Limits: Journal of Mathematics and Its Applications
E-ISSN: 2579-8936
P-ISSN: 1829-605X
Vol. 22, No. 3, Nopember 2025, 267-280
DOI: https://doi.org/10.12962/limits.v22i3.5737

Perbandingan Kinerja *Hybrid Classification* SVM-RF dan SVM-NN Terhadap Faktor Risiko Anemia Ibu Hamil di Indonesia dengan Pendekatan *Clustering* K-Means

Asyifah Qalbi [1] *, Erfiani [2], Budi Susetyo [3]
[1,2,3]IPB University; Jl. Raya Dramaga, Kabupaten Bogor, 16680
[1,2,3]Statistika dan Sains Data, IPB University, Indonesia
e-mail: asyifahqalbi@apps.ipb.ac.id

Diajukan: 2 September 2025, Diperbaiki: 8 Agustus 2025, Diterima: 29 Agustus 2025

Abstrak

Klasifikasi merupakan salah satu topik yang paling banyak diteliti oleh para peneliti dari bidang *machine learning* dan *data mining*. Metode *machine learning* yang sering digunakan antara lain Support Vector Machine (SVM), Random Forest (RF) dan Neural Network (NN). Namun, SVM tidak selalu memberikan nilai akurasi yang baik. Sebagai contoh, ketika diterapkan pada data yang sangat tidak seimbang, SVM akan mengalami tantangan. Selain itu, tidak terdapat satu metode terbaik yang bisa diterapkan untuk semua masalah klasifikasi. Saat ini, pendekatan metode *hybrid* untuk penggunaan *data mining* menjadi semakin populer seperti metode *hybrid* SVM-RF, SVM-NN dan KMeans-SVM. Pada penelitian ini, metode *hybrid* SVM-RF dan SVM-NN digunakan untuk mengklasifikasikan faktor risiko anemia pada ibu hamil di Indonesia dengan pendekatan K-Means untuk mengelompokkan data yang salah klasifikasi oleh SVM. Hasil penelitian menunjukkan bahwa metode *hybrid* dapat meningkatkan kinerja model SVM. *Hybrid* SVM-RF memberikan nilai metrik evaluasi yang lebih tinggi dibandingkan dengan SVM-NN. Empat metrik evaluasi yang digunakan, yaitu *accuracy*, *balanced accuracy*, *sensitivity* dan *specificity* pada SVM-RF masing-masing bernilai sebesar 0,989; 0,989; 0,988; dan 0,989. Peubah yang berkontribusi secara umum berdasarkan SHAP Global terhadap klasifikasi faktor risiko anemia pada ibu hamil secara berurutan adalah Usia, Tablet Fe, Status Bekerja, Pendidikan, Status Gizi dan ANC.

Kata Kunci: *Hybrid*, Klasifikasi, SVM-RF, SVM-NN, Anemia

Abstract

Classification is one of the most researched topics by researchers from the field of machine learning and data mining. Machine learning methods that are often used include Support Vector Machine (SVM), Random Forest (RF) and Neural Network (NN). However, SVM does not always provide good accuracy. For example, when applied to highly imbalanced data, SVM will experience challenges. In addition, there is no single best method that can be applied to all classification problems. Currently, hybrid method approaches for data mining applications are becoming increasingly popular such as hybrid SVM-RF, SVM-NN and KMeans-SVM methods. In this study, a hybrid method of SVM-RF and SVM-NN was used to classify risk factors for anemia in pregnant women in Indonesia with a K-Means approach to cluster data misclassified by SVM. The results showed that the hybrid method can improve the performance of the SVM model. Hybrid SVM-RF provides a higher evaluation metric value compared to SVM-NN. The four evaluation metrics used, namely accuracy, balanced accuracy, sensitivity and specificity in SVM-RF are worth 0,989; 0,989; 0,988; and 0,989, respectively. The variables that contribute generally based on SHAP Global to the classification of risk factors for anemia in pregnant women in order are Age, Fe Tablet, Working Status, Education, Nutritional Status and ANC.

Keywords: *Hybrid, Classification, SVM-RF, SVM-NN, Anemia*

1 Pendahuluan

Klasifikasi merupakan salah satu topik yang banyak diteliti oleh para peneliti pada bidang *machine learning* dan *data mining* [1]. Klasifikasi merupakan cara mengelompokkan objek berdasarkan ciri – ciri yang dimiliki oleh objek [2]. *Machine learning* (ML) adalah bidang kecerdasan buatan yang membahas tentang pembangunan sistem komputasi adaptif. ML mempelajari cara membangun dan menganalisis algoritma yang mampu melakukan generalisasi dari data yang tersedia [3].

ML terbagi menjadi dua yaitu *supervised machine learning* dan *unsupervised machine learning*. Beberapa algoritma *supervised machine learning* yang sering digunakan yaitu Support Vector Machine (SVM), Random Forest (RF) dan Neural Network (NN) [2]. Salah satu jenis *unsupervised machine learning* yang sering digunakan adalah *clustering* dengan algoritma K-Means [4]. SVM merupakan algoritma pembelajaran yang menganalisis data untuk klasifikasi dan analisis regresi. Selain melakukan klasifikasi linier, SVM secara efisien dapat digunakan untuk melakukan klasifikasi non-linier dengan bantuan kernel yang secara implisit memetakan input ke dalam ruang peubah berdimensi tinggi [5]. Penelitian [6] menggunakan SVM untuk klasifikasi pada produksi minyak menyimpulkan bahwa SVM bisa signifikan meningkatkan presisi prediksi. Pada penelitian yang dilakukan oleh [7] menyimpulkan bahwa SVM berkinerja dengan baik dalam mengklasifikasikan 100 gambar dari dataset CIFAR-10 dengan menunjukkan hasil yang efisien yaitu akurasi sebesar 85,80%.

