

# Implementasi Modified K-Nearest Neighbor (MKNN) untuk Deteksi Penyakit Anemia

Putra Dwi Wira Gardha Yuniahans<sup>a,1,\*</sup>, Anggraini Puspita Sari<sup>a,2</sup>, Yisti Vita Via<sup>a,3</sup>

<sup>a</sup> Informatika, UPN "Veteran" Jawa Timur, Jl.Raya Rungkut Madya, Gunung Anyar, Surabaya

<sup>1</sup> putra22wir@gmail.com\*; <sup>2</sup>anggraini.puspita.if@upnjatim.ac.id; <sup>3</sup>yistivita.if@upnjatim.ac.id

\* Penulis Koresponden

No WA:+6289528618478

## INFO ARTIKEL

### Histori Artikel

Pengajuan 2024-06-27

Diperbaiki 2025-05-21

Diterima 2025-05-22

### Kata Kunci

Anemia

Diagnosis Dini

Klasifikasi

Machine Learning

MKNN

## ABSTRAK

Anemia merupakan sebuah kondisi di mana kadar hemoglobin dalam tubuh manusia berada di bawah kadar normal. Anemia dapat berdampak buruk bagi pasien yang menderita penyakit tersebut. Dampak yang dirasakan antara lain keterlambatan perkembangan psikomotor, peningkatan risiko infeksi penyakit, bahkan pada perempuan dapat menyebabkan kelahiran prematur. Oleh karena itu, diagnosis dini terhadap penyakit anemia sangat penting untuk mempercepat penanganan dan pengobatan. Salah satu pendekatan yang dapat digunakan untuk membantu proses diagnosis tersebut adalah metode *machine learning*, khususnya metode Modified K-Nearest Neighbor (MKNN). MKNN merupakan pengembangan dari metode KNN, dengan penambahan proses seperti perhitungan validitas dan weight voting yang tidak terdapat pada metode KNN standar. Pada penelitian ini, metode MKNN diimplementasikan untuk mendeteksi penyakit anemia dan menghasilkan akurasi terbaik sebesar 84% pada parameter pembagian data 75:25 dengan nilai  $k=5$ . Dataset yang digunakan berasal dari Rumah Sakit Jemursari Surabaya, dengan jumlah data sebanyak 100 data pasien yang telah dikumpulkan dan dianalisis untuk keperluan pengujian metode ini.

## ABSTRACT

### Keyword

Anemia

Classification

Early Diagnosis

Machine Learning

MKNN

*Anemia is a condition in which the level of hemoglobin in the human body falls below the normal threshold. Anemia can have adverse effects on patients suffering from the disease. These effects include delayed psychomotor development, increased risk of infectious diseases, and in women, it can even lead to premature birth. Therefore, early diagnosis of anemia is crucial to accelerate treatment and recovery. One approach that can assist in the diagnosis process is the use of machine learning methods, particularly the Modified K-Nearest Neighbor (MKNN) algorithm. MKNN is an enhanced version of the standard KNN method, with additional processes such as validity calculation and weighted voting that are not present in the original KNN. In this study, the MKNN method was implemented to detect anemia and achieved the best accuracy of 84% using a data split ratio of 75:25 with  $k=5$ . The dataset used in this research was obtained from Jemursari Hospital in Surabaya, consisting of 100 patient records that were collected and analyzed for method evaluation.*

Ini adalah artikel akses terbuka di bawah lisensi [CC-BY-SA](#).



## 1. Pendahuluan

Anemia merupakan sebuah kondisi dimana kadar hemoglobin pada tubuh manusia lebih rendah dari kadar normalnya. Kadar hemoglobin sendiri pada tubuh manusia mempunyai peranan yang sangat penting yakni membawa oksigen ke semua jaringan tubuh [12]. Jika kadar hemoglobin tersebut rendah dapat menyebabkan kelelahan, sesak nafas, pusing, dan penurunan kinerja fisik. Penyebab anemia meliputi kekurangan zat besi, defisiensi vitamin B12, kerusakan sel darah merah, atau masalah genetik [13].

World Health Organization (WHO) memperkirakan angka prevalensi penyakit anemia di seluruh dunia terdapat 40% pada bayi dengan usia 6 hingga 59 bulan, serta sebanyak 30% perempuan usia 15 hingga 49 tahun tersebut terkena atau mengalami penyakit anemia [6]. Selain itu Kementerian Kesehatan Indonesia (KEMENKES) dalam hasil Riset Kesehatan Dasar (RISKESDAS) pada tahun 2007, 2013, hingga 2018 angka prevalensi penyakit anemia sendiri terdapat peningkatan. Seperti pada anak usia 0-59 bulan di tahun 2007 sebanyak 27,7%, pada tahun 2013 sebanyak 28,1%, hingga pada tahun 2018 meningkat ke angka 38,5%. Hal tersebut juga terjadi pada usia 5-14 tahun dan pada usia 15 hingga 24 tahun [7]. Hal tersebut juga tentu memiliki dampak negatif seperti keterlambatan perkembangan psikomotor [4], meningkatkan risiko terkena infeksi penyakit [3], penurunan prestasi anak disekolah hingga pada wanita dampak jangka panjang dari anemia dapat dikatakan jauh lebih berbahaya karena dapat meningkatkan risiko kelahiran prematur dan memiliki potensi berat badan bayi rendah [5].

