



Analisis Komparatif Kinerja Algoritma Machine Learning untuk Deteksi Stunting

Syahrani Lonang¹, Anton Yudhana^{2,*}, Muhammad Kunta Biddinika¹

¹Program Magister Informatika, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

²Program Studi Teknik Elektro, Universitas Ahmad Dahlan, Yogyakarta, Indonesia

Email: ¹lonangsyahrani3@gmail.com, ^{2,*}eyudhana@ee.uad.ac.id, ³muhammad.kunta@mti.uad.ac.id

Email Penulis Korespondensi: eyudhana@ee.uad.ac.id

Abstrak—Stunting merupakan masalah serius yang diakibatkan oleh kekurangan gizi kronis pada balita, menyebabkan pertumbuhan yang terhambat dan berdampak buruk pada kesehatan serta produktivitas jangka panjang. Oleh karena itu, deteksi dini stunting sangat penting untuk mengurangi dampak negatifnya. Studi-studi terdahulu yang memanfaatkan machine learning telah membuktikan keberhasilan metode ini dalam berbagai aplikasi kesehatan, seperti deteksi penyakit dan prediksi kondisi medis. Hasil penelitian ini merinci evaluasi komparatif lima algoritma klasifikasi, yaitu Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors (KNN), dan Support Vector Machine (SVM), dalam mengklasifikasi balita stunting. Dataset yang digunakan mengandung empat atribut penting: usia, jenis kelamin, berat badan, dan tinggi badan balita, serta label kelas biner yang membedakan antara balita yang mengalami stunting dan yang tidak. Hasil evaluasi menunjukkan bahwa KNN pada $K = 3$ menghasilkan akurasi tertinggi sebesar 94,85%, menjadikannya model terbaik dalam mengklasifikasikan stunting pada balita. Selain akurasi, metrik lain seperti presisi, recall, dan F1-score digunakan untuk menganalisis kemampuan algoritma dalam mengatasi masalah ini. KNN menonjol sebagai model terbaik dengan F1-score tertinggi sebesar 89,47%. KNN juga berhasil mempertahankan keseimbangan antara presisi dan recall, menjadikannya pilihan yang sangat baik dalam penanganan stunting pada balita. Selain itu, penggunaan metrik AUC dari kurva ROC juga menunjukkan keunggulan KNN dalam membedakan antara balita stunting dan non-stunting. Dengan kombinasi hasil evaluasi yang konsisten, baik dari segi akurasi maupun metrik evaluasi lainnya, penelitian ini membuktikan bahwa KNN adalah pilihan terbaik dalam mengatasi tugas klasifikasi stunting pada balita.

Kata Kunci: Machine Learning; Decision Tree; Random Forest; K-Nearest Neighbor; Logistic Regression; Stunting;

Abstract—Stunting is a serious problem caused by chronic malnutrition in children under five, causing stunted growth and having a negative impact on long-term health and productivity. Therefore, early detection of stunting is very important to reduce its negative impacts. Previous studies utilizing machine learning have proven the success of this method in various health applications, such as disease detection and the prediction of medical conditions. The results of this research are a comparative evaluation of five classifications, namely Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM), in classifying stunted toddlers. The dataset used contains four important attributes: age, gender, weight, and height of toddlers, as well as a binary class label that differentiates between toddlers who are stunted and those who are not. The evaluation results show that KNN at $K = 3$ produces the highest accuracy of 94.85%, making it the best model for classifying stunting in toddlers. Apart from accuracy, other metrics such as precision, recall, and F1-score are used to analyze the algorithm's ability to solve this problem. KNN stands out as the best model, with the highest F1-score of 89.47%. KNN also manages to maintain a balance between precision and recall, making it an excellent choice for treating stunting in toddlers. Apart from that, the use of the AUC metric from the ROC curve also shows the superiority of KNN in differentiating between stunted and non-stunting toddlers. With a combination of consistent evaluation results, both in terms of accuracy and other evaluation metrics, this research proves that KNN is the best choice for overcoming the task of classifying stunting in toddlers.

