnya dari X di mana C adalah hipotesis dan X adalah kumpulan data yang diberikan.

$$P(C|X) = P(X|C) \cdot \frac{P(C)}{P(X)} \tag{2}$$

Algoritma Naïve Bayes sangat efisien karena sederhana dalam proses komputasi[8]. Keunggulan lainnya adalah jumlah data latih yang dibutuhkan sedikit[9].

2.4 Algoritma Regresi Logistik

Regresi logistik adalah salah satu algoritma pembelajaran mesin yang handal digunakan untuk klasifikasi data dengan target bertipe kategori [10]. Regresi Logistik adalah jenis algoritma pembelajaran mesin klasifikasi statistik. Dasar dari algoritma ini diturunkan dari fungsi sigmoid (Gambar 3).



Gambar 3. Fungsi Sigmoid

Regresi Logistik sangat berguna dalam masalah klasifikasi biner. Karena klasifikasi saat ini yang bersangkutan adalah tipe biner dan sangat dapat diterapkan. Kelemahan dari regresi logistik adalah rentan terhadap *underfitting* dataset yang kelasnya tidak seimbang[11], sehingga menghasilkan akurasi yang rendah.

2.5 Normalisasi

Normalisasi data merupakan salah satu teknik pra-proses data. *Dataset* pada penelitian ini terjadi rentang nilai antara variabel tersebut sangat jauh, misalnya variabel *monetary* berskala ribuan dan *recency* berskala satuan. Berdasarkan kondisi tersebut, dibutuhkan proses normalisasi data. Tujuan dari normalisasi adalah untuk mengubah fitur menjadi pada skala yang sama. Selain itu, normalisasi data meningkatkan kinerja dan stabilitas pelatihan model. Dalam penelitian ini, metode normalisasi yang digunakan adalah *Z-Score*.

Z-Score merupakan ukuran penyimpangan data dari nilai rata-ratanya (μ) yang diukur dalam satuan standar deviasinya. *Z-Score* ini juga disebut dengan nilai standar atau nilai baku. Persamaan (3) merupakan formula untuk menghitung nilai *Z-score* dari setiap data yang diamati (x_i).

$$Z = \frac{x_i - \mu}{\text{standar deviasi}} \tag{3}$$

2.6 Evaluasi

Setiap algoritma pembelajaran mesin dievaluasi menggunakan akurasi (persamaan (4)), *precision* (persamaan (5)), *recall* (persamaan (6)) dan *F-1 score* (persamaan (7)). Pengukuran tersebut berdasarkan *confusion matrix* (Tabel 3).

Tabel 3. Confusion Matrix

		Aktual			
		Kelas 0		Kelas 1	
Prediksi	Kelas 0	True Positive (TP)	False Negative (FN)		
	Kelas 1	False Positive (FP)	True Negative (TN)		

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100\% \tag{4}$$

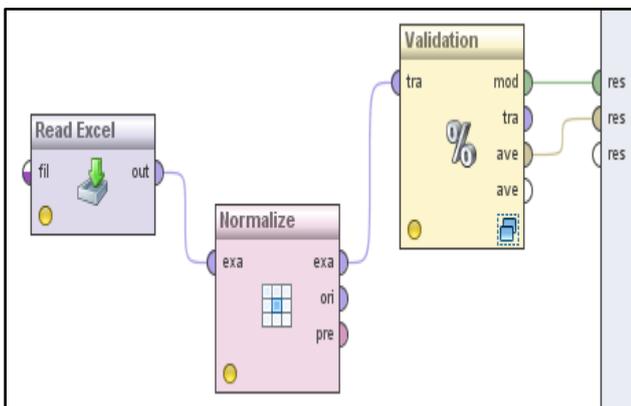
$$P = \frac{TP}{TP + FP} \cdot 100\% \tag{5}$$

$$Recall(R) = \frac{TP}{TP + FN} \cdot 100\% \tag{6}$$

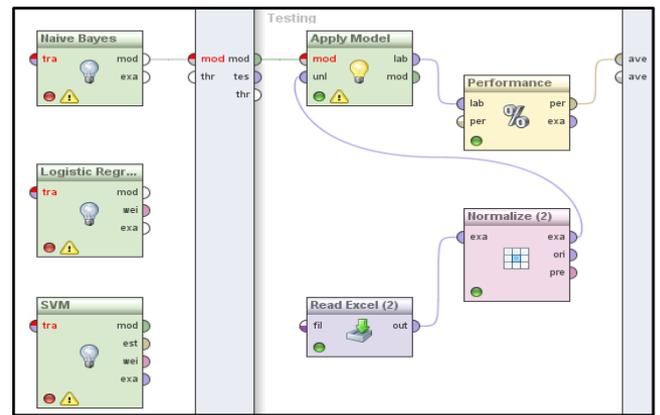
$$F1\ score = \frac{2PR}{P + R} \tag{7}$$