*Machine learning* adalah cabang dari kecerdasan buatan yang memungkinkan sistem untuk belajar dan membuat keputusan berdasarkan data tanpa pemrograman eksplisit. Teknik ini melibatkan penggunaan algoritma untuk menganalisis data, mengenali pola, dan membuat prediksi atau keputusan secara otomatis. Seperti pada penelitian [2] yang memperkenalkan model peramalan angin berbasis pembelajaran mendalam, yaitu 3DCNN-DConvLSTM, untuk meramalkan kecepatan dan arah angin secara bersamaan. Model ini menggabungkan jaringan saraf konvolusional tiga dimensi dan memori jangka panjang konvolusional mendalam, yang efektif dalam mengurangi waktu pelatihan dan kesalahan peramalan. Akurasi model dievaluasi dengan data dari AMEDAS, Jepang, dan dibandingkan dengan model peramalan lainnya. Contoh lain pada penelitian [8], yang meneliti mengenai prediksi hujan karena informasi dari prediksi hujan sering tidak akurat sehingga diperlukan bantuan *machine learning* untuk mengatasi permasalahan tersebut. Penelitian ini menunjukkan bahwa metode Neural Network dapat menghasilkan prediksi curah hujan dengan error 0.828, membuktikan akurasi yang baik.

Penelitian terbaru yang dilakukan oleh Jatoi et al. (2018) menunjukkan bahwa laporan hitung darah lengkap dapat dimanfaatkan secara efektif untuk mendeteksi penyakit melalui pendekatan data mining [11]. Dalam penelitian tersebut, digunakan metode klasifikasi Naive Bayes (NB) untuk mengklasifikasikan berbagai jenis penyakit berdasarkan parameter-parameter yang terdapat pada laporan hitung darah lengkap. Hasil yang diperoleh menunjukkan bahwa metode NB mampu memberikan tingkat akurasi yang sangat tinggi, yakni mencapai 98%. Temuan ini membuktikan bahwa algoritma NB memiliki potensi besar dalam analisis data medis, khususnya dalam membantu proses diagnosis penyakit secara cepat dan akurat berbasis data laboratorium.

Berdasarkan penjelasan-penjelasan sebelumnya maka penelitian ini dilakukan dengan judul Implementasi Metode Modified K-Nearest Neighbor (MKNN) Untuk Deteksi Penyakit Anemia. Metode MKNN sendiri merupakan metode modifikasi dari metode K-Nearest Neighbor (KNN). Tujuan dari metode MKNN sendiri yaitu untuk meningkatkan akurasi dari metode KNN. Penelitian sebelumnya yang menggunakan metode MKNN sendiri terdapat penelitian [1] dengan penelitian klasifikasi tingkat laju covid-19 untuk mitigasi penyebaran menggunakan metode

modified k-nearest neighbor (MKNN). Pada penelitian tersebut menggunakan data perkembangan kasus Covid19 dari KEMENKES dengan jumlah 510 data. Data tersebut akan dibagi menjadi dua bagian yakni 70% untuk data latih dan 30% untuk data uji. Hasil dari penelitian tersebut menghasilkan nilai akurasi sebesar 97% dengan parameter nilai  $k=3$  dan  $k=5$ .

Berdasarkan penelitian tersebut serta permasalahan yang ada pada penyakit anemia di Indonesia sendiri maka penelitian ini dilakukan untuk mendeteksi dini penyakit anemia. Tujuan untuk mendeteksi dini penyakit anemia yaitu agar penanganan pasien yang terdiagnosa penyakit anemia dapat ditangani secara cepat dan tepat, serta diharapkan penelitian ini juga dapat membantu praktisi kesehatan dalam pengambilan keputusan yang tepat dan cepat.

## 2. Metode

Pada penelitian ini terdapat beberapa langkah yang dilakukan sehingga penelitian ini dapat dilakukan hingga selesai. langkah-langkah tersebut seperti studi literatur, pengumpulan dataset, pengolahan data, implementasi metode, serta pengujian metode. Langkah-langkah dalam penelitian ini ditunjukkan pada Gambar 1.



**Gambar 1.** Diagram Alir Metode Penelitian

Pada Gambar 1 di atas menggambarkan alur proses deteksi penyakit anemia menggunakan metode Modified K-Nearest Neighbor (MKNN). Diagram tersebut menunjukkan tahapan-tahapan yang dilakukan, mulai dari input data dan pra-pemrosesan hingga menghasilkan klasifikasi akhir dengan metode MKNN. Berikut penjelasan masing-masing proses dari alur tersebut.

### 1. Studi Literatur

Studi literatur dalam implementasi Metode Modified K-Nearest Neighbor (MKNN) untuk deteksi penyakit anemia memberikan pemahaman yang mendalam tentang penyakit tersebut dan aplikasi metode MKNN dalam bidang kesehatan. Dengan mengumpulkan informasi dari sumber-sumber terpercaya seperti jurnal ilmiah dan artikel penelitian, studi literatur membantu merumuskan pendekatan penelitian yang tepat. Analisis literatur membantu mengidentifikasi kelemahan dan kelebihan metode MKNN dalam konteks deteksi penyakit anemia, serta memberikan landasan yang kuat untuk penelitian lebih lanjut dalam upaya meningkatkan diagnosis dan penanganan penyakit anemia.