Keywords: Machine Learning; Decision Tree; Random Forest; K-Nearest Neighbor; Logistic Regression; Stunting;

1. PENDAHULUAN

Stunting merupakan suatu kondisi dimana balita gagal tumbuh akibat kekurangan gizi kronis, sehingga mengakibatkan balita menjadi sangat pendek untuk usianya [1]. Dampak stunting sangat parah dan beragam, mencakup dampak buruk jangka pendek dan jangka panjang. Dalam jangka pendek, anak-anak yang mengalami stunting mungkin mengalami melemahnya sistem kekebalan tubuh, peningkatan kerentanan terhadap penyakit, dan perkembangan kognitif yang buruk. Dalam jangka panjang, stunting dapat menyebabkan penurunan produktivitas ekonomi dan mengganggu siklus kemiskinan [2], [3]. Karena dampaknya yang parah dan beragam, penting untuk mendeteksi stunting secara dini. Machine learning menawarkan pendekatan inovatif untuk deteksi dini stunting, memungkinkan penggunaan data berbagai atribut balita untuk memprediksi kondisi stunting. Hal ini dibutuhkan untuk mengidentifikasi balita yang berisiko mengalami stunting sehingga tindakan preventif dan intervensi yang sesuai dapat dilakukan. Machine learning dapat meningkatkan akurasi dalam deteksi stunting, membantu praktisi kesehatan dan pembuat kebijakan untuk menyalurkan intervensi lebih efektif, dan pada akhirnya, mengurangi prevalensi stunting serta menghindari dampak negatif jangka panjangnya. Oleh karena itu, penggunaan machine learning dalam mengatasi masalah stunting sangat relevan dan mendesak.

Machine learning sebagai bidang kecerdasan buatan yang sedang berkembang, menawarkan pendekatan inovatif untuk mendeteksi stunting [4]. Salah satunya adalah dengan menggunakan klasifikasi. Keuntungan machine learning dalam deteksi stunting terletak pada potensinya untuk meningkatkan akurasi, efisiensi, dan kecepatan dalam mendeteksi. Penerapan algoritma pembelajaran mesin dapat membantu praktisi kesehatan dan pembuat



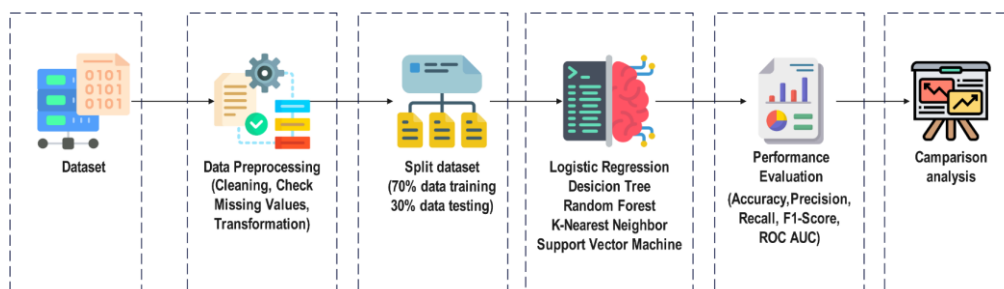
kebijakan menargetkan intervensi secara lebih efektif, sehingga pada akhirnya mengurangi prevalensi dan dampak stunting [5]. Dalam hal ini, machine learning telah diterapkan di sektor kesehatan seperti metode klasifikasi untuk tujuan deteksi penyakit [6]. Klasifikasi digunakan untuk tujuan prediksi diabetes [7], diagnosis covid 19 [8] dan mengklasifikasi balita kekurangan gizi [9].

