Institut Riset dan Publikasi Indonesia (IRPI)
MALCOM: Indonesian Journal of Machine Learning and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/malcom
Vol. 5 Iss. 3 July 2025, pp: 972-980
ISSN(P): 2797-2313 | ISSN(E): 2775-8575

Application of Ensemble Learning Technique for Classification of Anemia Types

Penerapan Teknik *Ensemble Learning* untuk Klasifikasi Jenis-jenis Anemia

Arjuna Priandika[1], Auliya Rahman Isnain[2*]

[1,2]Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Lampung

E-Mail: [1]priandikaarjuna@gmail.com, [2]auliyarahman@teknokrat.ac.id

Received Oct 16th 2024; Revised Jun 18th 2025; Accepted Jul 12th 2025; Available Online Jul 31th 2025, Published Jul 31th 2025
Corresponding Author: Arjuna Priandika

Abstract

Anemia is a medical condition that requires accurate diagnosis for effective treatment. This research explores the application of ensemble learning techniques, specifically the StackingClassifier, for the classification of anemia types. This technique combines three basic models: Random Forest, K-Nearest Neighbors (KNN), and Gradient Boosting, with Logistic Regression as the final estimator. The medical data used involved various hematological features, and preprocessing included cleaning, normalization, and data sharing. Model evaluation was performed using cross-validation with 10 folds. The results showed that the StackingClassifier achieved an overall accuracy of 98%, with excellent precision and recall in most classes. Classes such as Iron deficiency anemia, Leukemia, and Other microcytic anemia showed 100% precision, while some classes with small samples experienced lower recall. Overall, the model is effective in classifying anemia types with high accuracy and can be adapted to further improve medical diagnosis. This study highlights the potential of ensemble techniques in improving classification performance and suggests further exploration on data with uneven distributions.

Keyword: Anemia, Ensemble Learning, Medical Classification, Random Forest, StackingClassifier

Abstrak

Anemia merupakan kondisi medis yang memerlukan diagnosis yang akurat untuk penanganan yang efektif. Penelitian ini mengeksplorasi penerapan teknik *ensemble learning*, khususnya *stacking classifier*, untuk klasifikasi jenis-jenis anemia. Teknik ini menggabungkan tiga model dasar: *Random Forest*, *K-Nearest Neighbors* (*KNN*), dan *Gradient Boosting*, dengan *Logistic Regression* sebagai estimator akhir. Data medis yang digunakan melibatkan berbagai fitur hematologi, dan *preprocessing* meliputi pembersihan, normalisasi, serta pembagian data. Evaluasi model dilakukan menggunakan *cross-validation* dengan 10 lipatan. Hasil penelitian menunjukkan bahwa *stacking classifier* mencapai akurasi keseluruhan 98%, dengan *precision* dan *recall* yang sangat baik di sebagian besar kelas. Kelas-kelas seperti *Iron deficiency anemia*, *Leukemia*, dan *Other microcytic anemia* menunjukkan *precision* 100%, sementara beberapa kelas dengan sampel kecil mengalami *recall* yang lebih rendah. Secara keseluruhan, model ini efektif dalam mengklasifikasikan jenis-jenis anemia dengan akurasi tinggi dan dapat diadaptasi untuk meningkatkan diagnosis medis lebih lanjut. Penelitian ini menyoroti potensi teknik *ensemble* dalam memperbaiki performa klasifikasi dan menyarankan eksplorasi lebih lanjut pada data dengan distribusi yang tidak merata.

Kata Kunci: Anemia, *Ensemble Learning, Stacking Classifier,* Klasifikasi Medis, *Random Forest*

1. PENDAHULUAN

Anemia adalah salah satu kondisi medis yang paling umum di dunia, mempengaruhi jutaan orang dari berbagai usia dan latar belakang. Kondisi ini ditandai oleh penurunan jumlah sel darah merah atau kadar hemoglobin dalam darah, yang berfungsi untuk mengangkut oksigen ke seluruh tubuh [1]. Ketika jumlah sel darah merah atau hemoglobin berkurang, tubuh tidak mendapatkan oksigen yang cukup, yang dapat mengakibatkan berbagai gejala seperti kelelahan, pusing, sesak napas, dan dalam kasus yang parah, dapat

menyebabkan kerusakan organ atau kematian . Identifikasi dini dan akurat terhadap jenis anemia sangat penting untuk memberikan pengobatan yang tepat dan mengurangi risiko komplikasi lebih lanjut [2].

Terdapat berbagai jenis anemia yang berbeda penyebab dan dampaknya, termasuk *anemia defisiensi besi, anemia megaloblastik, anemia hemolitik, anemia aplastik,* dan anemia akibat penyakit kronis. Masing-masing jenis ini membutuhkan pendekatan pengobatan yang berbeda [3]. Misalnya, anemia defisiensi besi diobati dengan suplementasi zat besi, sementara anemia yang disebabkan oleh penyakit kronis mungkin memerlukan penanganan penyakit yang mendasarinya [4]. Oleh karena itu, diagnosis yang tepat mengenai jenis anemia merupakan langkah kritis dalam pengelolaan pasien.