### 2. Pengumpulan Dataset

Data penelitian yang digunakan diperoleh dari Rumah Sakit Islam Jemursari Surabaya, dengan total 100 data pasien yang telah didiagnosis mengalami penyakit anemia dan tidak terdiagnosa penyakit anemia. Proses pengumpulan data dilakukan dengan mengakses rekam medis pasien yang telah tercatat dalam sistem rumah sakit. Setiap data pasien mencakup informasi tentang kadar hemoglobin, jumlah sel darah merah, dan parameter klinis lainnya yang relevan untuk analisis klasifikasi penyakit anemia. Keterlibatan rumah sakit memastikan validitas dan keakuratan data yang digunakan dalam penelitian ini, serta memberikan landasan yang kuat untuk analisis dan kesimpulan yang akurat.

### 3. Pengolahan Data

Pengolahan data merupakan tahapan penting dalam proses pembuatan model dalam penelitian ini. Tahapan ini terdiri dari proses pembersihan data duplikat, penanganan missing value, serta proses mengubah data kategori menjadi data numerik. Pembersihan data duplikat merupakan proses pencarian data yang sama persis dalam suatu himpunan data, dan data tersebut akan dihapus salah satunya. Proses ini sangat penting karena keberadaan data duplikat dapat mengganggu tingkat akurasi model yang dibangun [14]. Selain itu, penanganan missing value juga diperlukan dengan cara mengidentifikasi nilai-nilai yang hilang pada dataset. Data yang tidak memiliki nilai tersebut dapat dihapus dari himpunan data atau, jika memungkinkan, diganti dengan nilai lain seperti nilai rata-rata (mean), median, atau modus, tergantung pada konteks dan jenis datanya [15].

Tahapan pengolahan data selanjutnya adalah mengubah data kategori menjadi data numerik. Hal ini penting dilakukan apabila dalam dataset terdapat variabel kategorik yang tidak dapat langsung diproses oleh algoritma *machine learning*. Proses ini dilakukan dengan teknik pengkodean seperti label encoding atau one-hot encoding. Sebagai contoh, variabel “jenis kelamin” yang memiliki kategori “laki-laki” dan “perempuan” dapat diubah menjadi nilai numerik seperti 0 dan 1, di mana 0 mewakili “laki-laki” dan 1 mewakili “perempuan”. Penanganan ini memungkinkan model untuk memahami informasi kategorikal sebagai bagian dari proses analisis data.

Dalam penelitian ini, proses evaluasi model dilakukan dengan menggunakan teknik validasi train-test split. Teknik ini membagi *dataset* menjadi dua bagian utama, yaitu data latih (training data) dan data uji (testing data) [9]. Penelitian ini menggunakan dua skenario pembagian data, yaitu 75:25 dan 85:15, di mana 75% atau 85% dari data digunakan untuk melatih model, sementara 25% atau 15% sisanya digunakan untuk menguji performa model. Penggunaan dua rasio pembagian ini bertujuan untuk membandingkan pengaruh proporsi data latih terhadap tingkat akurasi model. Teknik ini bersifat sederhana namun efektif dalam memberikan gambaran mengenai kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya. Dengan menerapkan pembagian train-test split secara sistematis, evaluasi kinerja model menjadi lebih objektif dan dapat dipercaya.

### 4. Implementasi Metode MKNN

Metode MKNN merupakan metode perkembangan atau modifikasi dari metode KNN. Langkah-langkah dari metode MKNN sendiri masih tidak jauh beda dari langkah-langkah metode KNN, dimana untuk memprediksi data baru pada metode MKNN juga masih melibatkan tetangga terdekat dari data tersebut. Perbedaan dari metode MKNN sendiri yaitu terdapat beberapa tambahan proses yang tidak dimiliki oleh metode KNN. Tambahan proses tersebut yaitu proses validasi data latih dan proses perhitungan pembobotan. Tujuan dari metode MKNN sendiri yaitu bertujuan untuk meningkatkan akurasi dari metode KNN. Pada Gambar 2 merupakan diagram alur proses dari implementasi metode MKNN.



**Gambar 2.** Diagram Alir Implementasi Metode MKNN

- Menentukan nilai k untuk proses perhitungan validitas pada data latih. Nilai k sendiri biasanya bernilai ganjil seperti 3, 5, 7 dan seterusnya.
- Menghitung jarak antara data latih, perhitungan jarak sendiri pada penelitian ini menggunakan persamaan euclidean distance. Persamaan tersebut ditunjukkan pada Persamaan 1.