Pada penelitian yang dilakukan oleh [10] meneliti tentang identifikasi penyakit dimensia menggunakan teknik machine learning. Ada tiga algoritma yang digunakan diantaranya SVM, KNN dan Linear Discriminant Analysis (LDA). Berdasarkan hasil pengujian KNN dan LDA mendapatkan akurasi yang seimbang sebesar 90%, sedangkan SVM hanya 76%. [11] menggunakan teknik machine learning untuk memprediksi balita stunting di Rwanda. Ada lima algoritma yang digunakan yaitu logistic regression, random forest, naïve bayes, extreme gradient boosting dan gradient boosting. Data yang digunakan adalah Rwanda Demographic and Health Survey 2019-2020 dengan basis model menggunakan stratified 10-fold cross validation. Performa terbaik dihasilkan oleh gradient boosting dengan akurasi sebesar 79.33%, precision 72.51% recall 94.49% dan f1-score 82.05%. [12] mengadopsi pendekatan klasifikasi data mining untuk prediksi diabetes. KNN dan SVM adalah algoritma yang digunakan, Dataset penelitian ini diambil dari UCI repository dengan sembilan atribut. Hasilnya setelah dibandingkan KNN menghasilkan akurasi lebih tinggi dibandingkan SVM. Akurasi sebesar 83% dan 80% pada KNN, SVM. Penelitian [13] mengklasifikasi kualitas sungai menggunakan bermacam-macam distances function pada algoritma KNN yang akan di bandingkan dengan algoritma DT dan SVM. Euclidean, Hamming, dan Entropy estimator adalah distance yang digunakan. Hasil penelitian ini menunjukkan KNN dengan entropy estimator mendapatkan akurasi tertinggi sebesar 99,90%. Lalu disusul oleh DT dengan 97,53% dan KNN-Hamming distance 90,12%. Penelitian ini mengisi celah pengetahuan dari penelitian sebelumnya yang telah menggunakan machine learning untuk deteksi stunting, terutama berfokus pada performa model. Meskipun penelitian [11] telah memanfaatkan berbagai algoritma, hasil performa model yang dibangun belum mencapai tingkat yang memuaskan. Oleh karena itu, penelitian ini bertujuan untuk mengisi kesenjangan dalam literatur dengan melakukan analisis komparatif lebih mendalam terhadap algoritma machine learning, khususnya (Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors, dan Support Vector Machine), dalam konteks deteksi stunting pada balita.

Paper ini menyajikan evaluasi komparatif lima algoritma pembelajaran mesin: Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), dan Support Vector Machine (SVM), untuk mengklasifikasi balita stunting. Dataset yang digunakan dalam penelitian ini diperoleh dari Elektronik Pencatatan dan Pelaporan Gizi Balita Berbasis Masyarakat (EPPGBM) Puskesmas Ubung, Lombok Tengah, Indonesia. Ini berisi informasi berharga, termasuk usia, jenis kelamin, berat badan, dan tinggi badan balita, serta label kelas biner: 0 untuk non-stunting dan 1 untuk stunting. Dimasukkannya keempat variabel prediktor ini didasarkan pada literatur yang menyoroti signifikansi keempat variabel prediktor ini dalam deteksi stunting. Penelitian ini bertujuan untuk mengidentifikasi algoritma yang bekerja paling optimal dalam mengklasifikasikan stunting pada balita dengan melihat performa pengujian setiap algoritma. Confusion matrix dimanfaatkan untuk melakukan evaluasi performa dengan metrik evaluasi seperti akurasi, presisi, recall, f1-score, dan AUC ROC dapat membantu menganalisis secara detail sehingga membantu dalam deteksi dan intervensi dini. Berdasarkan hasil algoritma yang paling optimal untuk deteksi stunting diharapkan dapat memberikan wawasan yang lebih baik untuk meningkatkan kualitas deteksi stunting menggunakan pendekatan machine learning.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan dataset yang didapatkan dari EPPGBM, dimana data tersebut akan melalui proses preprocessing untuk agar kompatibel untuk digunakan, lalu dataset akan dibagi, selanjutnya model klasifikasi akan dibangun menggunakan lima algoritma yang ditetapkan, performa setiap algoritma akan dievaluasi dan terakhir melakukan komparasi. Gambar 1 memperlihatkan langkah-langkah penelitian yang dilakukan.



