

## Percentage Split dan K-Fold Cross Validation pada Algoritma Support Vector Machine (SVM) pada Klasifikasi Penyakit Anemia

Herlina Hanum<sup>1</sup>, Syafrina Lamin<sup>2</sup>, Puspitahati<sup>3</sup>, Sugandi Yahdin<sup>4</sup>, Irmeilyana<sup>5</sup>, Anita Desiani<sup>6\*</sup>,  
Dite Geovani<sup>7</sup>, Redina An Fadhila Chaniago<sup>8</sup>

<sup>1,4,5,6,7,8</sup>Matematika, FMIPA Universitas Sriwijaya

<sup>2</sup>Biologi, FMIPA Universitas Sriwijaya

<sup>3</sup>Teknik Pertanian, Universitas Sriwijaya

e-mail: anitadesiani@unsri.ac.id

### Abstrak

Anemia merupakan salah satu penyakit yang memiliki pengaruh terhadap peningkatan angka kematian, penurunan tingkat produktivitas kerja, dan gangguan terhadap perkembangan saraf. Salah satu hal yang dapat dilakukan untuk mengurangi tingkat penderita penyakit anemia adalah deteksi dini dengan memanfaatkan teknologi informasi. Klasifikasi merupakan salah satu teknologi data mining yang dapat digunakan untuk memprediksi suatu pola yang terdapat dalam sejumlah data. Pada penelitian ini dilakukan klasifikasi penyakit anemia menggunakan metode support vector machine dengan kernel polynomial untuk dan metode latih yang digunakan ialah percentage split dan k-fold cross validation. Hasil yang diperoleh pada penelitian ini ialah pada metode latih percentage split memiliki tingkat akurasi sebesar 92,3% dan pada metode latih k-fold cross validation memberikan hasil akurasi sebesar 90%. Hal ini menunjukkan bahwa metode support vector machine cukup baik untuk digunakan dalam melakukan klasifikasi penyakit anemia.

**Kata Kunci:** Klasifikasi, Anemia, Support Vector Machine

### Abstract

Anemia is a disease that influences mortality rate escalation, decreases work productivity levels, and disrupts neurodevelopment. One of the things that can be done to reduce the level of anemia sufferers is early detection by utilizing information technology. Classification is one of the data mining technologies that can be used to predict a pattern contained in several data. In this study, the anemia classification was conducted using the support vector machine method with polynomial kernels, the training methods used were percentage split and k-fold cross-validation. The results obtained in this study are that the percentage split training method has an accuracy rate of 92.3% and the k-fold cross-validation training method gives an accuracy of 90%. This shows that the support vector machine method is good enough to be used in classifying anemia. ]

**Keywords:** Classification, Anemia, Support Vector Machine

## 1. PENDAHULUAN

Penyakit anemia merupakan penyakit yang terjadi karena gangguan pada hemoglobin yang terdapat di dalam sel darah merah sehingga tidak dapat menyalurkan oksigen ke seluruh bagian tubuh manusia. Beberapa faktor yang menyebabkan terjadinya anemia antara lain yaitu asupan energi, zat besi, protein, vitamin c, investasi cacing, kebiasaan minum kopi atau teh, pendapatan dalam keluarga, edukasi dari orang tua mengenai anemia, dan pola menstruasi. Penderita penyakit anemia memiliki daya tahan fisik yang rendah, menurunnya aktivitas fisik, dan sulit berkonsentrasi [1]. Apabila penyakit anemia menyerang dalam waktu yang lama dan tidak adanya tanggapan dari penderita maka dapat menyebabkan kerusakan pada otak, jantung, dan organ lainnya

dalam tubuh [2]. Salah satu hal yang dapat dilakukan untuk mengurangi tingkat penderita penyakit anemia adalah deteksi dini dengan teknologi informasi agar dapat memudahkan proses deteksi dini anemia pada masyarakat [3].