Kemajuan dalam teknologi informasi, terutama dalam bidang analisis data dan *machine learning*, telah membuka peluang baru untuk meningkatkan akurasi dan efisiensi diagnosis medis. Di antara berbagai pendekatan dalam *machine learning*, teknik *ensemble learning* telah menjadi salah satu yang paling menjanjikan. *Ensemble learning* merupakan metode yang menggabungkan beberapa model prediksi untuk menghasilkan hasil yang lebih akurat dan stabil dibandingkan dengan model tunggal [5]. Dengan menggabungkan berbagai algoritma yang berbeda, seperti *Random Forest, K-Nearest Neighbors (KNN),* dan *Gradient Boosting, ensemble learning* dapat mengatasi kelemahan yang ada pada masing-masing model dan meningkatkan kemampuan generalisasi terhadap data yang belum pernah dilihat sebelumnya [6].

Random Forest, misalnya, adalah teknik *ensemble* yang menggunakan kumpulan pohon keputusan untuk melakukan klasifikasi. Setiap pohon dalam hutan memprediksi kelas, dan kelas yang paling banyak dipilih oleh semua pohon menjadi prediksi akhir [7]. Penggunaan *Random Forest* dapat mengurangi risiko *overfitting* yang sering terjadi pada pohon keputusan tunggal, karena ia memanfaatkan banyak model untuk menghasilkan keputusan yang lebih stabil dan *robust*. Selain itu, *Random Forest* juga memiliki kemampuan untuk menangani data dengan jumlah fitur yang besar dan kompleks, menjadikannya pilihan yang baik untuk analisis data medis yang seringkali melibatkan banyak *variable* [8]. *KNN* adalah metode *non-parametrik* yang mengklasifikasikan data berdasarkan kedekatannya dengan data lain dalam ruang fitur. Meskipun konsepnya sederhana, *KNN* efektif dalam menangani masalah klasifikasi dengan distribusi data yang kompleks dan tidak memerlukan asumsi distribusi data sebelumnya. Kelebihan *KNN* terletak pada kemampuannya dalam menangani *dataset* yang tidak linier dan fleksibilitasnya dalam berbagai jenis data [9].

Dalam konteks klasifikasi jenis anemia, penggunaan teknik *ensemble* ini dapat menawarkan solusi yang lebih akurat dibandingkan metode tradisional. Beberapa penelitian sebelumnya telah mengaplikasikan teknik seperti *Random Forest*, *KNN*, dan *Gradient Boosting* untuk klasifikasi anemia, yang menunjukkan peningkatan akurasi prediksi. Misalnya, penelitian oleh El-Boghdady et al., (2023) menggunakan *Random Forest* untuk mendiagnosis anemia dan melaporkan bahwa teknik ini memberikan akurasi yang lebih tinggi dibandingkan dengan metode *SVM* (*Support Vector Machine*) [10]. Penelitian lainnya oleh Faradila et al., (2025) membandingkan *KNN* dengan metode *regresi logistik* dalam klasifikasi anemia, dan menemukan bahwa *KNN* lebih efektif dalam menangani *dataset* dengan distribusi yang tidak linier [11]. Sedangkan dalam penelitian oleh Noviandy et al., (2024), *Gradient Boosting* terbukti lebih unggul dalam meningkatkan akurasi klasifikasi pada data medis yang tidak seimbang, seperti pada kasus anemia dengan jumlah data pasien yang terbatas [12].

Penelitian ini mengusulkan penggunaan gabungan dari ketiga teknik *ensemble* tersebut, yang belum banyak dieksplorasi secara bersamaan dalam klasifikasi anemia. Pendekatan ini memungkinkan pemanfaatan kekuatan masing-masing model dalam menangani kompleksitas data medis, seperti variasi dalam parameter hematologi, serta mengurangi risiko *bias* atau kesalahan yang mungkin muncul jika hanya menggunakan satu model. Dengan demikian, penelitian ini berfokus pada integrasi informasi dari berbagai fitur medis yang relevan, yang sering kali tidak dapat dilakukan oleh model tunggal dengan tingkat akurasi yang sama. Hal ini menjadi *novelty* dari penelitian ini, karena belum ada studi yang secara komprehensif mengkombinasikan ketiga teknik *ensemble* untuk klasifikasi anemia dengan pendekatan yang lebih *holistic*. Dengan menggabungkan data medis yang relevan dengan model *ensemble learning* yang canggih, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan terhadap peningkatan diagnosis anemia. Selain itu, hasil penelitian ini diharapkan dapat membuka jalan bagi penerapan yang lebih luas dari teknik *ensemble learning* dalam bidang medis lainnya, mendukung perkembangan teknologi diagnostik yang lebih akurat dan efisien. Melalui penelitian ini, diharapkan juga dapat ditemukan metode baru yang dapat diadopsi oleh klinisi untuk meningkatkan pengelolaan dan pengobatan anemia serta kondisi medis lainnya yang kompleks.

2. METODE PENELITIAN

2.1. *Ensemble Learning*

Penelitian ini memanfaatkan teknik *ensemble learning*, yang menggabungkan beberapa model pembelajaran mesin untuk meningkatkan akurasi prediksi dibandingkan dengan model tunggal. Teknik *ensemble* bekerja dengan mengkombinasikan keputusan dari berbagai model untuk menghasilkan prediksi akhir yang lebih stabil dan akurat [13]. Beberapa metode *ensemble* yang digunakan dalam penelitian ini adalah *Random Forest, Gradient Boosting,* dan *AdaBoost*.