Gambar 1. Tahapan Penelitian

Tahap pertama adalah preprocessing data, yang melibatkan pembersihan dataset dari data yang tidak valid, pengecekan dan penanganan missing values, dan transformasi data agar sesuai dengan format yang diperlukan oleh algoritma machine learning. Langkah ini penting untuk memastikan dataset bersih dan siap digunakan dalam

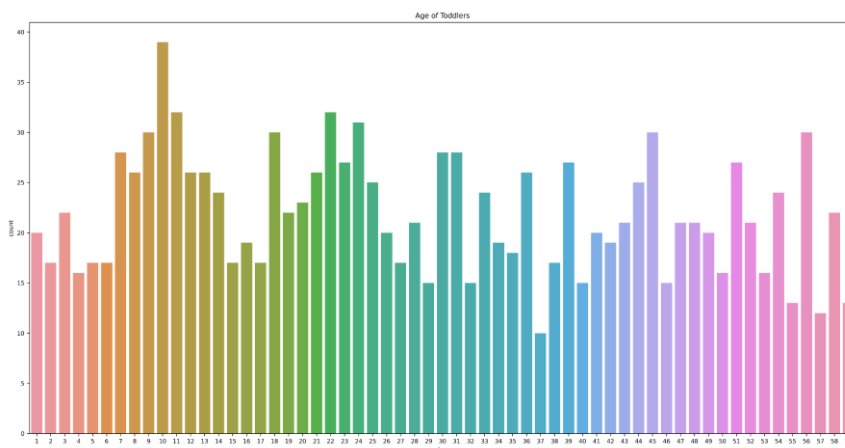


pembuatan model klasifikasi. Tahap kedua, data akan dibagi menjadi dua bagian, yaitu data training sebanyak 70% dari total data dan data testing sebanyak 30%. Pembagian ini memungkinkan kita untuk melatih model pada data tertentu dan menguji kinerja model pada data yang berbeda untuk menghindari overfitting. Tahap ketiga pembuatan model klasifikasi, model klasifikasi akan dibangun menggunakan lima algoritma berbeda, yaitu LR, DT, RF, KNN, dan SVM. Setiap algoritma akan menghasilkan model klasifikasi yang berpotensi untuk mengklasifikasikan balita stunting. Tahap keempat evaluasi performa model, performa model yang telah dibangun akan dievaluasi menggunakan confusion matrix dan berbagai metrik evaluasi, seperti akurasi, presisi, recall, F1-score, dan AUC ROC (Area Under the Curve Receiver Operating Characteristic). Evaluasi ini akan memberikan wawasan tentang sejauh mana masing-masing model dapat mengidentifikasi stunting pada balita. Tahap terakhir, perbandingan performa kelima model klasifikasi untuk menentukan model terbaik. Dalam tahap ini, model-model tersebut akan dibandingkan berdasarkan metrik evaluasi yang telah dianalisis sebelumnya. Tujuan akhir dari penelitian ini adalah untuk mengidentifikasi algoritma yang memberikan hasil terbaik dalam mengatasi tugas klasifikasi stunting pada balita.

2.1 Dataset

Dataset yang digunakan dalam penelitian ini adalah data Elektronik Pencatatan dan Pelaporan Gizi Balita Berbasis Masyarakat (EPPGBM) yang diperoleh dari Puskesmas Joggat, Lombok Tengah, Indonesia. Dataset ini terdiri dari total 1295 sampel data, yang terdiri dari 1000 balita yang tidak terindikasi mengalami stunting dan 295 balita yang terindikasi mengalami stunting. Dataset ini mengandung lima atribut atau fitur penting yang digunakan dalam analisis, yaitu:

1. Jenis kelamin balita: Atribut ini mengidentifikasi apakah balita adalah laki-laki atau perempuan.
2. Usia balita: Usia balita diukur dalam bulan, dengan rentang usia balita antara 1 hingga 60 bulan. Informasi usia ini memiliki relevansi signifikan dalam mengevaluasi status gizi balitas.
3. Tinggi badan balita: Atribut ini mencatat tinggi badan balita dalam satuan tertentu (misalnya, sentimeter).
4. Berat badan balita: Atribut ini mencatat berat badan balita dalam satuan tertentu (misalnya, kilogram).
5. Kelas biner: Kelas biner merupakan variabel target dalam dataset ini. Nilai 0 digunakan untuk mengidentifikasi balita yang tidak terindikasi mengalami stunting, sementara nilai 1 digunakan untuk mengidentifikasi balita yang terindikasi mengalami stunting.