Dalam bidang kesehatan, *data mining* merupakan salah satu bidang teknologi informasi yang dapat menjadi solusi untuk digunakan dalam deteksi dini sebuah penyakit [4]. *Data mining* merupakan teknik yang dilakukan terhadap kumpulan data besar untuk mengidentifikasi pola tersembunyi dalam data tersebut dengan menggunakan strategi kombinasional data analisis statistik pembelajaran mesin dan teknologi basis data. Penerapan *data mining* dalam bidang kesehatan adalah hal yang sangat penting karena dibutuhkannya pengembangan berbagai aplikasi dalam perawatan kesehatan yang terus berkembang [5]. Penerapan *data mining* dalam bidang kesehatan salah satunya adalah klasifikasi. Klasifikasi merupakan proses pembangunan model atau aturan klasifikasi tertentu dengan menggunakan *data training* sebagai tahapan pembelajaran dan *data testing* sebagai tahap pengujian. Metode klasifikasi yang sering digunakan salah satunya adalah *Support Vector Machine* (SVM) [6]. Metode SVM merupakan salah satu metode yang paling tepat yang digunakan dalam proses klasifikasi karena metode ini mampu bekerja untuk menentukan titik maksimal dan garis batas pemisah terbaik untuk memisahkan dua buah kelas [7]. Selain itu, kelebihan dari metode SVM ialah dapat memberikan hasil klasifikasi dengan tingkat akurasi yang tinggi [8]. Beberapa penelitian sebelumnya yang telah melakukan klasifikasi dengan metode SVM dalam bidang kesehatan adalah klasifikasi penyakit gigi dan mulut [6] yang memberikan akurasi sebesar 92%, klasifikasi kanker payudara [9] yang memberikan hasil akurasi sebesar 93,14%, klasifikasi penyakit ginjal kronis [10] yang memberikan hasil akurasi sebesar 90,56%, dan penyakit demam berdarah [11] yang memberikan hasil akurasi sebesar 93%.

Pada beberapa kasus dengan jumlah data yang besar cenderung tidak memiliki sebaran yang linier sehingga data tidak bisa diklasifikasikan dengan baik menggunakan metode SVM dengan kernel linier, oleh karena itu dikembangkan suatu kernel untuk melakukan klasifikasi terhadap data yang tidak tersebar secara linier yaitu kernel *polynomial* [12]. Beberapa penelitian sebelumnya yang menggunakan kernel *polynomial* antara lain yaitu klasifikasi *tweet public figure* [13] dengan tingkat akurasi sebesar 70%, klasifikasi vaksinasi Covid-19 [14] dengan tingkat akurasi sebesar 86%, dan klasifikasi untuk pengenalan wajah [15] dengan tingkat akurasi sebesar 80%. Berdasarkan permasalahan dan pengetahuan yang telah dijelaskan sebelumnya, maka pada penelitian ini akan dilakukan pembuatan sebuah sistem klasifikasi penyakit anemia terhadap data yang memiliki dua kelas menggunakan metode SVM biner dengan kernel *polynomial* karena data yang akan digunakan memiliki jumlah yang cukup besar, lalu akan diukur juga untuk hasil tingkat akurasinya serta tingkat presisi, *re-call*, dan *f1-score*.

## 2. METODE PENELITIAN

### 2.1 Dataset

Dataset yang digunakan pada penelitian ini merupakan hasil kumpulan data dari tes perhitungan darah lengkap yang dilakukan oleh analisis hematologi yang bertujuan untuk mengukur jumlah pasien di Pusat Diagnostik Eureka di Lucknow, India yang mengidap penyakit anemia. Periode pengumpulan data dilakukan selama empat bulan

yaitu dari September sampai Desember 2020 dan didapat sebanyak 364 data. Data diperoleh dari situs [www.kaggle.com](http://www.kaggle.com) dan bisa diakses melalui alamat URL <https://www.kaggle.com/datasets/manuelunmsm/anemia-prediction-based-on-machine-learning?>.

Dataset yang digunakan memiliki sebelas atribut, antara lain yaitu umur, jenis kelamin, jumlah sel darah merah, volume sel, tingkat rata-rata volume sel, tingkat rata-rata sel hemoglobin, konsentrasi hemoglobin eritrosit rata-rata, lebar distribusi sel darah merah, tingkat sel darah putih, tingkat hemoglobin, dan hasil uji. Seluruh atribut digunakan dalam penelitian ini karena memiliki data yang lengkap. Atribut yang menjadi variabel independen ialah umur, jenis kelamin, jumlah sel darah merah, volume sel, tingkat rata-rata volume sel, tingkat rata-rata sel hemoglobin, konsentrasi hemoglobin eritrosit rata-rata, lebar distribusi sel darah merah, tingkat sel darah putih, dan tingkat hemoglobin, sedangkan variabel hasil uji merupakan variabel terikat atau dependen. Pada Tabel 1 berikut ditampilkan enam *entry* teratas dari dataset yang digunakan beserta keterangan untuk setiap atribut yang ditunjukkan pada Tabel 2.