Meskipun demikian, SVM tidak selalu memberikan nilai akurasi yang baik ketika diterapkan pada data yang lain. Penelitian [8] dan [9] melakukan klasifikasi menggunakan SVM masing-masing menghasilkan akurasi sebesar 51,8% dan 52,1%. Kedua penelitian ini menggunakan data yang sangat tidak seimbang. Menurut [10], data yang sangat tidak seimbang akan membuat model bias pada kelas mayoritas dan menyebabkan akurasi secara keseluruhan akan rendah. Hal lain yang menjadi tantangan bagi SVM adalah ketika diterapkan pada data yang bervariasi dan berdimensi tinggi[11]. Selain itu, tidak terdapat satu metode terbaik untuk semua masalah klasifikasi. Setiap metode memiliki keterbatasan dan kinerjanya bergantung pada karakteristik data serta kemampuan algoritma yang digunakan [12].

Saat ini, pendekatan metode *hybrid* untuk penggunaan alat data mining menjadi semakin populer [12]. Pada penelitian yang dilakukan oleh [13] yang melakukan perbandingan metode SVM, KNN dan SVM-KNN dalam memprediksi solar flare menyimpulkan bahwa metode *hybrid* SVM-KNN memberikan hasil prediksi yang lebih baik dibandingkan dengan kedua metode tunggal tersebut.

Penelitian yang dilakukan oleh [12] membandingkan metode *hybrid* SVM-KNN dengan SVM-RF menyimpulkan bahwa SVM-RF bekerja lebih baik dalam melakukan klasifikasi dibandingkan dengan SVM-KNN. RF yang merupakan representasi dari algoritma berbasis Decision Tree adalah teknik *ensemble* berbasis bagging yang menghasilkan pengambilan sampel *bootstrap* [14]. RF membangun banyak subset (*bootstrap sampling*) dari data pelatihan dan melatih algoritma yang sama beberapa kali. Prediksi yang dihasilkan ditentukan sebagai rata-rata dari semua prediksi submodel. Seiring dengan bertambahnya jumlah pohon, RF dapat menghindari *overfitting* dan tidak terlalu terpengaruh oleh *outlier* [14].

Pada penelitian yang dilakukan oleh [15] menggunakan metode *hybrid* SVM-NN menghasilkan tingkat akurasi klasifikasi sebesar 98,08%. Arsitektur NN mencakup lapisan *input*, *hidden* dan *output* [16]. *Artificial Neural Network* (ANN) adalah pengklasifikasi pola yang sangat baik karena kemampuannya untuk mempelajari pola yang tidak dapat dipisahkan secara linier dan konsep yang berurusan dengan ketidakpastian, kebisingan dan kejadian acak [17]. Selain itu, penelitian [18] juga menggunakan model *hybrid* K-Means dengan SVM untuk memprediksi kanker menghasilkan nilai akurasi 99%. Algoritma K-Means merupakan salah satu algoritma pengelompokan yang paling sering digunakan karena prosesnya yang relatif cepat, mudah dipahami dan diimplementasikan [19]. Sehingga berdasarkan beberapa penelitian di atas, penggunaan metode klasifikasi dengan *hybrid* ML dapat meningkatkan tingkat akurasi hasil prediksi.

Anemia merupakan suatu kondisi tubuh yang ditandai dengan hasil pemeriksaan kadar hemoglobin (Hb) dalam darah lebih rendah dari normal. Anemia pada saat kehamilan akan meningkatkan risiko komplikasi perdarahan, melahirkan bayi Berat Badan Lahir Rendah (BBLR), Panjang Badan Lahir Rendah (PBLR) dan premature [20]. Berdasarkan laporan Riset Kesehatan Dasar (Riskesdas) 2018, prevalensi ibu hamil sebesar 48,9% [21]. Artinya, 5 dari 10 orang ibu hamil mengalami anemia. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja model *hybrid* SVM-RF dan SVM-NN pada data anemia ibu hamil di Indonesia dengan pendekatan K-Means untuk mengelompokkan data yang salah klasifikasi oleh SVM dan memperoleh peubah yang berkontribusi pada pembentukan klasifikasi status anemia ibu hamil di Indonesia. Hasil dari penelitian ini diharapkan bisa memberikan kontribusi dalam pembuatan kebijakan oleh pihak terkait.

2 Metode Penelitian

2.1 Data

Data yang digunakan pada penelitian ini adalah data kesehatan yang terhimpun dalam Riskesdas 2018 yang diintegrasikan dengan pelaksanaan Survei Sosial Ekonomi Nasional (Susenas) Maret 2018 oleh Badan Pusat Statistik (BPS). Amatan yang akan diteliti yaitu dari kelompok ibu hamil sebanyak 3783 yang tersebar di 34 Provinsi di Indonesia. Data ini digunakan untuk mengidentifikasi peubah penting yang akan digunakan pada model. Hasil prediksi algoritma yang digunakan akan dibandingkan dengan hasil pemeriksaan ibu hamil. Tabel 1 memberikan informasi peubah yang digunakan pada data empiris.