Gambar 2. Distribusi usia balita pada dataset

Berdasarkan Gambar 2, atribut usia balita yang ada pada dataset memiliki rentan nilai 1 sampai dengan 60 bulan yang jumlahnya bervariasi.

2.2 Data Preprocessing

Data preprocessing adalah langkah penting dalam analisis data dan machine learning, hal itu dapat meningkatkan kualitas data mentah, sehingga compatible untuk diproses lebih lanjut [14]. Pembersihan data, transformasi data adalah proses yang dapat dilakukan pada data. Pembersihan data melibatkan penanganan nilai yang hilang dan menghilangkan noise [15]. Transformasi data mencakup penskalaan, agregasi, atau normalisasi data untuk mempersiapkannya untuk dianalisis [16].

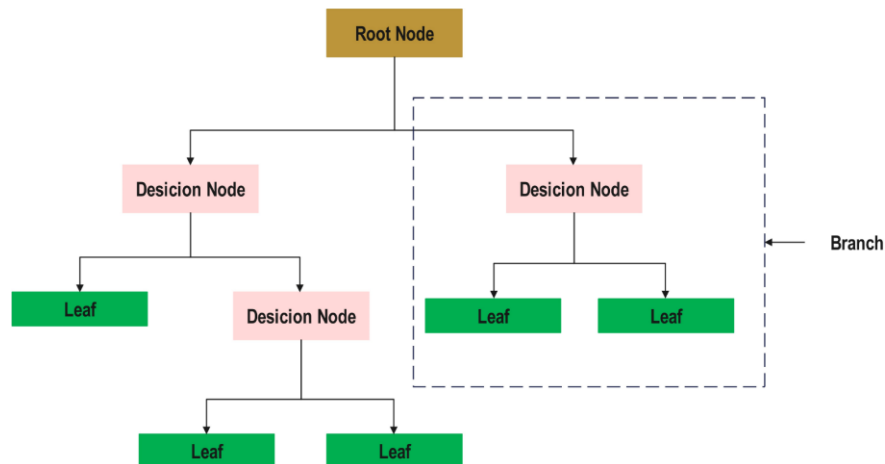
2.3 Logistic Regression (LR)

Logistic Regression (LR) adalah algoritma pembelajaran mesin yang biasanya digunakan untuk tugas klasifikasi yang kelas targetnya bersifat kategorikal, multinomial, atau ordinal. Hasilnya disajikan dalam format probabilitas. Metodologi yang digunakan menggabungkan persamaan regresi linier bersama dengan fungsi sigmoid, yang biasa disebut fungsi logistik, untuk memperkirakan probabilitas dalam rentang 0 hingga 1. Variabel acak mengambil nilai 1 dengan probabilitas p dan nilai 0 dengan probabilitas $1-p$ [17].



2.4 Decision Tree (DT)

Pohon keputusan adalah metode klasifikasi untuk prediksi yang memiliki kemiripan dengan struktur hierarki atau pohon [18]. Menetapkan aturan keputusan dan data ke pohon keputusan merupakan operasi mendasar dari pohon keputusan. Menyederhanakan keputusan yang kompleks adalah manfaat paling signifikan dari penerapan pohon keputusan. Prosedur ini memungkinkan pengambil keputusan untuk membuat konsep penyelesaian masalah yang lebih efektif [19].