Tabel 1. Contoh *Entry* Dataset

<i>Age</i>	<i>Sex</i>	<b>RBC</b>	<b>PCV</b>	<b>MCV</b>	<b>MCH</b>	<b>MCHC</b>	<b>RDW</b>	<b>TLC</b>	<b>HGB</b>	<b>TEST</b>
28	0	5,66	34	60,1	17	28,2	20	11,1	9,6	1
41	0	4,78	44,5	93,1	28,9	31	13	7,02	13,8	0
40	1	4,65	41,6	89,5	28,8	32,2	13	8,09	13,4	0
76	0	4,24	36,7	86,6	26,7	30,8	14,9	13,41	11,3	1
20	1	4,14	36,9	89,1	27,8	31,2	13,2	4,75	11,5	1
24	0	4,29	40,1	93,5	29,6	31,7	14,5	13,96	12,7	1

Tabel 2. Keterangan Atribut

<b>Atribut</b>	<b>Keterangan</b>	<b>Range</b>
<i>Age</i>	Umur	11 – 89
<i>Sex</i>	Jenis Kelamin	0 = laki-laki, 1 = perempuan
<b>RBC</b>	Jumlah Sel Darah Merah	1,36 – 6,9
<b>PCV</b>	Volume Sel	13,1 – 56,9
<b>MCV</b>	Tingkat Rata-rata Volume Sel	55,7 – 124,1
<b>MCH</b>	Tingkat Rata-rata Sel Hemoglobin	14,7 – 41,4
<b>MCHC</b>	Konsentrasi Hemoglobin Eritrosit Rata-rata	23,6 – 50,2
<b>RDW</b>	Lebar Distribusi Sel Darah Merah	10,6 – 29,2
<b>TLC</b>	Tingkat Sel Darah Putih	2 – 42,42
<b>HGB</b>	Tingkat Hemoglobin	4,2 – 19,6
<b>TEST</b>	Hasil Uji	0 = positif anemia, 1 = negatif anemia

## 2.2 Penerapan Metode *Support Vector Machine* (SVM)

Pada penelitian ini, digunakan metode SVM biner karena data memiliki dua kelas yaitu positif anemia dan negatif anemia. Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua buah kelas pada *input space*. Data yang tersedia dinotasikan sebagai  $\mathbf{x}_i \in \mathbf{R}^d$  kemudian label masing-masing dinotasikan dengan  $\mathbf{y}_i \in \{-1, +1\}$  untuk  $i = 1, 2, \dots, l$  dimana  $l$  merupakan jumlah banyaknya data. Diasumsikan terdapat kedua kelas yaitu sampel positif (+1) dan sampel negatif (-1) dapat terpisah secara sempurna oleh *hyperplane* berdimensi  $d$ , yang didefinisikan pada Persamaan (1) dibawah ini :

$$\overline{\mathbf{w}}_i \cdot \overline{\mathbf{x}}_i + \mathbf{b} = 0 \quad (1)$$

Rumus yang digunakan untuk menghitung hasil prediksi dengan *hyperplane* tunggal dan sesuai dengan  $\mathbf{b}$  dan  $\mathbf{w}$  yang akan diperoleh didefinisikan pada Persamaan (2) dan (3) sebagai berikut :

$$f(\phi(x)) = \text{sign}(\mathbf{w} \cdot \phi(x)) + \mathbf{b} \quad (2)$$

$$f(\phi(x)) = \text{sign}(\sum_{i=1}^n \alpha_i \mathbf{y}_i \phi(x_i)^T \cdot \phi(x) + \mathbf{b}) \quad (3)$$