3. HASIL DAN PEMBAHASAN

Proses klasifikasi dilakukan dengan bantuan perangkat lunak *RapidMiner*. *RapidMiner* merupakan perangkat lunak yang bersifat terbuka (*open source*) untuk melakukan analisis terhadap data, text dan prediksi[13]. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. Gambar 4 merupakan proses persiapan data pada *RapidMiner*, sedangkan Gambar 5 merupakan gambaran implementasi algoritma *machine learning* di *RapidMiner*.



Gambar 4. Proproses Data dengan *RapidMiner*



Gambar 5. Penerapan Model Klasifikasi di RapidMiner

3.1 Hasil Algoritma Naïve Bayes

Percobaan pertama merupakan proses identifikasi sukarelawan yang akan mendonor atau tidak dengan menggunakan algoritma Naïve Bayes. Pendekatan algoritma Naïve Bayes berdasarkan peluang, sehingga tidak ada manual *input* parameter. Adapun performansi model dari algoritma Naïve Bayes tersaji pada tabel 4, dengan akurasi 82.09%, *precision* 88.34%, *recall* 91.51% dan *F-1 score* 89.90%.

Tabel 4. Confusion Matrix Algoritma Naïve Bayes Classifier

	Aktual 0	Aktual 1
pred. 0	1185	110
pred. 1	155	30

3.2 Hasil Algoritma Regresi Logistik

Percobaan selanjutnya merupakan proses identifikasi sukarelawan yang akan mendonor atau tidak dengan menggunakan algoritma Regresi logistik. Pendekatan algoritma Regresi Logistic adalah model regresi linear yang digunakan untuk prediksi biner. Adapun performansi model dari algoritma Regresi logistik tersaji pada Tabel 5, dengan akurasi 71.35%, *precision* 75.07%, *recall* 91.79% dan *F-1 score* 82.59%.

Tabel 5. Confusion Matrix Algoritma Regresi Logistik

	Aktual 0	Aktual 1
pred. 0	1006	90
pred. 1	334	50

3.3 Hasil Algoritma SVM

Percobaan selanjutnya merupakan proses identifikasi sukarelawan yang akan mendonor atau tidak dengan menggunakan algoritma Regresi Logistik. Pendekatan algoritma Regresi Logistic adalah model regresi linear yang digunakan untuk prediksi biner. Adapun performansi model dari algoritma Regresi Logistik tersaji pada Tabel 6 dengan

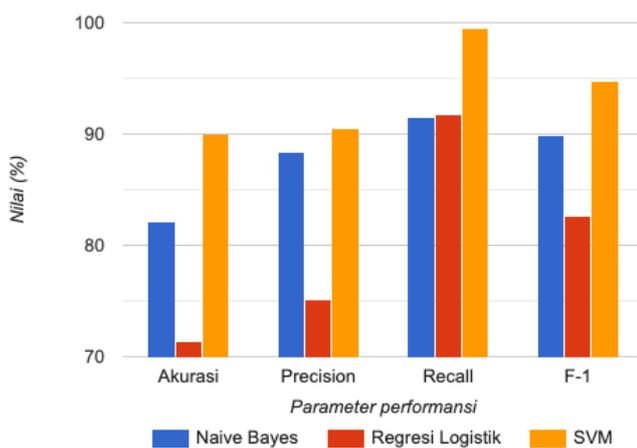
akurasi 90%, *precision* 90.54%, *recall* 99.46% dan *F-1 score* 94.79%.

Tabel 6. *Confusion Matrix* Algoritma SVM

	Aktual 0	Aktual 1
pred. 0	1332	140
pred. 1	8	0

3.4 Analisis Perbandingan Antar Algoritma dan Penelitian Sebelumnya

Bagian ini menyajikan analisis perbandingan performansi algoritma. Gambar 6 menampilkan perbandingan performansi ketiga model klasifikasi untuk mengidentifikasi sukarelawan. Dari empat parameter performansi, nilai algoritma SVM selalu lebih unggul dibandingkan dengan dua algoritma lainnya. Dengan kata lain, algoritma SVM memberikan performansi terbaik dibandingkan dengan algoritma Naïve Bayes dan Regresi Logistik. Adapun algoritma Regresi *Logistic* menjadi pilihan terburuk untuk mengidentifikasi sukarelawan pendonor.