Selain itu, penelitian ini juga menerapkan *Stacking Classifier*, sebuah teknik *ensemble* yang menggabungkan beberapa model dasar (*base learners*) melalui sebuah model meta (*meta-learner*) untuk menghasilkan prediksi akhir [14]. alam pendekatan *stacking*, prediksi dari model-model dasar digunakan sebagai fitur *input* bagi model meta, yang kemudian memberikan keputusan akhir. Teknik ini memungkinkan pemanfaatan kelebihan dari berbagai model dasar yang berbeda untuk mengatasi kelemahan masing-masing, sehingga meningkatkan akurasi dan *robustness* dari hasil prediksi. Pada penelitian ini, model dasar yang digunakan adalah *Random Forest*, *KNN*, dan *Gradient Boosting*, dengan *Logistic Regression* sebagai *estimator akhir* atau *meta-learner* yang mengintegrasikan *output* dari ketiga *model dasar* tersebut.

2.2. *Dataset*

Dataset yang digunakan dalam penelitian ini merupakan data hitung darah lengkap (*Complete Blood Count* atau *CBC*) yang telah diberi label dengan jenis anemia. Data ini dikumpulkan dari berbagai pemeriksaan *CBC* dan telah didiagnosis secara manual. *Dataset* ini diperoleh dari Kaggle, yang tersedia di link https://www.kaggle.com/datasets/ehababoelnaga/anemia-types-classification. *Dataset* ini mencakup beberapa fitur medis penting yang relevan untuk klasifikasi anemia [15], termasuk:

Penelitian ini memanfaatkan teknik *ensemble learning*, yang menggabungkan beberapa model pembelajaran mesin untuk meningkatkan akurasi prediksi dibandingkan dengan model tunggal. Teknik *ensemble* bekerja dengan mengkombinasikan keputusan dari berbagai model untuk menghasilkan prediksi akhir yang lebih stabil dan akurat [16]. Beberapa metode ensemble yang digunakan dalam penelitian ini adalah *Random Forest, Gradient Boosting,* dan *AdaBoost*.

1. *Hemoglobin (HGB)*: Jumlah *hemoglobin* dalam darah, yang penting untuk transportasi oksigen.
2. *Platelet (P1T)*: Jumlah *trombosit* dalam darah, yang terlibat dalam proses pembekuan darah.
3. *White Blood Cell (WBC)*: Jumlah sel darah putih, yang vital untuk respon imun tubuh.
4. *Red Blood Cell (RBC)*: Jumlah sel darah merah, yang bertanggung jawab untuk transportasi oksigen.
5. *Mean Corpuscular Volume (MCV)*: Volume rata-rata dari satu sel darah merah.
6. *Mean Corpuscular Hemoglobin (MCH)*: Jumlah rata-rata *hemoglobin* per sel darah merah.
7. *Mean Corpuscular Hemoglobin Concentration (MCHC)*: Konsentrasi rata-rata *hemoglobin* dalam sel darah merah.
8. *Platelet Distribution Width (PDW)*: Pengukuran variabilitas dalam distribusi ukuran *trombosit* dalam darah.
9. *Procalcitonin (PCT)*: Tes *prokalsitonin* yang dapat membantu dalam diagnosis sepsis akibat infeksi bakteri atau untuk menilai risiko tinggi pengembangan *sepsis*.
10. Diagnosis: Jenis anemia berdasarkan parameter-parameter *CBC*.

Dataset ini mencakup beberapa jenis anemia, yang dapat mencakup *anemia defisiensi besi*, *anemia megaloblastik*, *anemia hemolitik*, dan jenis-jenis lain yang didiagnosis berdasarkan parameter-parameter *CBC* yang tersedia. Setiap jenis anemia dikategorikan secara eksplisit dalam fitur "Diagnosis" yang menjadi target klasifikasi dalam penelitian ini.

2.3. *Pre-processing* data

Preprocessing data adalah langkah krusial dalam pipeline analisis data yang bertujuan untuk mempersiapkan data mentah menjadi format yang dapat diolah oleh model pembelajaran mesin [8]. Proses ini mencakup beberapa tahap penting yang memastikan kualitas dan konsistensi data, serta meminimalkan bias yang dapat mempengaruhi hasil analisis. Langkah-langkah dalam preprocessing data ini meliputi pembersihan data, normalisasi data, dan pembagian data, yang semuanya penting untuk memastikan integritas data dan efektivitas model yang dikembangkan.

1. Pembersihan Data

Tahap pertama dalam *preprocessing* adalah pembersihan data, yang melibatkan identifikasi dan penanganan entri yang tidak lengkap atau tidak valid [17]. Data yang hilang atau tidak konsisten dapat menyebabkan distorsi dalam model pembelajaran mesin jika tidak ditangani dengan benar. Untuk menjaga kualitas *dataset*, entri dengan nilai yang hilang melebihi ambang batas tertentu dihapus. Ini dilakukan untuk memastikan bahwa data yang digunakan dalam pelatihan model adalah data yang representatif dan lengkap. Untuk nilai hilang yang tersisa, teknik *imputasi* digunakan untuk menggantikan nilai-nilai yang hilang dengan nilai yang masuk akal, seperti rata-rata atau median dari fitur tersebut [18]. *Imputasi* ini bertujuan untuk mempertahankan integritas statistik dari data sambil menghindari penghapusan data yang berpotensi penting.