Tabel 1. Peubah Data Empiris

	Peubah	**Tipe Data**	**Keterangan**
Y	Status Anemia	Kategorik	1 = Ya, 2 = Tidak
X1	Usia	Numerik	Tahun
X2	Konsumsi tablet Fe	Numerik	Tablet
X3	Pemeriksaan kehamilan (ANC)	Kategorik	1 = Ya, 2 = Tidak
X4	Status gizi (LILA)	Kategorik	1 = KEK, 2 = Normal
X5	Status bekerja	Kategorik	1 = Tidak, 2 = Ya
X6	Pendidikan	Kategorik	1 = Tidak/belum pernah sekolah 2 = Tidak tamat SD/MI 3 = Tamat SD/MI 4 = Tamat SLTP/MTS 5 = Tamat SLTA/MA 6 = Tamat D1/D2/D3 7 = Tamat PT

2.2 Metode Analisis Data

Analisis data menggunakan *software* R Studio. Adapun tahapan analisis yang dilakukan adalah sebagai berikut.

1. Melakukan eksplorasi data
2. Pemodelan klasifikasi
   a. Membagi data menjadi 2 bagian yaitu data latih 80% dan data uji 20%.
   b. Melakukan *preprocessing* data.
   c. Menerapkan SMOTE pada data latih.
   d. Membangun model dengan algoritma SVM-RF dan SVM-NN
      1) Membangun model SVM menggunakan data latih dengan kernel RBF [22].
      2) Mengidentifikasi data yang salah diklasifikasikan oleh SVM dengan menggunakan nilai epsilon (ξ).

         Persamaan nilai ξ diberikan oleh [23]

$$\xi_j = \max\{1 - y_j f(x_j), 0\} \quad (1)$$

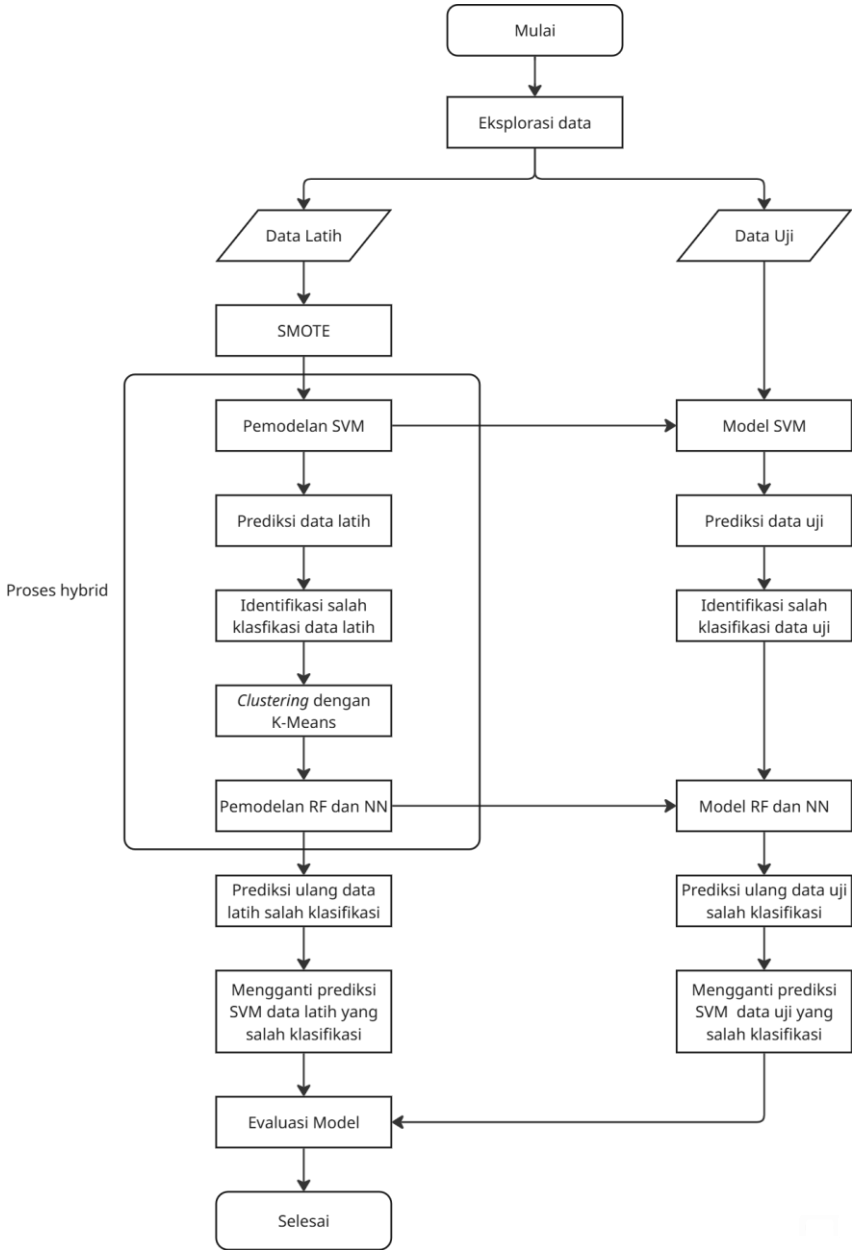
$f(x_j)$ merupakan fungsi keputusan. Jika suatu objek memiliki nilai $\xi_j > 1$ maka objek tersebut salah klasifikasi [24]. Persamaan fungsi keputusan didefinisikan oleh [25]