Keterangan :  $\mathbf{w}$  = nilai weight

$\mathbf{x}$  = nilai variabel input

$\mathbf{b}$  = nilai bias

Selanjutnya, untuk mencari lebar margin  $y$  atau jarak dari  $\mathbf{x}^+$  (data yang terletak pada kelas  $y = +1$ ) ke *hyperplane* atau jarak dari  $\mathbf{x}^-$  (data yang terletak pada kelas  $y = -1$ ) ke *hyperplane* dilakukan dengan memaksimalkan  $\|\mathbf{w}\|$  dengan syarat yang didefinisikan pada Persamaan (4) dan (5) sebagai berikut:

$$\mathbf{w} = \sum_{i=1}^n \alpha_i \mathbf{y}_i \mathbf{K}(x, x_i) \quad (4)$$

$$\mathbf{b} = -\frac{1}{2}(\mathbf{w} \cdot \mathbf{x}^+ + \mathbf{w} \cdot \mathbf{x}^-) \quad (5)$$

Dengan kendala  $\mathbf{y}_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + \mathbf{b}) \geq 1, i = 1, 2, \dots, n$  merupakan *hyperplane* dengan *margin* maksimum. Persamaan yang digunakan untuk mencari nilai *alpha* didefinisikan pada Persamaan (6) berikut:

$$\alpha_n = \frac{n}{\sum K(n*n)} \quad (6)$$

## 2.3 Evaluasi Hasil

Pada tahap ini, evaluasi hasil dilakukan dengan metode *confusion matrix* yang menampilkan hasil evaluasi model menggunakan tabel matriks. Kelas pertama dianggap positif dan kelas kedua dianggap negatif. Evaluasi menggunakan *confusion matrix* menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* (D. Putra & Wibowo, 2020). *Confusion matrix* adalah tabel yang menyatakan klasifikasi jumlah data uji yang diprediksi benar dan yang salah (Normawati & Prayogi, 2021). Contoh *confusion matrix* untuk klasifikasi biner ditunjukkan pada Tabel 3.

Tabel 3. *Confusion Matrix*

<i>Class</i>		<i>Prediction</i>	
		<i>Positive</i>	<i>Negative</i>
<i>Actual</i>	<i>Positive</i>	TP ( <i>True Positive</i> )	FN ( <i>False Negative</i> )
	<i>Negative</i>	FP ( <i>False Postive</i> )	TN ( <i>True Negative</i> )

### 3. HASIL DAN PEMBAHASAN

Dataset yang digunakan berjumlah sebanyak 364 data dengan 11 atribut. Atribut yang menjadi label klasifikasi ialah hasil uji pasien menderita penyakit anemia atau tidak. Pada penelitian ini, digunakan metode latihan *percentage split* dan *k-fold cross validation*.

#### 3.1 *Percentage Split*

Pada metode latihan *percentage split* dilakukan pembagian data latih sebanyak 75% dan 25% sisanya menjadi data uji. Hasil klasifikasi penyakit anemia menggunakan algoritma SVM dengan pembagian data latih dan data uji dengan *percentage split* akan diukur sejauh mana tingkat keberhasilannya dengan menggunakan matriks *confusion*. Matriks *confusion* yang diperoleh dari perhitungan metode SVM dengan pembagian data latih dan data uji dengan *percentage split* dapat dilihat pada Tabel 4 dibawah ini.

Tabel 4. Matriks *Confusion SVM* dengan *Percentage Split*

<b>Kelas</b>		<b>Prediksi</b>	
		<b>Positif Anemia</b>	<b>Negatif Anemia</b>
<b>Aktual</b>	<b>Positif Anemia</b>	33	6
	<b>Negatif Anemia</b>	1	51

Tabel 4 dapat dilihat bahwa sebanyak 33 kasus yang positif diprediksi positif, 6 kasus yang positif diprediksi negatif, 1 kasus yang negatif diprediksi positif, dan 51 kasus yang negatif diprediksi negatif. Selain itu, dari tabel 2 dapat dihitung nilai akurasi yang diperoleh dari total data yang berhasil diprediksi dengan tepat sebesar 92,3%. Hasil ini menunjukkan metode SVM sangat baik dalam memprediksi klasifikasi penyakit anemia. Hasil pengukuran *precision* untuk kelas positif anemia adalah sebesar 89% dan untuk kelas negatif anemia sebesar 97%. Hasil pengukuran *recall* untuk kelas positif anemia didapat sebesar 98% dan untuk kelas negatif anemia sebesar 85%. Hasil pengukuran *f1-score* untuk kelas positif anemia ialah sebesar 94% dan untuk kelas negatif anemia sebesar 90%.