2. Normalisasi Data

Setelah data dibersihkan, langkah berikutnya adalah normalisasi data. Normalisasi adalah proses transformasi fitur-fitur dalam *dataset* sehingga memiliki rentang nilai yang seragam, yang penting untuk

memastikan bahwa model pembelajaran mesin tidak terpengaruh oleh skala fitur yang berbeda [19]. Normalisasi dilakukan dengan menggunakan teknik *min-max scaling*, di mana setiap nilai fitur dinormalisasi ke dalam rentang [0, 1] menggunakan persamaan 1.

$$x' = \frac{x - min(X)}{max(X) - min(X)} \quad (1)$$

di mana *x* adalah nilai fitur asli, *"min" (X)* dan *max(X)* adalah nilai minimum dan maksimum dari fitur tersebut. Proses ini memastikan bahwa semua fitur memiliki kontribusi yang seimbang dalam model dan memudahkan proses pelatihan.

3. Pembagian Data

Pembagian data adalah tahap yang membagi *dataset* menjadi dua subset: set pelatihan dan set pengujian. Set pelatihan digunakan untuk melatih model, sementara set pengujian digunakan untuk mengevaluasi performa model pada data yang tidak terlihat sebelumnya. Pembagian ini biasanya dilakukan dengan proporsi 80% untuk set pelatihan dan 20% untuk set pengujian [20]. Pembagian ini bertujuan untuk memastikan bahwa model dilatih dengan data yang representatif dan diuji dengan data yang relevan untuk mengukur kemampuannya dalam generalisasi.

2.4. Implementasi Model

Penelitian ini menerapkan tiga teknik pembelajaran mesin yang berbeda yaitu *Random Forest, KNN,* dan *Gradient Boosting* untuk klasifikasi jenis-jenis anemia berdasarkan data hitung darah lengkap *CBC*. Masing-masing model memiliki karakteristik dan pendekatan yang berbeda dalam menangani masalah klasifikasi, dan implementasi serta evaluasi dilakukan untuk menentukan model mana yang memberikan kinerja terbaik.

Random Forest dibangun dengan 10 pohon keputusan. Pada setiap *split*, model memilih fitur terbaik dari subset fitur yang dipilih secara acak. Prediksi akhir dari *Random Forest* adalah hasil voting mayoritas dari semua pohon menggunakan persamaan 2.

$$\hat{y} = \text{mode}\{\text{h}_\text{m}(\text{x})\}_{\text{m=1}}^{\text{M}} \quad (2)$$

di mana $\hat{y}$ adalah prediksi akhir, *M* adalah jumlah pohon dalam hutan, dan $h_m(x)$adalah prediksi dari pohon ke-m untuk *input* x.

Model *KNN* dibangun dengan menentukan nilai K, yaitu jumlah tetangga terdekat yang akan dipertimbangkan untuk melakukan klasifikasi. Jarak yang paling umum digunakan adalah *Euclidean distance*, yang didefinisikan seperti pada persamaan 3.

$$d(x, x') = \sqrt{\sum_{i=1}^{n}(x_i - x_i')^2} \quad (3)$$

di mana x dan x' adalah dua vektor fitur, dan n adalah jumlah fitur.

Persamaan umum *KNN* dapat dirumuskan pada persamaan 4.

$$\hat{y} = \text{mode}\{y_1, y_2, \ldots, y_K\} \quad (4)$$

di mana $y_1, y_2, \ldots, y_K$ adalah label kelas dari tetangga terdekat berdasarkan jarak *Euclidean*, dan $\hat{y}$ adalah prediksi kelas dari data x. *KNN* adalah algoritma non-parametrik yang bekerja dengan membandingkan titik data baru dengan tetangga-tetangganya yang paling dekat. Metode ini digunakan karena kemampuannya untuk menangani data yang tidak linier tanpa perlu asumsi distribusi data sebelumnya. Seperti yang dijelaskan oleh Sadrabadi et al., (2020) "*KNN* adalah metode yang sangat intuitif, sederhana, dan efektif dalam berbagai aplikasi klasifikasi" [21].

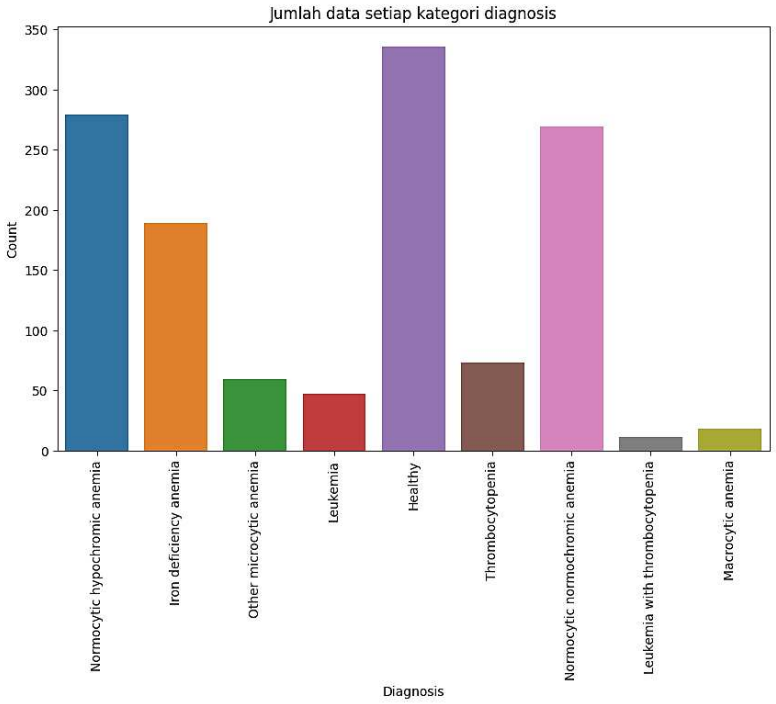
Gradient Boosting dibangun dengan serangkaian pohon keputusan sebagai model dasar, di mana setiap pohon dilatih untuk meminimalkan *residual error* dari model sebelumnya. Proses ini dilakukan dalam beberapa iterasi, dengan setiap iterasi memperbaiki kesalahan prediksi dari iterasi sebelumnya [22]. Model *Gradient Boosting* dihasilkan dengan menambahkan kontribusi dari masing-masing pohon yang ditimbang oleh *learning rate* (η) seperti pada rumus 5.