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

dimana,

d = jarak antara koordinat a dan b

a, b = vektor a dan vektor b

i = setiap data

n = jumlah keseluruhan data

bi, bi = vektor a ke i dan vektor b ke i

- Menghitung nilai validitas pada data latih sesuai nilai k yang telah ditentukan sebelumnya. Perhitungan validitas sendiri ditunjukkan pada Persamaan 2.

$$validitas(x) = \frac{1}{k} \sum_{i=1}^k S(lbl(x), lbl(n_i(x))) \quad (2)$$

dimana,

k = jumlah titik terdekat

lbl(x) = kelas x

ni(x) = label kelas titik terdekat x

Fungsi S dalam persamaan 2.4 merupakan fungsi untuk mencocokkan antara titik x dan titik ke i sesuai dengan nilai k terdekat. Untuk mendefinisikan fungsi S tersebut maka didefinisikan dalam Persamaan 3.

$$S(x, i) = \left\{ \begin{array}{l} 1 \text{ jika } x \\ = i \\ 0 \text{ jika } x \neq \\ i \end{array} \right\} \quad (3)$$

- d. Menentukan nilai k untuk proses perhitungan pembobotan atau weight voting. Nilai k sendiri biasa bernilai ganjil.
- e. Menghitung jarak antara data uji dan data latih. Perhitungan jarak menggunakan persamaan euclidean distance yang ditunjukkan pada Persamaan 1.
- f. Melakukan perhitungan pembobotan atau weight voting sesuai nilai k yang telah ditentukan sebelumnya. Perhitungan weight voting sendiri melibatkan hasil dari perhitungan jarak antara data uji dan data latih pada tahapan yang ke 5 sebelumnya serta pada hasil perhitungan validitas yang terdapat pada tahapan ke 3. Perhitungan weight voting ditunjukkan pada Persamaan 4.

$$w(i) = Validitas(i) \times \frac{1}{d_e + 0,5} \quad (4)$$

dimana,

$w(i)$  = perhitungan weight voting

Validitas(i) = nilai validitas

$d_e$  = jarak euclidean

- g. Mengurutkan hasil dari perhitungan weight voting dari yang terbesar hingga yang terkecil. Sehingga diambil nilai weight voting sebanyak k yang telah ditentukan pada tahapan ke 4, dan nilai weight voting tersebut dijumlahkan sesuai kelasnya. Sehingga untuk memprediksi kelas pada data baru yaitu membandingkan apakah kelas anemia lebih besar dibanding kelas non anemia. Jika iya maka hasil prediksinya adalah kelas anemia, begitupun sebaliknya.

## 5. Pengujian

Setelah implementasi dari metode MKNN sendiri, maka selanjutnya metode tersebut akan dilakukan pengujian. Pengujian untuk metode MKNN sendiri bertujuan untuk mengetahui sejauh mana metode mampu mengklasifikasikan data dengan benar dan mengidentifikasi keterbatasan atau area di mana model perlu diperbaiki [10]. Skenario pengujian dalam penelitian ini yaitu menggunakan parameter split data 75:25 dan 85:15 serta pada nilai k sendiri yang akan digunakan dalam pengujian ini yaitu  $k=5$ , dan  $k=7$ .

Lebih lanjut, untuk mengetahui efektivitas dari metode MKNN, dilakukan pula perbandingan hasil pengujian dengan metode K-Nearest Neighbor (KNN) menggunakan parameter yang sama dengan hasil akurasi terbaik pada metode MKNN. Perbandingan ini bertujuan untuk mengetahui sejauh mana modifikasi yang dilakukan pada metode MKNN mampu meningkatkan akurasi dibandingkan metode KNN konvensional. Dari hasil pengujian ini diharapkan dapat diperoleh informasi mengenai metode mana yang

memberikan performa terbaik dalam mendeteksi penyakit anemia berdasarkan data yang digunakan.

### 3. Hasil dan Analisis

Hasil dari penelitian implementasi metode MKNN untuk mendeteksi penyakit anemia sendiri terdapat pengujian yang dilakukan. Pengujian tersebut dilakukan dengan menggunakan parameter split data dengan rasio 75:25 dan 85:15 serta dengan menggunakan parameter nilai k dengan nilai 5, dan 7. Berdasarkan skenario yang telah ditentukan tersebut diharapkan agar penelitian ini dapat memberikan pemahaman lebih mengenai keefektifan dari metode MKNN dalam mendeteksi penyakit anemia.

#### 1. Hasil Pengujian 75:25

Pada pengujian dengan rasio 75:25 parameter nilai k yang akan digunakan yaitu nilai k=5, dan k=7. Split data dengan rasio 75:25 mengartikan bahwa 75% dataset akan digunakan untuk data latih, dan 25% dari dataset akan digunakan sebagai data uji. 75% dataset yang akan digunakan sebagai data latih pada penelitian ini yaitu berjumlahkan 75 data dan 25% data yang akan digunakan sebagai data uji berjumlahkan 25 data. Pada pengujian ini dengan parameter k=5 mendapatkan nilai akurasi sebesar 84% yang ditunjukkan pada Gambar 3 serta pada pengujian dengan menggunakan parameter nilai k=7 mendapatkan nilai akurasi sebesar 80% yang ditunjukkan pada Gambar 4.