$$f(x_j) = \sum_{i=1}^{n} y_i \alpha_i k(x_i, x_j) + b \quad (2)$$

3) Menerapkan K-Means pada data yang salah diklasifikasikan menjadi beberapa klaster.
4) Data dari setiap klaster tersebut kemudian diimplementasikan ke dalam RF dan NN untuk dilatih.
5) Melakukan klasifikasi ulang pada data yang salah klasifikasi menggunakan algoritma RF dan NN berdasarkan klaster terdekat.
6) Mengganti data yang salah diklasifikasikan SVM dengan data yang sudah diklasifikasikan ulang oleh RF dan NN.

3. Melakukan evaluasi kinerja model pada data uji dengan menggunakan metrik statistik berdasarkan pada *confusion matriks* yaitu *accuracy, sensitivity, specificity* dan *balanced accuracy*.

Analisis data yang telah dipaparkan digambarkan pada Gambar 1.



Gambar 1. Diagram Alir Analisis Data

3 Hasil dan Pembahasan

3.1 Eksplorasi data

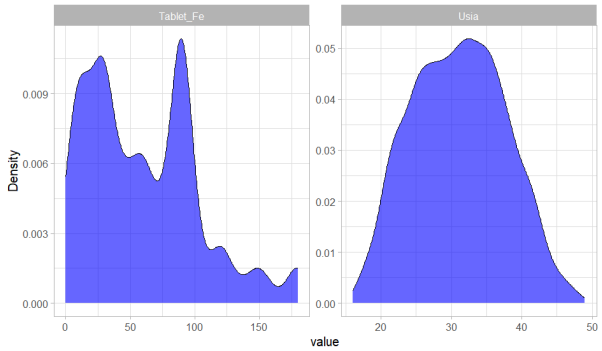
Data empiris yang digunakan sebanyak 3783 amatan, 3366 amatan berkategori tidak anemia sedangkan 417 merupakan kategori anemia. Hal ini menandakan bahwa terjadi ketidakseimbangan kelas, proporsi kelas mayoritas yaitu tidak anemia sebesar 89% sedangkan kelas minoritas yaitu anemia sebesar 11%. Pada peubah numerik, distribusi peubah usia (X1) tampak mengikuti sebaran normal sehingga dilakukan pengujian hipotesis untuk menguji asumsi distribusi sebagai berikut:

H_0: Peubah usia berdistribusi normal

H_1: Peubah usia berdistribusi tidak normal

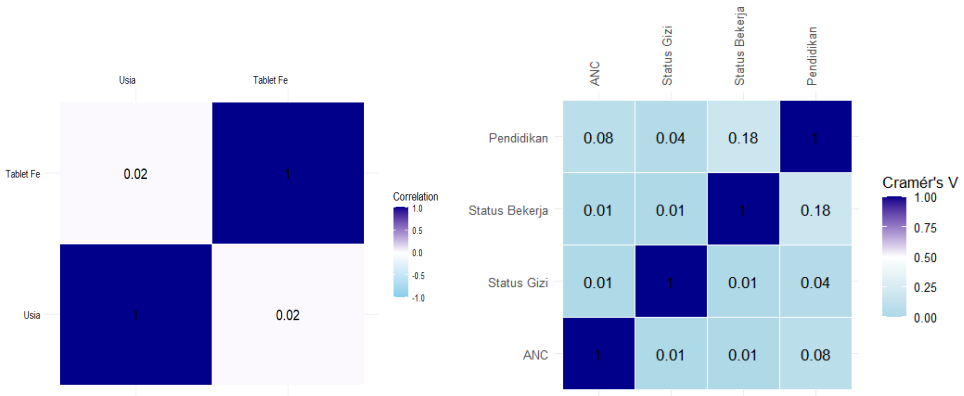
Tolak H_0 jika *p-value* < 0,05

Berdasarkan analisis menggunakan uji Shapiro-Wilk dan Kolmogorov-Smirnov diperoleh hasil pada kedua uji memberikan *p-value* < 0,05. Hal ini menyimpulkan bahwa tolak H_0 artinya peubah usia tidak berdistribusi normal, walaupun pada visualnya terlihat simetris dan menyerupai distribusi normal. Peubah konsumsi tablet fe (X2) sebaran nilainya skew kanan dengan nilai skewnya sebesar 0,71. Visualisasinya ditunjukkan pada Gambar 2.



Gambar 2. Plot Distibusi X1 dan X2

Korelasi antar peubah numerik ditunjukkan pada Gambar 3(a) sedangkan korelasi pada peubah kategorik ditunjukkan pada Gambar 3(b)



(a) Peubah Tipe Numerik (b) Peubah Tipe Kategorik

Gambar 3. Korelasi Peubah Prediktor

Pada peubah tipe numerik menggunakan korelasi spearman untuk melihat hubungan antara usia dan tablet fe. Berdasarkan pada hasil analisis yang telah dilakukan, Gambar 3(a) menunjukkan kedua peubah tersebut memiliki nilai korelasi sebesar 0,02 artinya hubungannya kecil. Sedangkan peubah tipe kategorik ditunjukkan oleh Gambar 3(b) menggunakan cramer's V menghasilkan nilai korelasi yang juga kecil yaitu kurang dari 0,2 artinya hubungan peubah kategorik tersebut juga kecil. Nilai korelasi yang rendah menunjukkan bahwa model yang dibangun tidak menghadapi masalah multikolinearitas yang serius. Sebaliknya, jika nilai korelasi tinggi maka model yang dibangun tidak cukup baik karena dapat membuat model tidak stabil. Model tidak mampu