$$F_M(x) = F_{M-1}(x) + \eta \cdot h_M(x) \quad (5)$$

3. HASIL DAN PEMBAHASAN

3.1. *Exploratory Data Analysis*

Untuk memperoleh wawasan lebih lanjut tentang distribusi jumlah data dalam setiap kategori diagnosis, sebuah *plot histogram* digunakan. *Plot* ini menggambarkan frekuensi atau jumlah sampel yang tersedia untuk masing-masing kategori diagnosis, memberikan visualisasi yang jelas tentang ketidakseimbangan dalam *dataset*. Hasil visualisasi distribusi jumlah data per kategori diagnosis dapat dilihat Gambar 1.



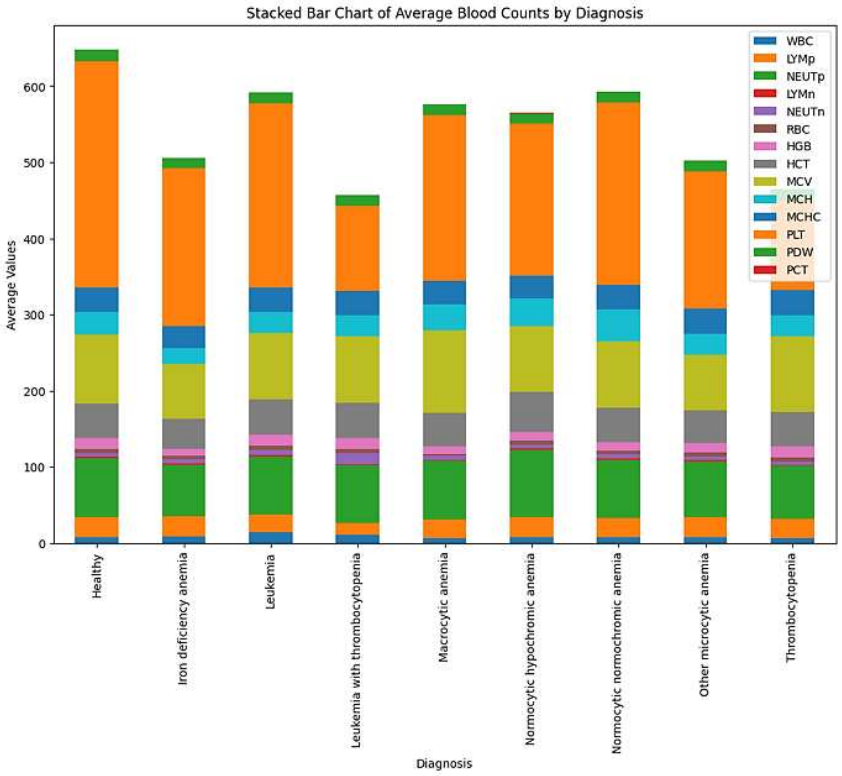
Gambar 1. Distribusi Data per Kategori

Hasil analisis distribusi jumlah data untuk setiap kategori diagnosis menunjukkan ketidakmerataan yang signifikan antara kelas-kelas dalam *dataset.* Kategori *Healthy* memiliki jumlah sampel tertinggi yaitu 336, diikuti oleh *Normocytic hypochromic anemia* dengan 279 sampel dan *Normocytic normochromic anemia* dengan 269 sampel. Jumlah data yang lebih besar pada kategori-kategori ini memungkinkan model untuk mempelajari karakteristik dengan lebih mendalam dan menghasilkan klasifikasi yang lebih akurat untuk kondisi sehat serta *anemia normositik*. Sebaliknya, kategori *Iron deficiency anemia*, meskipun memiliki jumlah sampel yang relatif lebih besar dengan 189 sampel, masih kurang dibandingkan dengan kategori utama, namun tetap memberikan informasi yang memadai untuk klasifikasi.

Sebaliknya, beberapa kategori menunjukkan jumlah data yang sangat rendah, seperti *Leukemia with thrombocytopenia* dengan hanya 11 sampel, *Macrocytic anemia* dengan 18 sampel, dan *Leukemia* dengan 47 sampel. Kategori-kategori ini mengalami ketidakseimbangan yang signifikan dibandingkan dengan kategori lainnya, yang dapat mempengaruhi performa model dalam hal akurasi dan generalisasi. Kategori dengan data terbatas mungkin menghadapi risiko *overfitting* atau *underfitting*, sehingga mengharuskan penggunaan teknik khusus seperti *oversampling* untuk kategori minor atau *undersampling* untuk kategori mayor untuk menangani ketidakseimbangan ini secara efektif.