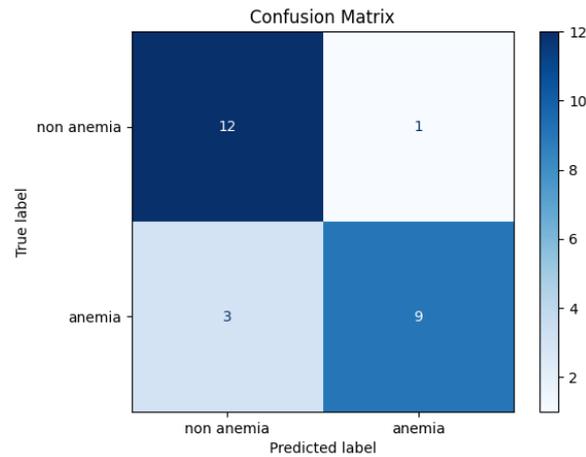
	precision	recall	f1-score	support
non anemia	0.80	0.92	0.86	13
anemia	0.90	0.75	0.82	12
accuracy			0.84	25
macro avg	0.85	0.84	0.84	25
weighted avg	0.85	0.84	0.84	25

Gambar 3. Akurasi Pengujian Split Data 75:25 k=5

	precision	recall	f1-score	support
non anemia	0.75	0.92	0.83	13
anemia	0.89	0.67	0.76	12
accuracy			0.80	25
macro avg	0.82	0.79	0.79	25
weighted avg	0.82	0.80	0.80	25

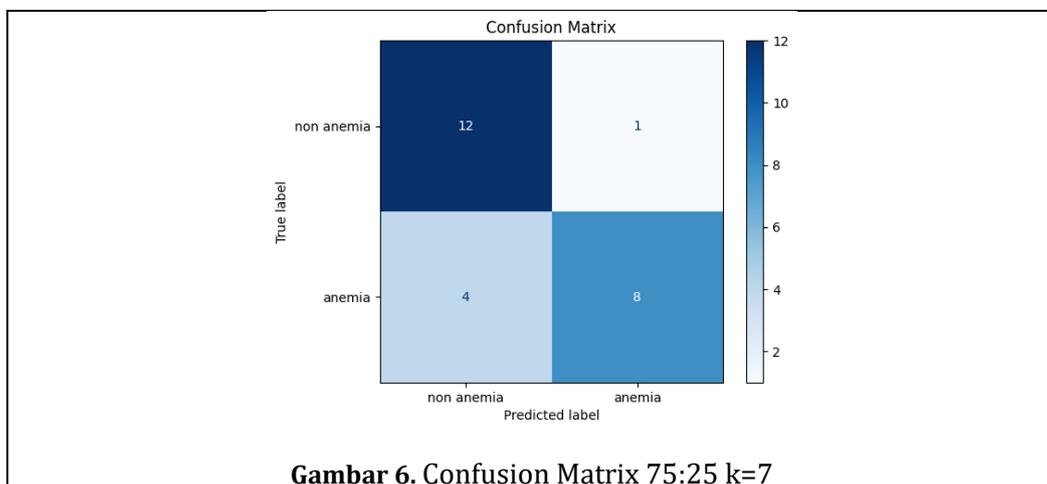
Gambar 4. Akurasi Pengujian Split Data 75:25 k=7

Pada pengujian dengan parameter nilai k=5 tersebut 84% dari total data uji berhasil diprediksi benar oleh metode MKNN. 84% tersebut berarti berjumlahkan 21 data dari 25 total data uji, data yang diprediksi benar tersebut antara lain 12 data dengan kelas non anemia dan 9 data dengan kelas anemia. Selain itu data yang diprediksi salah oleh metode MKNN berjumlahkan 4 data, data-data tersebut antara lain 1 data dengan kelas sebenarnya adalah kelas non anemia tetapi oleh metode MKNN sendiri memprediksinya sebagai kelas anemia, dan 3 data dengan kelas sebenarnya adalah kelas anemia tetapi oleh metode MKNN sendiri memprediksikannya sebagai kelas anemia. Visualisasi data yang diprediksi benar dan yang diprediksi salah tersebut ditunjukkan pada Gambar 5.



**Gambar 5.** Confusion Matrix 75:25 k=5

Selanjutnya nilai akurasi yang didapatkan pada pengujian dengan parameter nilai k=7 tersebut yakni 80% dimana dapat diartikan bahwa 80% dari total data uji dapat diprediksi dengan benar. Total data yang diprediksi benar tersebut yakni 20 data dari 25 data total keseluruhan. 20 data tersebut diantaranya 12 data pada kelas non anemia dan 8 data pada kelas anemia. Sisanya 5 data yang diprediksi salah oleh metode MKNN sendiri antara lain 1 data dengan kelas sebenarnya non anemia tetapi metode MKNN sendiri memprediksinya sebagai kelas anemia, dan 4 data dengan kelas sebenarnya adalah kelas anemia tetapi oleh metode MKNN sendiri memprediksinya sebagai kelas non anemia. Visualisasi dari pengujian dengan parameter nilai k=7 tersebut ditunjukkan pada Gambar 6.



**Gambar 6.** Confusion Matrix 75:25 k=7

## 2. Hasil Pengujian 85:15

Pada pengujian dengan rasio 85:15 parameter nilai k yang akan digunakan yaitu nilai k=5, dan k=7. Split data dengan rasio 85:15 mengartikan bahwa 85% dataset akan digunakan untuk data latih, dan 15% dari dataset akan digunakan sebagai data uji. 85% dataset yang akan digunakan sebagai data latih pada penelitian ini yaitu berjumlahkan 85 data dan 15% data yang akan digunakan sebagai data uji berjumlahkan 15 data. Pada pengujian ini dengan menggunakan parameter nilai k=5 mendapatkan nilai akurasi sebesar 80% yang

ditunjukkan pada Gambar 67 serta pada pengujian dengan parameter nilai  $k=7$  mendapatkan nilai akurasi sebesar 73% yang ditunjukkan pada Gambar 8.