Gambar 3. Decision Tree

Berdasarkan Gambar 3, pohon keputusan terdiri dari simpul-simpul fundamental yang dikenal sebagai root dapat dilihat pada gambar diatas. Root node mewakili keseluruhan kumpulan data dan berfungsi sebagai titik awal untuk proses pengambilan keputusan. Decision/Internal Node mewakili keputusan tentang fitur masukan. Leaf/Terminal Node adalah sebuah node yang tidak memiliki node anak dan mewakili label kelas atau nilai numerik. Splitting adalah proses membagi sebuah node menjadi dua atau lebih sub-node berdasarkan kriteria pemisahan dan fitur yang dipilih. Node yang terbagi menjadi satu atau lebih node anak dikenal sebagai parent node. Node yang muncul ketika parent node dipecah disebut sebagai child nodes [20].

2.5 K-Nearest Neighbor

K-Nearest Neighbor merupakan bagian dari algoritma supervised machine learning yang mudah diimplementasikan dan mampu menyelesaikan masalah yang kompleks. Akan tetapi KNN sering disebut lazy learning karena mengharuskan setiap data pelatihan disimpan dan setiap data baru yang akan diprediksi akan diukur kesamaannya ke seluruh data latih [21]. Untuk mengklasifikasikan data baru yang belum dimasukkan ke dalam kelas tertentu, algoritma KNN menentukan k (tetangga) titik data mana yang jaraknya paling dekat dengan data baru tersebut. Untuk data baru, kelas prediksi akan ditentukan menggunakan frekuensi kelas dari kumpulan data terdekat yaitu k. Untuk mencegah jarak yang sama selama prosedur klasifikasi, k biasanya diberi nilai ganjil. Dengan menggunakan fungsi jarak seperti Manhattan distance, jarak atau kedekatan tetangga dihitung [22]. Rumus Manhattan distance dapat dilihat pada persamaan 1 dibawah.

$$\text{Manhattan}_{d(x,y)} = \sum_{i=1}^n |x_i - y_i| \quad (1)$$

2.6 Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah model klasifikasi yang menggunakan hyperplane sebagai perantara untuk membedakan berbagai kategori. Biasanya, SVM digunakan untuk tugas klasifikasi biner; namun, penerapannya dapat diperluas untuk mencakup skenario klasifikasi multikelas. Hyperplane di SVM ditentukan dengan menghitung margin maksimum untuk setiap kategori dan mencari titik tengahnya. Penentuan margin maksimum didasarkan pada pemeriksaan titik-titik berdekatan yang terletak antar kategori [23], [24]. Keterbatasan utama teknik ini adalah memerlukan lebih banyak memori saat memproses data dalam jumlah besar. Algoritma ini berhasil mendeteksi ketidaklinieran dan ketersebaran data informasi. SVM adalah metode yang kompleks, dan beberapa penelitian telah membuktikan kelayakannya [25].

2.7 Evaluasi Performa

Untuk mengevaluasi model klasifikasi, confusion matrix bisa dimanfaatkan dalam pembelajaran mesin memungkinkan analisis yang lebih mendalam terhadap kinerja algoritma yang akan dibandingkan [26]. Confusion matrix adalah tabel yang sering digunakan untuk menggambarkan kinerja model klasifikasi pada sekumpulan data uji yang nilai sebenarnya diketahui. Jika confusion matrix telah berhasil didapatkan dari setiap algoritma yang diimplementasikan, maka nilai metrik seperti akurasi, precision, recall, f1-score, AUC-ROC curve. Tabel confusion matrix dapat dilihat pada gambar dibawah ini.