Gambar 6. Perbandingan Algoritma

Algoritma Naïve Bayes membutuhkan data *training* data yang kecil untuk mengestimasi parameter dengan perhitungan yang mudah dilakukan tetapi memiliki kelemahan pada probabilitas jika kondisinya nol maka probabilitas prediksi akan memberikan nilai nol juga. Regresi Logistik memiliki kelemahan yaitu *underfitting* pada *class* yang tidak seimbang sehingga menghasilkan akurasi yang rendah. Algoritma SVM memiliki performansi paling tinggi karena metode ini lebih mudah digunakan dengan mencari nilai minimum batas antara dua kelas *hyperplane*.

Terdapat penelitian sebelumnya yang telah menggunakan dataset yang sama [2]. Sehingga pada bagian ini menyajikan perbandingan dengan hasil penelitian sebelumnya (Tabel 7). Hasil penelitian [2] menunjukkan bahwa algoritma *Random Tree* merupakan model klasifikasi terbaik dibandingkan dengan dua algoritma lain yang digunakan pada penelitian

tersebut. Sedangkan performansi algoritma SVM + PCA pada penelitian[14] tidak lebih baik dibandingkan dengan algoritma[2]. Hasil penelitian pada jurnal ini menunjukkan bahwa, model klasifikasi yang dihasilkan algoritma SVM lebih unggul dibandingkan dengan penelitian sebelumnya [2].

Tabel 7. Perbandingan dengan Penelitian Sebelumnya

Penelitian	Algoritma	<i>F1-Score</i> (%)
[2]	Naïve Bayes	75
	J48	80.88
	Random Tree	93.18
[14] Saat ini	SVM + PCA	78.2
	Naïve Bayes	89.90
	Regresi Logistik	82.59
	SVM	94.79

Terdapat kesamaan algoritma yang digunakan dengan penelitian [2], yakni algoritma Naïve Bayes, namun performansi model berbeda. Perbedaan mendasar yang menyebabkan perbedaan hasil adalah pelibatan proses normalisasi data pada penelitian yang tidak digunakan pada penelitian sebelumnya. Sehingga algoritma Naïve Bayes dan proses normalisasi pada penelitian ini lebih baik dibandingkan dengan performansi Naïve Bayes pada penelitian sebelumnya. Dengan kata lain, proses normalisasi memberikan dampak positif terhadap performansi model klasifikasi Naïve Bayes.

Terdapat pula kesamaan algoritma yang digunakan pada penelitian [14], yakni algoritma SVM. Perbedaan mendasar ialah penelitian sebelumnya menggunakan fitur seleksi fitur (PCA) yang berfungsi untuk mengurangi fitur yang tersedia sedangkan pada penelitian ini tidak. Perbedaan lainnya ialah penelitian ini melibatkan proses normalisasi sedangkan penelitian sebelumnya tidak. Alasan tidak menggunakan fitur seleksi pada penelitian ini ialah karena jumlah variabel hanya 4. Sedangkan alasan pelibatan normalisasi data ialah *range* data antar variabel yang tidak seimbang, sehingga perlu di normalisasi.

4. KESIMPULAN

Berdasarkan hasil percobaan yang telah dilakukan, dapat disimpulkan bahwa performansi algoritma Naïve Bayes, algoritma Regresi Logistik dan algoritma SVM yang digunakan dalam penelitian ini cukup baik. Setiap algoritma klasifikasi memiliki karakteristik yang berbeda-beda hal ini memungkinkan hasil setiap performansi algoritma berbeda-beda pula. Setelah melakukan pengujian terhadap tiga algoritma klasifikasi dan membandingkan dengan penelitian sebelumnya hasil penelitian dengan Naïve Bayes 89.90%, Regresi Logistik 82.59% dan SVM 94.79%. SVM merupakan algoritma yang memiliki performansi paling tinggi. Sehingga pada kasus prediksi donor darah untuk mempertahankan cadangan darah yang sesuai dengan syarat kesehatan dengan bentuk model data *numeric* dengan pendekatan *machine learning* ialah *Support Vector Machine* karena performa yang dimiliki sesuai untuk

mengklasifikasikan dua kelas yang berbeda yaitu pendonor akan melakukan donor kembali atau tidak melakukan donor. Penambahan Metode *Z-Score* berkontribusi positif dalam meningkatkan performansi model klasifikasi.