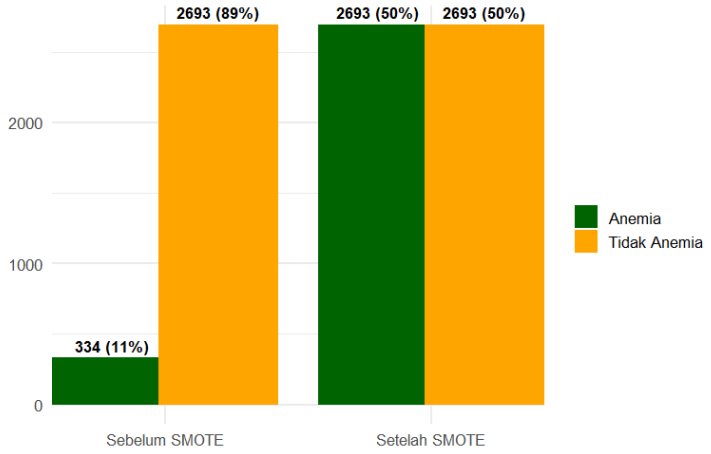
mengenal efek individual masing-masing peubah terhadap target sehingga menurunkan kinerja model [26].

Korelasi tinggi pada data empiris umumnya disebabkan oleh adanya hubungan nyata atau kedekatan konsep antar peubah. Kondisi seperti ini dapat menyulitkan model untuk memisahkan kontribusi masing-masing peubah karena informasi yang diberikan tumpang tindih. Akibatnya, model berisiko membangun pola yang tidak stabil dan interpretasi terhadap peubah prediktor berpotensi bias. Pada data yang digunakan dalam penelitian ini (Gambar 3), korelasi antar peubah prediktor baik numerik maupun kategorik relatif rendah sehingga risiko multikolinearitas dari data empiris tergolong kecil.

3.2 Pemodelan dan Evaluasi

Pemodelan data dilakukan dengan menggunakan metode *hybrid* SVM-RF dan SVM-NN. *Prepocessing data* yang dilakukan yaitu standarisasi, pengkodean ulang dan penyeimbangan kelas. Standarisasi dilakukan menggunakan z-score untuk menyamakan skala data, agar setiap data memiliki peluang yang sama untuk berkontribusi pada model. Standarisasi dilakukan pada data latih terlebih dahulu untuk mendapatkan nilai mean dan standar deviasi, kemudian nilai ini digunakan ke data latih atau data baru agar tidak terjadi kebocoran data. Pengkodean ulang dilakukan menggunakan *one hot encoding*, hal ini karena model *machine learning* hanya dapat belajar pada data numerik [27]. Penyeimbangan kelas menggunakan SMOTE, hal ini dilakukan agar model dapat belajar dengan baik pada kedua kelas.

Gambar 4 adalah visualisasi sebelum dan setelah data latih dilakukan penyeimbangan. Sebelum dilakukan penyeimbangan, jumlah data latih kategori anemia sebanyak 334 amatan. Setelah dilakukan penyeimbangan, jumlah amatan antara kategori anemia dengan tidak anemia sudah seimbang yaitu masing-masing sebanyak 2693 amatan.



Gambar 4. Penyeimbangan Kelas Pada Data Latih

Tahapan selanjutnya, melakukan pemodelan menggunakan SVM dengan kernel RBF. Hasil evaluasi model menggunakan SVM diberikan pada Tabel 2.

Tabel 2. Evaluasi Model SVM

Accuracy	*Sensitivity*	*Specificity*	*Balanced accuracy*
0,646	0,470	0,667	0,569

Berdasarkan pada Tabel 2, diperoleh hasil metrik evaluasi model SVM masih cukup rendah karena data yang diprediksi dengan benar hanya berkisar 64%. Kinerja model SVM belum cukup baik dalam melakukan klasifikasi faktor risiko anemia pada ibu hamil karena nilai *sensitivity* yang rendah (47%), serta nilai *balanced accuracy* sebesar 0,569 menunjukkan kemampuan rata-rata model dalam mengklasifikasikan kedua kelas masih cukup rendah.

Tahap selanjutnya, mengidentifikasi data yang salah klasifikasi menggunakan nilai epsilon. Jika nilai epsilon data yang sudah diprediksi oleh SVM > 1 maka data tersebut dikelompokkan sebagai data yang salah klasifikasi. Dataset yang salah klasifikasi tersebut selanjutnya dilakukan *clustering* menggunakan K-Means, diperoleh empat klaster menggunakan metode elbow. Keempat klaster ini memuat peubah-peubah data dengan karakteristik yang berbeda dari data yang salah diklasifikasikan oleh SVM, sehingga RF dan NN masing-masing memiliki 4 model. Keempat model RF dan NN digunakan untuk memprediksi ulang data yang salah diklasifikasikan oleh SVM, dengan penentuan model berdasarkan klaster terdekat dari setiap titik data. Tabel 3 menunjukkan hasil klasifikasi model *hybrid* SVM-RF dan SVM-NN, diperoleh hasil bahwa pada semua metrik evaluasi menunjukkan metode *hybrid* mampu meningkatkan kinerja model SVM. 98% data bisa diprediksi dengan benar oleh kedua metode.