Confusion Matrix

Actual Class	0	True Negative (TN)	False Positive (FP)
	1	False Negative (FN)	True Positive (TP)
		0	1
		Predicted Class	

Gambar 4. Tabel Confusion Matrix

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \times 100\%, \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\%, \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\%, \quad (5)$$

Berdasarkan Gambar 4, istilah "TN" mewakili singkatan dari true negative, yang mengindikasikan jumlah contoh negatif yang berhasil diklasifikasikan secara benar. Sama halnya dengan "TP," yang merujuk kepada true positive, menggambarkan jumlah contoh positif yang diklasifikasikan dengan tepat. Perlu dicatat bahwa "FP" merujuk pada positif palsu, yang mencerminkan jumlah contoh negatif sebenarnya yang salah diklasifikasikan sebagai positif, sementara "FN" mengartikan negatif palsu, yakni jumlah contoh positif sebenarnya yang keliru diklasifikasikan sebagai negatif [27],[28]. Akurasi mewakili persentase semua prediksi yang akurat. Presisi, yang sering disebut sebagai akurasi prediksi positif, mencerminkan sejauh mana prediksi positif sebenarnya adalah positif. Dalam istilah yang lebih sederhana, presisi mengukur sejauh mana nilai positif yang telah diidentifikasi adalah benar dibandingkan dengan total jumlah nilai positif. Recall, atau dikenal juga sebagai TP rate, menggambarkan sejauh mana kasus positif sebenarnya berhasil teridentifikasi oleh model. F1-Score merupakan rata-rata harmonik dari sensitivitas dan presisi, kedua faktor tersebut dianggap signifikan [29], [30].

3. HASIL DAN PEMBAHASAN

Dalam pembahasan mengenai hasil penelitian ini, bahasa pemrograman yang digunakan adalah Python dengan library Scikit-Learn. Terdapat lima algoritma klasifikasi yang akan dibandingkan, yaitu LR, DT, RF, KNN, dan SVM. Setiap model klasifikasi dievaluasi melalui confusion matrix, dengan pengukuran metrik seperti akurasi, presisi, recall, f1-score, dan AUC ROC. Kemudian, performa masing-masing algoritma dibandingkan untuk menemukan algoritma yang memberikan hasil terbaik. Analisis hasil ini menjadi fokus dalam pembahasan kami, yang akan kami bahas lebih lanjut. Hasil pengujian setiap algoritma yang memanfaatkan confusion matrix dapat dilihat pada tabel 1 dibawah.

Tabel 1. Confusion matrix hasil pengujian lima algoritma

Algoritma	TP	TN	FP	FN
LR	65	272	20	32
DT	88	279	13	9
RF	81	286	6	16
KNN	85	284	8	12
SVM	42	291	1	55

Berdasarkan tabel 1, metrik performa seperti akurasi, presisi, recall, f1-score dapat dihitung dengan menggunakan persamaan 2-5. Contoh menghitung metrik performa dari algoritma KNN.

$$\text{Accuracy} = (85+284)/(85+284+8+15) \times 100\% = 94,85\%$$

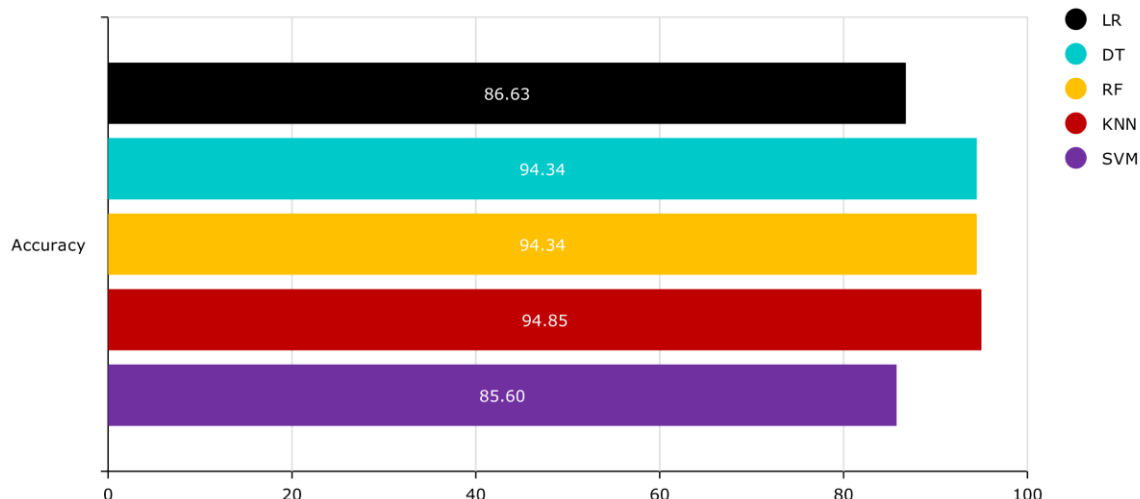
$$\text{Precision} = (85)/(85+8) \times 100\% = 91,39\%$$