DAFTAR PUSTAKA

- [1] Y. Widyastuti, D. Sari, and S. R. Febrianti, "Manajemen Transfusi Masif Pada Pediatrik," *Jurnal komplikasi Anestesi*, vol. 3, no. Dic, 2016.
- [2] E. Alajrami *et al.*, "A comparative study of classification algorithm on blood transfusion," *International Journal of Academic Engineering Research*, vol. 3, no. 6, 2014.
- [3] A. Sharma and P. C. Gupta, "Predicting the Number of Blood Donors through their Age and Blood Group by using Data Mining Tool," *International Journal of Communication and Computer Technologies*, vol. 01, no. 6, 2012.
- [4] M. Khalilinezhad, B. Minaei, G. Vernazza, and S. Dellepiane, "Prediction of healthy blood with data mining classification by using Decision Tree, Naive Bayesian and SVM approaches," in *Sixth International Conference on Graphic and Image Processing (ICGIP 2014)*, 2015, vol. 9443. doi: 10.1117/12.2179871.
- [5] A. S. Alkahtani and M. Jilani, "Predicting return donor and analyzing blood donation time series using data mining techniques," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 8, 2019, doi: 10.14569/ijacsa.2019.0100816.
- [6] I. M. Karo Karo, S. Nadia Amalia, dan Dian Septiana, P. Ilmu Komputer, and P. Matematika, "Klasifikasi Kebakaran Hutan Menggunakan Feature Selection dengan Algoritma K-NN, Naive Bayes dan ID3," *Journal of Software Engineering, Information and Communication Technology*, vol. 3, no. 1, pp. 121–126, 2022.
- [7] Y. I. Kurniawan, "Perbandingan Algoritma Naive Bayes dan C.45 dalam Klasifikasi Data Mining," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 4, 2018, doi: 10.25126/jtiik.201854803.
- [8] H. Pramudia and A. Nugroho, "Sistem Informasi Kerusakan Laptop Menggunakan Metode Naive Bayes," *Teknologi Elektro, Universitas Mercu Buana*, vol. 8, no. 3, 2017.
- [9] R. Hasudungan and W. J. Pranoto, "Implementasi Teorema Naive Bayes Pada Prediksi Prestasi Mahasiswa," *Jurnal Rekayasa Teknologi Informasi (JURTI)*, vol. 5, no. 1, 2021, doi: 10.30872/jurti.v5i1.4996.
- [10] I. M. Karo Karo, M. Farhan, M. Fudzee, S. Kasim, and A. A. Ramli, "Karonese Sentiment Analysis: A New Dataset and Preliminary Result," *JOIV: International Journal on Informatics Visualization*, vol. 6, no. 2–2, pp. 523–530, 2022, [Online]. Available: www.joiv.org/index.php/joiv
- [11] I. M. Karo Karo, M. F. M. Fudzee, S. Kasim, and A. A. Ramli, "Sentiment Analysis in Karonese Tweet using Machine Learning," *Indonesian Journal of Electrical Engineering and Informatics*, vol. 10, no. 1, pp. 219–231, Mar. 2022, doi: 10.52549/ijeei.v10i1.3565.
- [12] A. Prayoga, H. A. Tawakal, and R. Aldiansyah, "Pengembangan Metode Deteksi Tingkat Kematangan Buah Melon Berdasarkan Tekstur Kulit Buah dengan menggunakan Metode Ekstraksi Ciri Statistik dan Support Vector Machine (SVM)," *Jurnal Teknologi Terpadu*, vol. 4, no. 1, 2018, doi: 10.54914/jtt.v4i1.112.
- [13] D. Novianti, "Implementasi Algoritma Naive Bayes Pada Data Set Hepatitis Menggunakan Rapid Miner," *Paradigma - Jurnal Komputer dan Informatika*, vol. 21, no. 1, 2019, doi: 10.31294/p.v21i1.4979.
- [14] M. Darwiche, M. Feuilloy, G. Bousaleh and D. Schang, "Prediction of blood transfusion donation," 2010 Fourth International Conference on Research Challenges in Information Science (RCIS), 2010, pp. 51-56, doi: 10.1109/RCIS.2010.5507363.