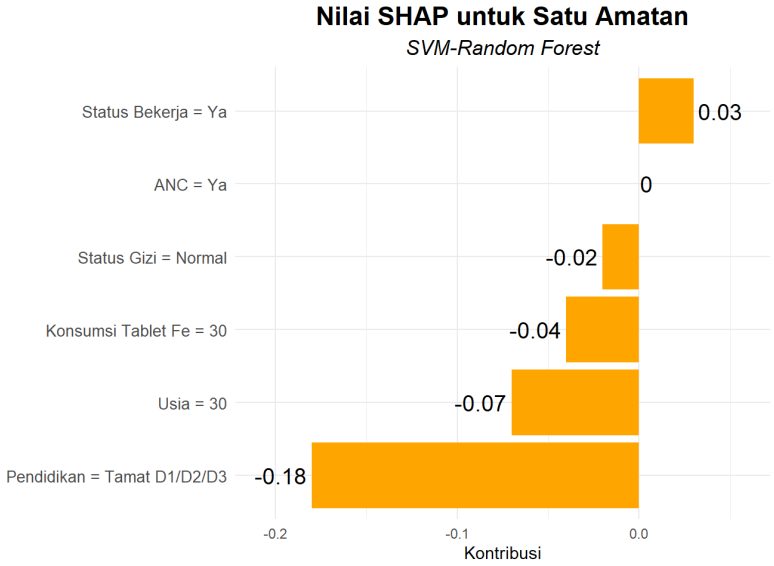
Tabel 3. Metrik Evaluasi Model *Hybrid*

Model	*Accuracy*	*Sensitivity*	*Specificity*	*Balanced accuracy*
SVM-RF	0,989	0,988	0,989	0,989
SVM-NN	0,985	0,952	0,989	0,971

SVM-RF menunjukkan metrik evaluasi yang lebih tinggi dibandingkan dengan SVM-NN. Nilai *sensitivity* SVM-RF menunjukkan kemampuan memprediksi kelas positif atau pada penelitian ini sebagai kelas minoritas lebih tinggi 3% dibandingkan dengan SVM-NN. Selain itu, Nilai *balanced accuracy* SVM-RF sebesar 0,989 atau 98,9%, menunjukkan bahwa kemampuan rata-rata model dalam memprediksi kelas anemia dan tidak anemia sangat baik. RF memiliki kemampuan secara efisien dalam memproses data dengan berbagai karakteristik data dan kelas [12].

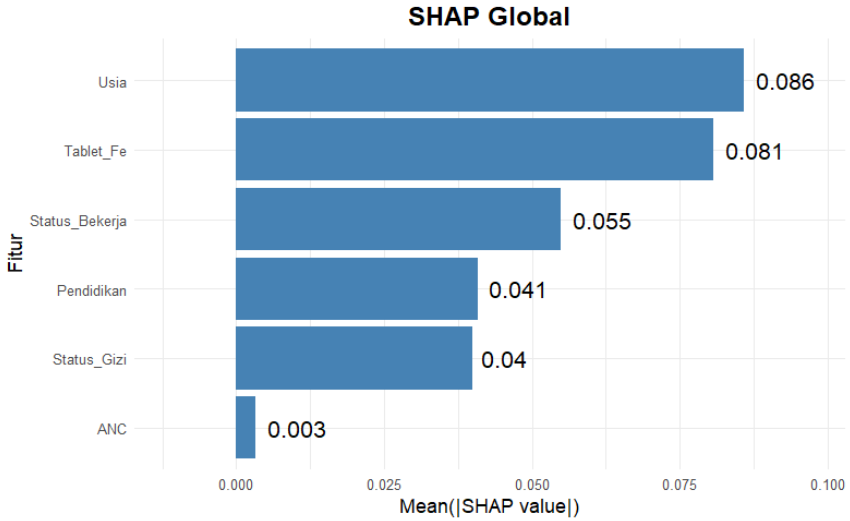
Analisis kontribusi peubah terhadap kinerja model dilakukan menggunakan SHAP. Analisis SHAP mengidentifikasi peubah-peubah yang memberikan kontribusi positif dan negatif. Kontribusi positif artinya meningkatkan kecenderungan prediksi ke arah ibu hamil anemia.

Kontribusi negatif bermakna peubah tersebut mendorong ke arah prediksi ibu hamil tidak anemia. Pada Gambar 5 menunjukkan ilustrasi interpretasi secara lokal setiap peubah terhadap satu amatan menggunakan nilai SHAP.



Gambar 5. SHAP Lokal

Berdasarkan pada Gambar 5, Status bekerja memberikan kontribusi positif terhadap kemungkinan ibu hamil masuk dalam kategori anemia, hal tersebut sejalan dengan hasil penelitian yang dilakukan oleh [28]. Status gizi normal, Konsumsi Tablet fe sebanyak 30, usia 30 tahun dan Pendidikan dengan lulusan D1/D2/D3 berkontribusi negatif terhadap kemungkinan ibu hamil masuk dalam kategori anemia. Hal ini sesuai dengan beberapa penelitian yang dilakukan oleh [28], [29], [30]. ANC tidak memberi kontribusi apapun terhadap kemungkinan ibu hamil masuk dalam kategori anemia. Hasil analisis SHAP Lokal ini hanya berlaku untuk satu amatan, sehingga perlu analisis lebih lanjut untuk seluruh amatan.