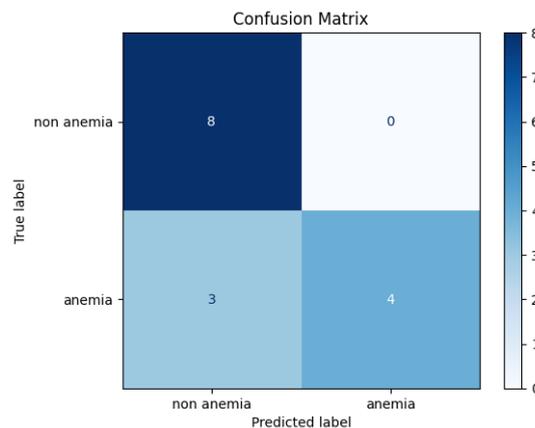
	precision	recall	f1-score	support
non anemia	0.73	1.00	0.84	8
anemia	1.00	0.57	0.73	7
accuracy			0.80	15
macro avg	0.86	0.79	0.78	15
weighted avg	0.85	0.80	0.79	15

Gambar 7. Akurasi Pengujian Split Data 85:15  $k=5$

	precision	recall	f1-score	support
non anemia	0.67	1.00	0.80	8
anemia	1.00	0.43	0.60	7
accuracy			0.73	15
macro avg	0.83	0.71	0.70	15
weighted avg	0.82	0.73	0.71	15

Gambar 8. Akurasi Pengujian Split Data 85:15  $k=7$

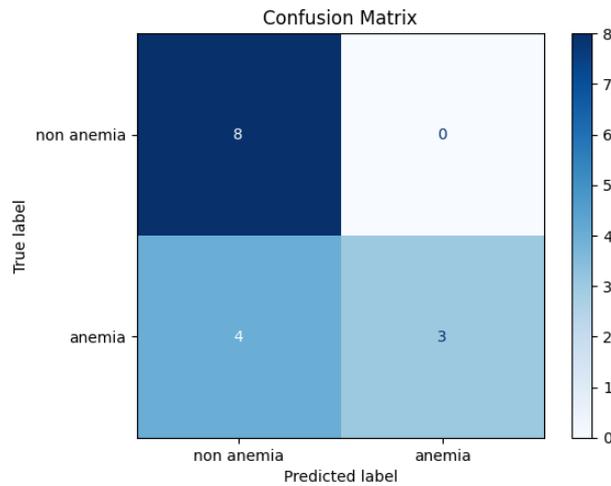
Pada pengujian dengan parameter nilai  $k=5$  tersebut menghasilkan nilai akurasi 80%, dimana hal tersebut dapat disimpulkan 80% dari total data uji yang digunakan dapat diprediksi benar oleh metode MKNN. Total data yang diprediksi benar tersebut yakni 12 data dari 15 total data uji. 12 data tersebut diantaranya 8 data pada kelas non anemia dan 4 data pada kelas anemia. Sisa data yang diprediksi salah yakni 3 data pada kelas anemia tetapi oleh metode MKNN memprediksinya sebagai kelas non anemia. Visualisasi dari data yang diprediksi benar dan salah tersebut ditunjukkan pada Gambar 9.



Gambar 9. Confusion Matrix 85:15  $k=5$

Selanjutnya pada pengujian dengan parameter nilai  $k=7$  mendapatkan nilai akurasi sebesar 73% dimana hal tersebut dapat disimpulkan bahwa 73% dari total data uji yang digunakan dapat diprediksi benar oleh metode MKNN. Total dari jumlah data yang diprediksi benar tersebut berjumlah 11 data dari total data uji yakni 15 data. 11 data tersebut diantaranya 8 data pada kelas non anemia dan 3 data pada kelas anemia. Selain itu pada data yang salah diprediksi berjumlah 4 data pada kelas anemia yang

diprediksi oleh metode MKNN sebagai kelas non anemia. Visualisasi dari data yang diprediksi benar dan salah tersebut ditunjukkan pada Gambar 10.



**Gambar 10.** Confusion Matrix 85:15 k=7

### 3. Analisis Pengujian

Berdasarkan hasil pengujian yang telah dilakukan sebelumnya maka hasil pengujian tersebut akan disatukan dan akan diambil nilai akurasi yang paling baik. Kumpulan-kumpulan nilai akurasi dari pengujian sebelumnya ditunjukkan pada Tabel 1. Kumpulan nilai akurasi dari pengujian sebelumnya didapatkan nilai akurasi yang paling baik didapatkan pada parameter split data 75:25 dengan nilai k=5. Nilai akurasi yang didapatkan pada pengujian dengan parameter tersebut yakni nilai akurasi sebesar 84%.