#### 3.2 *K-fold Cross Validation*

Pada metode latihan *k-fold cross validation* nilai *k-fold* yang dipilih adalah 10. Matriks *confusion* yang diperoleh dari perhitungan metode SVM dengan pembagian data latih dan data uji dengan *k-fold cross validation* dapat dilihat pada Tabel 5.

Tabel 5. *Confusion Matrix SVM dengan K-fold Cross Validation*

Kelas		Prediksi	
		Positif Anemia	Negatif Anemia
Aktual	Positif Anemia	130	26
	Negatif Anemia	12	196

Dari Tabel 5 diatas dapat dilihat bahwa sebanyak 130 kasus yang positif diprediksi positif, 26 kasus yang sebenarnya positif diprediksi negatif, 12 kasus yang negatif diprediksi positif, dan 196 kasus yang negatif diprediksi negatif. Selain itu, dari tabel 3 dapat dihitung nilai akurasi yang diperoleh dari total data yang berhasil diprediksi secara benar sebesar 90%. Hasil ini menunjukkan metode SVM sangat baik dalam memprediksi klasifikasi penyakit anemia. Hasil pengukuran *precision* untuk kelas positif anemia didapat sebesar 88% dan untuk kelas negatif anemia sebesar 92%. Hasil pengukuran *recall* untuk kelas positif anemia adalah sebesar 94% dan untuk kelas negatif anemia sebesar 83%. Hasil pengukuran *f1-score* untuk kelas positif anemia ialah sebesar 91% dan untuk kelas negatif anemia ialah sebesar 87%.

### 3.3 Perbandingan Hasil Kedua Metode *Data Split*

Hasil prediksi menggunakan algoritma SVM dengan metode latihan *percentage split* dan *k-fold cross validation* menunjukkan bahwa algoritma SVM dengan kedua metode latihan tersebut baik dalam memprediksi pengklasifikasian penyakit anemia. Perbandingan hasil prediksi dari kedua metode dapat dilihat pada Tabel 6 dan Tabel 7 berikut.

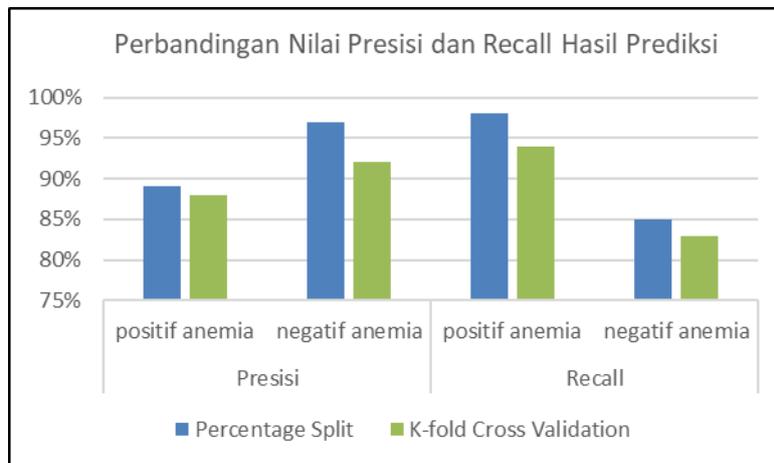
Tabel 6. Nilai Akurasi dan Presisi Metode *Percentage Split* dan *K-fold Cross Validation*

Metode	Presisi (%)		Akurasi (%)
	Positif Anemia	Negatif Anemia	
<i>Percentage Split</i>	89	97	92,3
<i>K-fold Cross Validation</i>	88	92	90

Tabel 7. Nilai *Recall* Metode *Percentage Split* dan *K-fold Cross Validation*

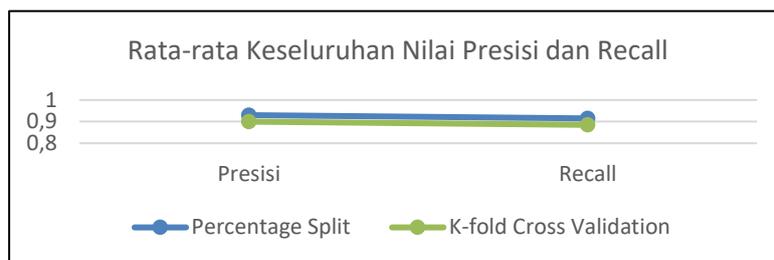
Metode	<i>Recall</i> (%)	
	Positif Anemia	Negatif Anemia
<i>Percentage Split</i>	98	85
<i>K-fold Cross Validation</i>	94	83