Gambar 6. SHAP Global

Gambar 6 menunjukkan SHAP Global yaitu rata-rata kontribusi peubah dari seluruh amatan terhadap kinerja model *hybrid* SVM-RF dalam melakukan klasifikasi anemia pada ibu hamil di Indonesia. Berdasarkan hasil analisis diperoleh peubah yang memberikan kontribusi yang paling besar ke kontribusi paling kecil terhadap klasifikasi faktor risiko anemia pada ibu hamil berturut-turut adalah Usia, Tablet Fe, Status Bekerja, Pendidikan, Status Gizi dan ANC. Usia dan Tablet fe menjadi peubah yang memiliki kontribusi besar. Berdasarkan pada jurnal [29] ibu hamil berisiko tinggi mengalami anemia ketika (usia <20 dan >35 tahun) dan (konsumsi tablet fe <30 tablet pada semester 1, < 60 tablet pada semester 2 dan <90 tablet pada semester 3). Ibu hamil berisiko rendah ketika (usia 20-35 tahun) dan (konsumsi ≥30 tablet pada semester 1, ≥60 tablet pada semester 2 dan ≥90 tablet pada semester 3) [29].

4 Simpulan

Model SVM-RF menunjukkan kinerja yang sedikit lebih baik dibandingkan dengan SVM-NN. *Hybrid* SVM-RF menunjukkan kinerja yang sangat baik. Hasil keempat metrik evaluasinya yaitu *accuracy* sebesar 0,989; *sensitivity* sebesar 0,988; *specificity* sebesar 0,989 dan *balanced accuracy* sebesar 0,989. Pada analisis kontribusi peubah menggunakan SHAP Global, peubah yang memberikan kontribusi yang paling besar terhadap prediksi anemia pada ibu hamil adalah Usia dan Tablet Fe.

Saran untuk penelitian selanjutnya yaitu mengembangkan model dengan penambahan peubah prediktor CRP dan serum feritin serta mengembangkan pendekatan *hybrid* dengan mendeteksi amatan yang berpotensi salah klasifikasi menggunakan REL-U (*Relative Uncertainty*). Lampiran koding analisi dapat diunduh pada link https://drive.google.com/file/d/1jUAInNpIpuCesFg8ybVYUcOmLymCKdHP/view?usp=sharing

5 Daftar Pustaka

[1] A. A. Soofi and A. Awan, “Classification Techniques in Machine Learning: Applications and Issues,” *Journal of Basic & Applied Sciences*, vol. 13, pp. 459–465, 2017.

[2] F. Y. Osisanwo, J. E. T. Akinsola, O. Awodele, J. O. Hinmikaiye, O. Olakanmi, and J. Akinjobi, “Supervised Machine Learning Algorithms: Classification and Comparison,” *International Journal of Computer Trends and Technology*, vol. 48, no. 3, pp. 128–138, 2017, doi: 10.14445/22312803/ijctt-v48p126.

[3] J. Dj Novakovi, A. Veljovi, S. S. Ili, ˇ Zeljko Papi, and M. Tomovi, "Evaluation of Classification Models in Machine Learning," *Theory and Applications of Mathematics & Computer Science*, vol. 7, no. 1, pp. 39–46, 2017.

[4] S. Naeem, A. Ali, S. Anam, and M. M. Ahmed, "An Unsupervised Machine Learning Algorithms: Comprehensive Review," *International Journal of Computing and Digital Systems*, vol. 13, no. 1, pp. 911–921, 2023, doi: 10.12785/ijcds/130172.

[5] B. Mahesh, "Machine Learning Algorithms - A Review," *International Journal of Science and Research (IJSR)*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/ART20203995.

[6] M. Minghui and Z. Chuanfeng, "Application of Support Vector Machines to a Small-Sample Prediction," *Advances in Petroleum Exploration and Development*, vol. 10, no. 2, pp. 72–75, 2015, doi: 10.3968/7830.

[7] B. Raj S., S. L. Gibson Moses, G. Chinnadurai, J. V. Anand, V. Janakiraman, and P. Anand, "Support Vector Machine for Image Classification," *Journal of the Asiatic Society of Mumbai*, vol. 97, no. 4, pp. 1–10, 2023, [Online]. Available: https://www.researchgate.net/publication/371492398

[8] S. Annisa, A. Saefuddin, and Erfiani, "Blood Glucose Levels of Non Invasive Tool Using Support Vector Machine on Multi-Class Imbalanced Data," *Int J Sci Eng Res*, vol. 9, no. 11, pp. 1649–1652, 2018.

[9] Y. D. Evitasari, W. J. Pranoto, and N. A. Verdikha, "Evaluasi Support Vector Machine Dengan Optimasi Metode Genetic Algorithm Pada Klasifikasi Banjir Kota Samarinda," *Jurnal Sains Komputer dan Teknologi Informasi*, vol. 6, no. 1, pp. 49–53, 2023, doi: 10.33084/jsakti.v6i1.5462.

[10] S. Rezvani, F. Pourpanah, C. P. Lim, and Q. M. J. Wu, "Methods for class-imbalanced learning with support vector machines: a review and an empirical evaluation," *Soft comput*, 2024, doi: 10.1007/s00500-024-09931-5.