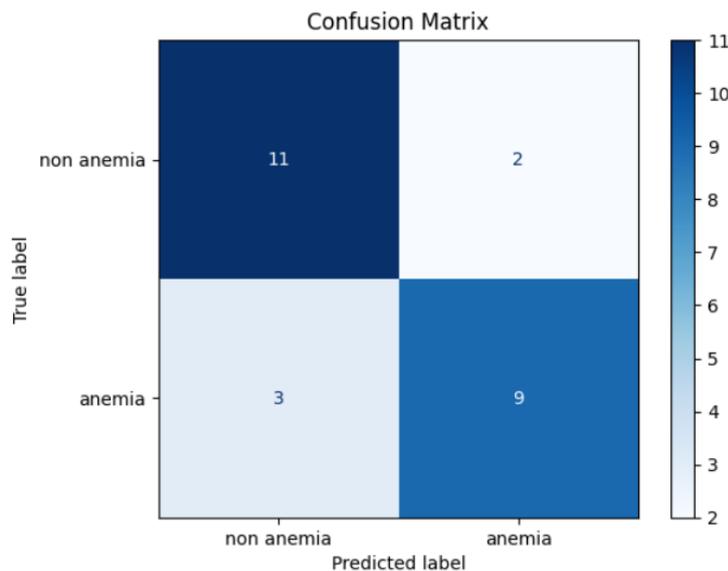
Nilai k	Split Data 75:25	Split Data 85:15
5	<b>84%</b>	80%
7	80%	73%

Pada Tabel 1 diatas dapat disimpulkan parameter yang menghasilkan nilai akurasi paling baik menggunakan metode MKNN adalah pada parameter dengan split data 75:25 dengan nilai k=5, maka dari itu dengan parameter yang sama akan dilakukan pengujian menggunakan metode KNN untuk mengevaluasi metode MKNNnya. hasil dari pengujian menggunakan metode KNN dengan parameter yang sama menghasilkan nilai akurasi sebesar 80% yang ditunjukkan pada Gambar 11. Dimana terdapat 2 data dengan kelas non-anemia dan 3 data dengan kelas anemia diprediksi salah yang ditunjukkan pada Gambar 12.

Berdasarkan analisa tersebut perbandingan antara metode MKNN dan metode KNN dengan parameter yang sama, nilai akurasi yang didapatkan lebih baik pada metode MKNN daripada metode KNN sendiri. Maka dari itu metode MKNN berhasil meningkatkan nilai akurasi dari metode KNN sendiri.

	precision	recall	f1-score	support
non anemia	0.79	0.85	0.81	13
anemia	0.82	0.75	0.78	12
accuracy			0.80	25
macro avg	0.80	0.80	0.80	25
weighted avg	0.80	0.80	0.80	25

Gambar 11. Akurasi Metode KNN 75:25 dengan k=5



Gambar 12. Confusion Matrix KNN 75:25 k=5

#### 4. Simpulan

Pengujian metode MKNN untuk mendeteksi penyakit anemia dilakukan melalui beberapa percobaan dengan menggunakan kombinasi parameter yang berbeda. Tujuan dari skenario percobaan ini adalah untuk menemukan konfigurasi parameter terbaik yang menghasilkan akurasi klasifikasi tertinggi. Berdasarkan hasil pengujian, diperoleh bahwa pada parameter split data 75:25 dan nilai  $k = 5$ , metode MKNN memberikan performa terbaik dengan nilai akurasi sebesar 84%, di mana dari total 25 data uji, sebanyak 21 data berhasil diprediksi dengan benar oleh model.

Sebagai bentuk pengembangan, penelitian selanjutnya disarankan untuk melakukan pengujian lebih lanjut dengan menggunakan variasi parameter lainnya, baik dari segi rasio pembagian data maupun nilai  $k$ , guna mengeksplorasi kemungkinan peningkatan performa. Selain itu, penambahan variabel-variabel penting lainnya dalam dataset juga dapat dilakukan untuk memperkaya informasi dan meningkatkan kualitas klasifikasi.

Adapun hasil perbandingan antara metode MKNN dan KNN pada parameter yang sama, yaitu split data 75:25 dan nilai  $k = 5$ , menunjukkan bahwa MKNN menghasilkan akurasi lebih tinggi sebesar 84%, dibandingkan dengan KNN yang hanya memperoleh akurasi sebesar 80%. Hal ini membuktikan bahwa modifikasi yang dilakukan pada algoritma MKNN, seperti penambahan proses validitas dan pembobotan (weight voting), mampu meningkatkan akurasi klasifikasi dibandingkan metode KNN standar. Dengan demikian, MKNN dapat dianggap sebagai alternatif

yang lebih baik dalam mendeteksi penyakit anemia pada dataset yang digunakan dalam penelitian ini.