Berdasarkan Tabel 6 diketahui bahwa metode latihan *percentage split* memberikan nilai akurasi dan presisi yang lebih tinggi dibanding dengan metode latihan *k-fold cross validation*. Nilai *recall* yang diberikan dengan metode *percentage split* juga lebih baik dibanding dengan metode *k-fold cross validation* seperti yang tertera pada Tabel 7. Hasil presisi dan *recall* kedua metode latihan tersebut disajikan secara ringkas dalam Gambar 1.



Gambar 1. Perbandingan Nilai Presisi dan *Recall* Hasil Prediksi

Selanjutnya, perbandingan rata-rata pada nilai presisi dan *recall* dari kedua metode latih dapat dilihat pada Gambar 2.



Gambar 2. Rata-rata Keseluruhan Nilai Presisi dan *Recall*

Pada Gambar 2 terlihat bahwa nilai presisi dan *recall* dari kedua metode yang digunakan tidak berbeda jauh dan baik untuk digunakan dalam prediksi klasifikasi penyakit anemia. Selain itu, perbandingan hasil pada penelitian ini terhadap penelitian sebelumnya mengenai klasifikasi anemia dapat dilihat pada Tabel 8 berikut.

Tabel 8. Perbandingan Algoritma Klasifikasi pada Data Penyakit Anemia

Algoritma Klasifikasi	Dataset	Akurasi	Hasil		
			Presisi	Recall	F1-Score
KNN (Gowtham et al., 2020)	<i>Thalassemia and Sickle Cell Society (TSCS)</i>	0,77	0,72	0,68	0,69
Logistic Regression (Gowtham et al., 2020)	<i>Thalassemia and Sickle Cell Society (TSCS)</i>	0,78	0,77	0,68	0,69
CART (Khan et al., 2021)	<i>Bangladesh Demographic and Health Survey (BDHS)</i>	0,62	-	0,71	-

<i>Random Forest</i> (Khan et al., 2021)	<i>Bangladesh Demographic and Health Survey (BDHS)</i>	0,68	-	0,70	-
Zero R (Mohammed et al., 2020)	Dataset Anemia	0,40	-	-	-
PART (Mohammed et al., 2020)	Dataset Anemia	0,85	-	-	-
SVM (kernel <i>polynomial</i> )	Pusat Diagnostik Eureka di Lucknow, India	0,92	0,93	0,91	0,92

Dari Tabel 8 diatas dapat dilihat bahwa algoritma klasifikasi SVM dengan kernel *polynomial* memberikan hasil akurasi, presisi, *recall* dan *f1-score* yang paling tinggi dibandingkan dengan algoritma lain pada beberapa penelitian yang telah dilakukan. Hal tersebut menunjukkan bahwa algoritma SVM merupakan algoritma klasifikasi yang sangat baik untuk digunakan dalam memprediksi pengklasifikasian penyakit anemia.

#### 4. KESIMPULAM

Berdasarkan hasil yang didapatkan pada penelitian ini dapat dilihat bahwa metode support vector machine sangat baik dalam melakukan klasifikasi penyakit anemia yang dilakukan dalam penelitian ini. Selain itu, secara umum hasil pengukuran precision, recall, f1-score serta akurasi menggunakan metode percentage split lebih unggul dibanding menggunakan metode k-fold cross validation, sehingga dapat disimpulkan bahwa metode percentage split bekerja lebih baik untuk algoritma SVM pada penelitian ini.