Untuk memberikan wawasan lebih mendalam mengenai distribusi fitur hematologi di berbagai *diagnosis* anemia, berikut ini ditampilkan serangkaian *plo*t distribusi yang dapat dilihat pada Gambar 2. Berdasarkan Gambar 2, analisis rata-rata fitur berdasarkan diagnosis menunjukkan perbedaan signifikan dalam profil hematologi antar kategori. Untuk jumlah sel darah putih *(WBC), "Leukemia"* memiliki nilai tertinggi, menandakan peningkatan jumlah sel darah putih khas pada kondisi ini, sedangkan *"Macrocytic anemia"* dan *"Thrombocytopenia"* menunjukkan nilai lebih rendah. Persentase *limfosit (LYMp)* cenderung rendah pada *"Leukemia with thrombocytopenia"* dan lebih tinggi pada *"Iron deficiency anemia"*. Persentase *neutrofil (NEUTp)* sangat tinggi pada *"Normocytic hypochromic anemia"*, menandakan respon inflamasi yang kuat. Nilai *RBC* dan *HGB* lebih rendah pada *"Iron deficiency anemia"* dan *"Macrocytic anemia",* sesuai dengan ciri khas anemia tersebut. *MCV* menunjukkan nilai sangat tinggi pada *"Macrocytic anemia"*, mencerminkan sel darah merah yang lebih besar. *MCH* dan *MCHC* tertinggi pada *"Normocytic normochromic anemia",* menandakan konsentrasi *hemoglobin* yang lebih tinggi. *PLT* sangat rendah pada

"Leukemia with thrombocytopenia" dan tinggi pada kategori lainnya. *PDW* dan *PCT* menunjukkan variasi yang lebih besar pada *"Thrombocytopenia",* mencerminkan perubahan dalam ukuran dan proporsi *trombosit.*



Gambar 2. Distribusi Rata-rata Fitur

3.2. Hasil Implementasi Model

Dalam penelitian ini, teknik *ensemble learning* diterapkan menggunakan *Stacking Classifier* yang merupakan teknik *ensemble* yang menggabungkan beberapa model pembelajaran dasar (*base models*) melalui model meta (*meta-learner*) untuk meningkatkan akurasi prediksi. Dalam penelitian ini, tiga model dasar digunakan, yaitu *Random Forest, KNN,* dan *Gradient Boosting*. *Output* dari ketiga model ini kemudian menjadi *input* bagi model *Logistic Regression* yang berperan sebagai *meta-learner*. *Logistic Regression* mengintegrasikan prediksi dari ketiga model dasar untuk menghasilkan prediksi akhir yang lebih akurat. Konfigurasi dari model yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Konfigurasi Model

Komponen	Konfigurasi
Random Forest	n_estimators=10, random_state=42
K-Nearest Neighbors	n_neighbors=10
Gradient Boosting	Standart
Estimator	Logistic Regression
Cross-Validation	Cv=10

Berdasarkan Tabel 1, implementasi teknik *ensemble learning* melalui *Stacking Classifier* dalam penelitian ini melibatkan tiga model dasar, yaitu *Random Forest, KNN,* dan *Gradient Boosting*. Konfigurasi model dasar *Random Forest* diatur dengan 10 pohon keputusan (*n_estimators=10*), bersama dengan pengaturan *random_state=42* untuk memastikan hasil yang konsisten dan terukur. Model ini efektif dalam menangani variabilitas data dengan memanfaatkan keputusan dari berbagai pohon. *KNN*, diatur dengan 10 tetangga terdekat (*n_neighbors=10*), berfungsi dengan mengklasifikasikan data berdasarkan kedekatannya dengan tetangga terdekat, yang memungkinkan pemodelan hubungan lokal dalam data. Sementara itu, *Gradient Boosting* menggunakan konfigurasi standar untuk meningkatkan akurasi melalui metode *boosting*, yang memperbaiki prediksi model sebelumnya dengan menambahkan model-model baru yang mengurangi kesalahan.

Estimator akhir dalam konfigurasi ini adalah *Logistic Regression*, yang mengintegrasikan prediksi dari ketiga model dasar untuk menghasilkan keputusan akhir yang lebih akurat. *Logistic Regression* bertindak

sebagai model akhir yang menyaring dan menggabungkan hasil dari model-model dasar, memastikan prediksi yang lebih terintegrasi dan lebih baik. Evaluasi model dilakukan menggunakan teknik *cross-validation* dengan 10 lipatan (*cv=10*). Metode ini memastikan evaluasi yang *robust* dan mencegah *overfitting*, dengan memberikan gambaran yang lebih akurat tentang kemampuan model dalam menggeneralisasi pada data yang tidak terlihat sebelumnya. Dengan konfigurasi ini, *Stacking Classifier* mampu menggabungkan kekuatan model-model dasar untuk mencapai hasil klasifikasi yang optimal, menunjukkan performa yang tinggi dalam mengidentifikasi berbagai jenis anemia secara efektif.

3.3. Hasil Evaluasi Model

Penelitian ini menerapkan *Stacking Classifier* untuk klasifikasi jenis-jenis anemia, yang menggabungkan *Random Forest Classifier, KNN Classifier,* dan *Gradient Boosting Classifier* dengan *Logistic Regression* sebagai estimator akhir. StackingClassifier menggabungkan tiga model pembelajaran mesin (*Random Forest Classifier, KNN Classifier,* dan *Gradient Boosting Classifier*) dengan *Logistic Regression* sebagai model final. Hasil evaluasi dari *Stacking Classifier* berdasarkan laporan klasifikasi dapat dilihat pada Tabel 2.