## Referensi

- [1] Cholissodin, I., Marvela Evanita, F., Tedjasulaksana, J. J., Wahyuditomo, K. W., & Korespondensi, P. (2021). Klasifikasi tingkat laju data covid-19 untuk mitigasi penyebaran menggunakan metode modified k-nearest neighbor (mknn). 8(3). <https://doi.org/10.25126/jtiik.2021834400>
- [2] Sari, A. P., Suzuki, H., Kitajima, T., Yasuno, T., Prasetya, D. A., & Arifuddin, R. (2022). Short-Term Wind Speed and Direction Forecasting by 3DCNN and Deep Convolutional LSTM. *IEEJ Transactions on Electrical and Electronic Engineering*, 17(11), 1620-1628.
- [3] Spezia, J., Carvalho, L. F. da S., Camargo-Filho, M. F. de A., Furman, A. E., Utiyama, S. R. da R., & Henneberg, R. (2018). Prevalence of anemia in schools of the metropolitan region of Curitiba, Brazil. *Hematology, Transfusion and Cell Therapy*, 40(2), 151-155. <https://doi.org/10.1016/j.htct.2017.11.007>
- [4] Gumilang, L., Nurlaelasari, D., Dhamayanti, M., Tina Dewi Judistiani, R., Martini, N., Yogi Pramartirta, A., & Anak Rumah Sakit Umum Pusat Hasan Sadikin Bandung Fakultas Kedokteran, K. (2021). Gambaran Faktor Risiko Kejadian Anemia Pada Balita. *Jurnal Kebidanan Malahayati*, 7(4), 681-687. <https://doi.org/10.33024>
- [5] Levy, T. S., de la Cruz Góngora, V., & Villalpando, S. (2015). Anemia: Causes and Prevalence. In *Encyclopedia of Food and Health* (pp. 156-163). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-384947-2.00029-5>
- [6] Who. (2023, May 12). WHO calls for accelerated action to reduce anaemia, <URL:<https://www.who.int/news/item/12-05-2023-who-calls-for-accelerated-action-to-reduce-anaemia>>
- [7] Kemenkes. (2022, November 16). Remaja Bebas Anemia: Konsentrasi Belajar Meningkatkan, Bebas Prestasi. <URL:<https://ayosehat.kemkes.go.id/remaja-bebas-anemia-konsentrasi-belajar-meningkat-bebas-prestasi>>
- [8] Hermawan, A., & Avianto, D. (2021). Prediksi Curah Hujan Wilayah Provinsi Yogyakarta dengan Algoritma Neural Network. *JASIEK (Jurnal Aplikasi Sains, Informasi, Elektronika dan Komputer)*, 3(1), 47-54.
- [9] Sivakumar, M., Parthasarathy, S., & Padmapriya, T. (2024). *Trade-off between training and testing ratio in machine learning for medical image processing*. *PeerJ Computer Science*, 10, e2245.
- [10] Rizky, J. L., & Putra, Z. P. Analisis Perbandingan Algoritma Pembelajaran Mesin untuk Meningkatkan Akurasi dan Klasifikasi Tumor Otak. *IJAI (Indonesian Journal of Applied Informatics)*, 9(1), 31-44.
- [11] Jatoi, S., Aamir Panhwar, M., Sulleman Memon, M., Ahmed Baloch, J., & Saddar, S. (2018). Mining Complete Blood Count Reports For Disease Discovery. In *IJCSNS International Journal of Computer Science and Network Security* (Vol. 18, Issue 1).
- [12] Ulva, T., Stikes, K., & Kendal, M. (2022). Peran Edukasi Gizi Dalam Pencegahan Anemia Pada Remaja Di Indonesia: Literature Review. 4(1).
- [13] Cappellini, M. D., Musallam, K. M., & Taher, A. T. (2020). Iron deficiency anaemia revisited. *Journal of internal medicine*, 287(2), 153-170.
- [14] Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural networks*, 106, 249-259.
- [15] Li, W., Zhong, K., Wang, J., & Chen, D. (2021). A dynamic algorithm based on cohesive entropy for influence maximization in social networks. *Expert Systems with Applications*, 169, 114207.



Putra Dwi Wira Gardha Yuniahans, Surabaya, 22 Juni 2002. Memperoleh gelar sarjana dari Universitas Pembangunan Nasional “Veteran” Jawa Timur pada 2024. Minat Penelitiannya adalah Mechine Learning dan Web Development.

Alamat Email: [putra22wir@gmail.com](mailto:putra22wir@gmail.com)



Anggraini Puspita Sari, memperoleh gelar Sarjana Teknik dari Universitas Brawijaya pada 2009, Magister Teknik dari Universitas Brawijaya pada 2012, Selain itu mendapatkan gelar Doktor Teknik (Dr. Eng.) dari Tokushima University pada tahun 2021. Saat ini, ia bekerja sebagai Asisten Profesor di bidang Informatika, Universitas Pembangunan Nasional Veteran Jawa Timur, Surabaya, Minat penelitian beliau saat ini meliputi peramalan, pembangkit listrik tenaga angin, kecerdasan buatan, mikroelektronika, dan teknik elektro.

Ia merupakan anggota Forum Pendidikan Tinggi Teknik Elektro Indonesia (FORTEI Indonesia), anggota Institute of Electrical and Electronics Engineers (IEEE), serta anggota Persatuan Insinyur Indonesia (PII).

Alamat Email: [anggraini.puspita.if@upnjatim.ac.id](mailto:anggraini.puspita.if@upnjatim.ac.id)



Yisti Vita Via lahir di Lamongan, 25 April 1986, ia memperoleh gelar Sarjana Teknologi Informasi dari Politeknik Elektronika Negeri Surabaya (PENS), Indonesia pada 2008. Selanjutnya, ia meraih gelar Magister Teknik Informatika dari Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Indonesia pada 2012. Saat ini, ia merupakan dosen di Departemen Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur, Surabaya, Indonesia. Minat penelitiannya meliputi pengembangan sistem cerdas dengan menerapkan algoritma yang terinspirasi dari alam dan algoritma optimasi.

Alamat Email: [yistivia.if@upnjatim.ac.id](mailto:yistivia.if@upnjatim.ac.id)