#### REFERENSI

- [1] R. L. Rohimah, N. W. Mulyati, N. A. Shafa, R. Fitriani, and R. N. Putri, "Health Education About Anemia in Family," *Inspirasi Masy. Madani*, vol. 01, no. 01, pp. 32–41, 2021.
- [2] M. Yanto, Y. Yuhandri, and K. Khairiazaz, "Komparasi Metode Naive Bayes dan Certainty Factor untuk Mendiagnosa Penyakit Anemia," *Ilm. Komputasi*, vol. 19, no. 1, pp. 1–8, 2020.
- [3] E. Nidianti, G. Nugraha, I. A. N. Aulia, S. K. Syadzila, S. S. Suciati, and N. D. Utami, "Pemeriksaan Kadar Hemoglobin dengan Metode POCT (Point of Care Testing) sebagai Deteksi Dini Penyakit Anemia Bagi Masyarakat Desa Sumbersono, Mojokerto," *J. Surya Masy.*, vol. 2, no. 1, p. 29, 2019, doi: 10.26714/jsm.2.1.2019.29-34.
- [4] A. Nurmasani and Y. Pristyanto, "Algoritme Stacking untuk Klasifikasi Penyakit Jantung pada Dataset Imbalanced Class," *Pseudocode*, vol. 8, no. 1, pp. 21–26, 2021,

doi: 10.33369/pseudocode.8.1.21-26.

- [5] D. Derisma, “Perbandingan Kinerja Algoritma untuk Prediksi Penyakit Jantung dengan Teknik Data Mining,” *J. Appl. Informatics Comput.*, vol. 4, no. 1, pp. 84–88, 2020, doi: 10.30871/jaic.v4i1.2152.
- [6] A. M. Puspitasari, D. E. Ratnawati, and A. W. Widodo, “Klasifikasi Penyakit Gigi dan Mulut Menggunakan Metode Support Vector Machine,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 2, pp. 802–810, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [7] A. Budianto, R. Ariyuana, and D. Maryono, “Perbandingan K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM) dalam Pengenalan Karakter Plat Kendaraan Bermotor,” *J. Ilm. Pendidik. Tek. dan Kejuru.*, vol. 11, no. 1, p. 27, 2019, doi: 10.20961/jiptek.v11i1.18018.
- [8] N. Indriani, E. Rainarli, and K. E. Dewi, “Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen,” *J. Inf. Telekomunikasi Elektron.*, vol. 9, no. 4, p. 416, 2017, doi: 10.20895/infotel.v9i4.312.
- [9] F. Ma’arif and T. Arifin, “Optimasi Fitur Menggunakan Backward Elimination dan Algoritma SVM untuk Klasifikasi Kanker Payudara,” *J. Inform.*, vol. 4, no. 1, pp. 46–53, 2017.
- [10] H. Amalia, “Perbandingan Metode Data Mining Svm dan NN Untuk Klasifikasi Penyakit Ginjal Kronis,” *Pilar Nusant.*, vol. 14, no. 1, p. 1, 2018, [Online]. Available: [www.bsi.ac.id](http://www.bsi.ac.id)
- [11] C. A. Hasibuan, M. A. Mukid, and A. Prahutama, “Klasifikasi Diagnosa Penyakit Demam Berdarah Dengue (Dbd) Menggunakan Support Vector Machine (Svm) Berbasis Gui Matlab,” *J. Gaussian*, vol. 6, no. 2, pp. 171–180, 2017, [Online]. Available: <http://ejournal-s1.undip.ac.id/index.php/gaussian>
- [12] A. Pratama, R. C. Wihandika, and D. E. Ratnawati, “Implementasi Algoritme Support Vector Machine (SVM) untuk Prediksi Ketepatan Waktu Kelulusan Mahasiswa,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 4, pp. 1704–1708, 2018.
- [13] A. Mustika and M. Affandes, “Penerapan Metode Support Vector Machine dalam Klasifikasi Sentimen Tweet Public Figure,” in *Sentra*, 2017, pp. 978–979.
- [14] T. M. P. Aulia, N. Arifin, and R. Mayasari, “Perbandingan Kernel Support Vector Machine (SVM) dalam Penerapan Analisis Sentimen Vaksinasi Covid-19,” *SINTECH (Science Inf. Technol. J.)*, vol. 4, no. 2, pp. 139–145, 2021, doi: 10.31598/sintechjournal.v4i2.762.
- [15] R. Saragih, T. D. Sunoto, J. J. Jarden, and D. M. Hanif, “Penerapan Kernel Jamak pada Program Berbasis PCA untuk Pengenalan Wajah dengan Variasi Iluminasi,” *J. Ecotipe (Electronic, Control. Telecommun. Information, Power Eng.)*, vol. 7, no. 2, pp. 85–91, 2020, doi: 10.33019/jurnalecotipe.v7i2.1902.