Tabel 2. Hasil Evaluasi *Ensemble Learning*

Kelas	Precision	Recall	F1-Score
Healthy	95%	100%	98%
Iron deficiency anemia	100%	97%	99%
Leukemia	100%	100%	100%
Leukemia with thrombocytopenia	100%	67%	80%
Macrocytic anemia	100%	67%	80%
Normocytic hypochromic anemia	100%	98%	80%
Normocytic normochromic anemia	98%	100%	99%
Other microcytic anemia	100%	100%	100%
Thrombocytopenia	95%	95%	95%
Accuracy			98%
Macro Average	99%	91%	94%
Weighted Average	98%	98%	98%

Berdasarkan Tabel 2, menunjukkan performa yang sangat baik dalam klasifikasi jenis-jenis anemia dengan akurasi keseluruhan mencapai 98%. Model ini menunjukkan *precision* tertinggi untuk *kelas Iron deficiency anemia, Leukemia,* dan *Other microcytic anemia,* dengan nilai precision 100%, menandakan bahwa model sangat akurat dalam mengidentifikasi jenis-jenis anemia ini tanpa banyak *false positives*. *Precision* untuk kelas *Healthy* adalah 95%, menunjukkan ketepatan yang sangat baik dalam mengklasifikasikan kondisi sehat.

Dalam hal *recall, Healthy* mencapai nilai sempurna 100%, yang menunjukkan bahwa model berhasil mendeteksi semua kasus sehat dalam *dataset*. Kelas *Leukemia* juga memiliki *recall* 100%, menandakan deteksi penuh untuk semua kasus *leukemia.* Namun, beberapa kelas dengan jumlah sampel kecil, seperti *Leukemia with thrombocytopenia* dan *Macrocytic anemia*, menunjukkan *recall* yang lebih rendah 67%, mengindikasikan bahwa model menghadapi tantangan dalam mendeteksi semua kasus dari kelas-kelas ini.

F1-Score, yang menggabungkan *precision* dan *recall,* menunjukkan performa yang sangat baik di sebagian besar kelas. Kelas *Leukemia, Other microcytic anemia,* dan *Iron deficiency anemia* mencapai nilai *F1-Score* tertinggi yaitu 100%, mencerminkan keseimbangan yang sangat baik antara *precision* dan *recall*. Sebaliknya, *Leukemia with thrombocytopenia* dan *Macrocytic anemia* memiliki *F1-Score* yang lebih rendah 80%, menunjukkan adanya *trade-off* antara *precision* dan *recall* dalam klasifikasi kelas-kelas ini. Secara keseluruhan, *Stacking Classifier* membuktikan kemampuannya dalam mengklasifikasikan jenis-jenis anemia dengan akurasi dan keandalan yang tinggi. Meskipun terdapat beberapa tantangan dalam mendeteksi kasus-kasus dari kelas dengan jumlah sampel kecil, model ini secara efektif menangani klasifikasi multi-kelas dengan performa yang sangat memuaskan.

4. KESIMPULAN

Penelitian ini berhasil mencapai tujuan utamanya, yaitu mengembangkan model klasifikasi yang akurat untuk mendeteksi berbagai jenis anemia menggunakan teknik *ensemble learning* dengan *Stacking Classifier.* Hasil yang diperoleh menunjukkan bahwa kombinasi model *Random Forest, KNN*, dan *Gradient Boosting* yang diintegrasikan melalui *Logistic Regression* mampu memberikan akurasi yang tinggi, dengan performa yang konsisten dalam berbagai pengujian menggunakan *cross-validation.* Berdasarkan evaluasi terbaik yang diperoleh, model *Stacking Classifier* menunjukkan akurasi keseluruhan mencapai 98%. Model ini memiliki *precision* tertinggi pada kelas *Iron deficiency anemia, Leukemia,* dan *Other microcytic anemia,* masing-masing dengan nilai *precision* 100%, yang menandakan kemampuan model yang sangat akurat dalam mengidentifikasi jenis-jenis anemia ini tanpa banyak *false positives.* Selain itu, *recall* terbaik tercatat pada

kelas *Healthy* dan *Leukemia*, keduanya dengan nilai 100%, yang menunjukkan bahwa model berhasil mendeteksi semua kasus pada kelas ini. *F1-Score* juga mencapai 100% untuk kelas *Iron deficiency anemia, Leukemia,* dan *Other microcytic anemia*, yang mencerminkan keseimbangan yang sangat baik antara *precision* dan *recall.* Namun, kelas dengan jumlah sampel kecil, seperti *Leukemia with thrombocytopenia* dan *Macrocytic anemia*, menunjukkan *recall* lebih rendah pada 67% dan *F1-Score* 80%, mencerminkan tantangan dalam mendeteksi semua kasus dari kelas-kelas ini. Secara keseluruhan, *Stacking Classifier* terbukti efektif dalam mengklasifikasikan berbagai jenis anemia dengan akurasi dan keandalan yang tinggi. Penelitian ini menunjukkan potensi besar teknik *ensemble learning* dalam mengatasi masalah klasifikasi medis yang kompleks, dan dapat dilanjutkan dengan pengujian lebih lanjut pada *dataset* yang lebih besar serta eksperimen dengan model dasar lain untuk meningkatkan performa.

REFERENSI

[1] R. R. Muradova, "Modern approaches to the pharmacotherapy of anemia," *International journal of medical sciences and clinical research*, vol. 4, no. 6, pp. 89–94, 2024, doi: 10.37547/ijmscr/volume04issue06-14.