[11] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, "An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review," *Information (Switzerland)*, vol. 15, no. 4, 2024, doi: 10.3390/info15040235.

[12] L. A. Demidova, I. A. Klyueva, and A. N. Pylkin, "Hybrid approach to improving the results of the SVM classification using the random forest algorithm," *Procedia Comput Sci*, vol. 150, pp. 455–461, 2019, doi: 10.1016/j.procs.2019.02.077.

[13] R. Li, H. N. Wang, H. He, Y. M. Cui, and Z. Le Du, "Support vector machine combined with K-nearest neighbors for solar flare forecasting," *Chinese Journal of Astronomy and Astrophysics*, vol. 7, no. 3, pp. 441–447, 2007, doi: 10.1088/1009-9271/7/3/15.

[14] G. W. Cha, H. J. Moon, and Y. C. Kim, "Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables," *Int J Environ Res Public Health*, vol. 18, no. 16, Aug. 2021, doi: 10.3390/ijerph18168530.

[15] P. Nanglia, S. Kumar, A. N. Mahajan, P. Singh, and D. Rathee, "A hybrid algorithm for lung cancer classification using SVM and Neural Networks," *ICT Express*, vol. 7, no. 3, pp. 335–341, 2021, doi: 10.1016/j.icte.2020.06.007.

[16] C. Y. Ku, C. Y. Liu, Y. J. Chiu, and W. Da Chen, "Deep Neural Networks with Spacetime RBF for Solving Forward and Inverse Problems in the Diffusion Process," *Mathematics*, vol. 12, no. 9, 2024, doi: 10.3390/math12091407.

[17] B. Debska and B. Guzowska-Świder, "Application of artificial neural network in food classification," *Anal Chim Acta*, vol. 705, no. 1–2, pp. 283–291, 2011, doi: 10.1016/j.aca.2011.06.033.

[18] R. Kumar, G. Ganapathy, and J.-J. Kang, "A Hybrid Mod K-Means Clustering with Mod SVM Algorithm to Enhance the Cancer Prediction," *International Journal of Internet, Broadcasting and Communication*, vol. 13, no. 2, pp. 231–243, 2021, [Online]. Available: http://dx.doi.org/10.7236/IJIBC.2021.13.2.231

[19] Y. P. Raykov, A. Boukouvalas, F. Baig, and M. A. Little, "What to do when K-means clustering fails: A simple yet principled alternative algorithm," *PLoS One*, vol. 11, no. 9, pp. 1–28, 2016, doi: 10.1371/journal.pone.0162259.

[20] Kemenkes RI, *Buku Saku Pencegahan Anemia Pada Ibu Hamil Dan Remaja Putri*, vol. 5, no. 4. 2023. [Online]. Available: http://dx.doi.org/10.1016/j.snb.2010.05.051

[21] Kemenkes, *Laporan Riskesdas 2018 Nasional.pdf*. 2018.

[22] L. A. Demidova, "Two-stage hybrid data classifiers based on svm and knn algorithms," *Symmetry (Basel)*, vol. 13, no. 4, 2021, doi: 10.3390/sym13040615.

[23] Y. Y. Colin Campbell, *Learning with Support Vector Machines (Synthesis Lectures on Artificial Intelligence and Machine Learning)*. 2011.

[24] X. Wang and P. M. Pardalos, "A Survey of Support Vector Machines with Uncertainties," *Annals of Data Science*, vol. 1, no. 3–4, pp. 293–309, 2014, doi: 10.1007/s40745-014-0022-8.

[25] B. Scholkopf and A. J. Smola, *(book)Learning with Kernels*, vol. 53, no. 9. London, England: The MIT Press, 2002.

[26] K. I. Sundus, B. H. Hammo, M. B. Al-Zoubi, and A. Al-Omari, "Solving the multicollinearity problem to improve the stability of machine learning algorithms applied

to a fully annotated breast cancer dataset," *Inform Med Unlocked*, vol. 33, no. September, p. 101088, 2022, doi: 10.1016/j.imu.2022.101088.

[27] F. Bolikulov, R. Nasimov, A. Rashidov, F. Akhmedov, and Y. I. Cho, "Effective Methods of Categorical Data Encoding for Artificial Intelligence Algorithms," *Mathematics*, vol. 12, no. 16, 2024, doi: 10.3390/math12162553.

[28] G. K. Dewi, I. Istianah, and S. Septiani, "Analisis Risiko Anemia Pada Ibu Hamil," *Jurnal Ilmiah Kesehatan (JIKA)*, vol. 4, no. 1, pp. 67–80, 2022, doi: 10.36590/jika.v4i1.223.

[29] I. Tanziha, M. R. M. Damanik, L. J. Utama, and R. Rosmiati, "Faktor Risiko Anemia Ibu Hamil Di Indonesia," *Jurnal Gizi dan Pangan*, vol. 11, no. 2, pp. 143–152, 2016, doi: 10.25182/jgp.2016.11.2.%p.

[30] G. S. Yulfani Putri, S. Sulistiawati, and M. A. Cahya Laksana, "Analisis faktor-faktor risiko anemia pada ibu hamil di Kabupaten Gresik tahun 2021," *Jurnal Riset Kebidanan Indonesia*, vol. 6, no. 2, pp. 119–129, 2023, doi: 10.32536/jrki.v6i2.220.