[2] M. L. Smith, "Anemia: Evaluation of Suspected Anemia," *FP Essent*, vol. 530, pp. 7–11, 2023.

[3] V. Yu. Pavlova and E. V Kazakovtseva, "Anemia of chronic diseases," *Vestnik Volgogradskogo gosudarstvennogo medicinskogo universiteta*, 2024, doi: 10.19163/1994-9480-2024-21-2-21-28.

[4] R. Allende, L. D. de Entresotos, and S. C. Diez, "Anaemia of chronic diseases: Pathophysiology, diagnosis and treatment.," *Med Clin (Barc)*, vol. 156, no. 5, pp. 235–242, 2021, doi: 10.1016/J.MEDCLE.2020.07.022.

[5] L.-Y. Wu and S.-S. Weng, "Ensemble Learning Models for Food Safety Risk Prediction," *Sustainability*, vol. 13, no. 21, p. 12291, 2021, doi: 10.3390/SU132112291.

[6] A. Das, K. Jain, D. Majumder, and R. Karmakar, "An Ensemble Learning Paradigm for Efficient Stock Market Predictions," pp. 1–7, 2024, doi: 10.1109/icccnt61001.2024.10726086.

[7] Z. Khan *et al.*, "Ensemble of optimal trees, random forest and random projection ensemble classification," *Advanced Data Analysis and Classification*, vol. 14, no. 1, pp. 97–116, 2020, doi: 10.1007/S11634-019-00364-9.

[8] M. Maindola *et al.*, "Utilizing Random Forests for High-Accuracy Classification in Medical Diagnostics," pp. 1679–1685, 2024, doi: 10.1109/ic3i61595.2024.10828609.

[9] G. F. C. de Castro and R. Tinós, "K-Nearest Neighbors based on the Nk Interaction Graph," 2022, doi: 10.5753/eniac.2022.227174.

[10] A. M. El-Boghdady, S. Kishk, M. M. Ashour, and E. AbdElhalim, "Machine-learning based stacked ensemble model for accurate multi classification of CBC Anemia," 2023, doi: 10.58491/2735-4202.3144.

[11] Z. Faradila, A. Homaidi, and J. Prasetyo, "Classification of Anaemia Status Using The K-Nearest Neighbor Algorithm," *G-Tech*, vol. 9, no. 1, pp. 436–444, 2025, doi: 10.70609/gtech.v9i1.6377.

[12] T. R. Noviandy, G. M. Idroes, R. Suhendra, T. K. Bakri, and R. Idroes, "Advanced Anemia Classification Using Comprehensive Hematological Profiles and Explainable Machine Learning Approaches," *Infolitika Journal of Data Science*, vol. 2, no. 2, pp. 72–81, 2024, doi: 10.60084/ijds.v2i2.237.

[13] C. Gong, Y. Lin, J. Cao, and J. Wang, "Research on Enterprise Risk Decision Support System Optimization based on Ensemble Machine Learning," 2024, doi: 10.20944/preprints202410.0948.v1.

[14] G. Ghaly, N. T. Amanda, R. Ardhani, B. Sartono, and A. R. Firdawanti, "Analisis kinerja model stacking berbasis random forest dan svm dalam klasifikasi rumah tangga berdasarkan garis kemiskinan makanan di provinsi jawa barat," *Jurnal Lebesgue*, vol. 5, no. 3, pp. 2244–2265, 2024, doi: 10.46306/lb.v5i3.856.

[15] Y. Pamungkas, R. D. Indriani, and Z. B. Syulthoni, "Implementation of Synthetic Minority Over-Sampling Technique in the Anaemia Classification Using the LSTM and Bi-LSTM Algorithms," pp. 359–364, 2024, doi: 10.1109/eecsi63442.2024.10776106.

[16] C. Gong, Y. Lin, J. Cao, and J. Wang, "Research on Enterprise Risk Decision Support System Optimization based on Ensemble Machine Learning," 2024, doi: 10.20944/preprints202410.0948.v1.

[17] P. Gupta, N. K. Sehgal, and J. M. Acken, "Practical Aspects in Machine Learning," *Synthesis lectures on engineering, science, and technology*, pp. 281–330, 2024, doi: 10.1007/978-3-031-59170-9_9.

[18] K. H. Erian, P. H. Regalado, and J. M. Conrad, "Missing data handling for machine learning models," vol. 10, no. 2, pp. 123–132, 2021, doi: 10.11591/IJRA.V10I2.PP123-132.

[19] D. Ozsahin, M. T. Mustapha, A. S. Mubarak, Z. S. Ameen, and B. Uzun, "Impact of feature scaling on machine learning models for the diagnosis of diabetes," *2022 International Conference on Artificial Intelligence in Everything (AIE)*, pp. 87–94, 2022, doi: 10.1109/aie57029.2022.00024.

[20] V. R. Joseph and A. Vakayil, "SPlit: An Optimal Method for Data Splitting," *arXiv: Machine Learning*, 2020, doi: 10.1080/00401706.2021.1921037.

[21] A. N. Sadrabadi, S. M. Znjirchi, H. Z. Ahmadi, and A. Hajimoradi, "An optimized K-Nearest Neighbor algorithm based on Dynamic Distance approach," 2020, doi: 10.1109/ICSPIS51611.2020.9349582.
[22] H. Chen, "Understanding Gradient Boosting Classifier: Training, Prediction, and the Role of ——— b3_j," 2024, doi: 10.48550/arxiv.2410.05623.