$$\text{Recall} = (85)/(85+12) \times 100\% = 87,62\%$$



$F1\text{-Score} = 2 \times (91,39 \times 87,62) / (91,39 + 87,62) \times 100\% = 89,47\%$

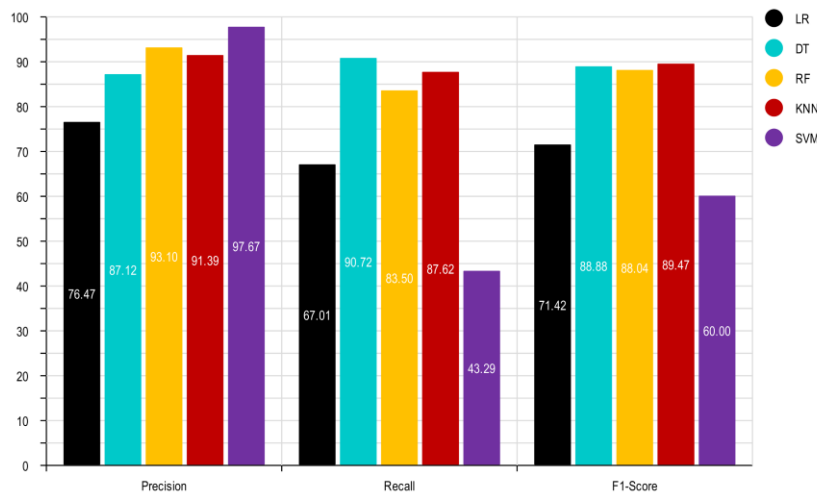
Performa metric evaluasi setiap algoritma dapat dilihat pada gambar 5, 6, dan 7 dibawah.



Gambar 5. Akurasi setiap algoritma

Berdasarkan Gambar 5, hasil evaluasi menunjukkan adanya perbedaan signifikan dalam akurasi antara algoritma-algoritma tersebut. Dalam hal ini, akurasi menjadi parameter utama dalam mengukur performa model. Hasil menunjukkan bahwa KNN pada $K = 3$ menghasilkan akurasi tertinggi sebesar 94,85%, menunjukkan bahwa model KNN memiliki tingkat keakuratan yang baik dalam mengklasifikasikan stunting. Meskipun Random Forest juga menghasilkan akurasi yang sama dengan KNN, penting untuk mempertimbangkan faktor-faktor lain seperti waktu komputasi dan kompleksitas model. DT juga menghasilkan akurasi yang tinggi sebesar 94,34%, namun, model ini cenderung rentan terhadap overfitting. LR menunjukkan akurasi yang lebih rendah sebesar 86,63%, dan SVM memiliki akurasi terendah sebesar 85,60%. Masing-masing akurasi yang dihasilkan dapat dilihat pada gambar diatas.

Penggunaan metrik tambahan seperti presisi, recall, F1-score, dan analisis kurva ROC dapat memberikan wawasan lebih mendalam tentang kemampuan model dalam mengatasi permasalahan klasifikasi stunting. Dengan menganalisis metrik-metrik ini secara komprehensif, penelitian dapat mengidentifikasi model terbaik yang tidak hanya memiliki akurasi tinggi tetapi juga keseimbangan yang baik antara presisi dan recall untuk penanganan masalah stunting pada balita secara efektif. Masing-masing hasil presisi, recall, dan F1-score dari kelima algoritma dapat dilihat pada gambar 6 dibawah.



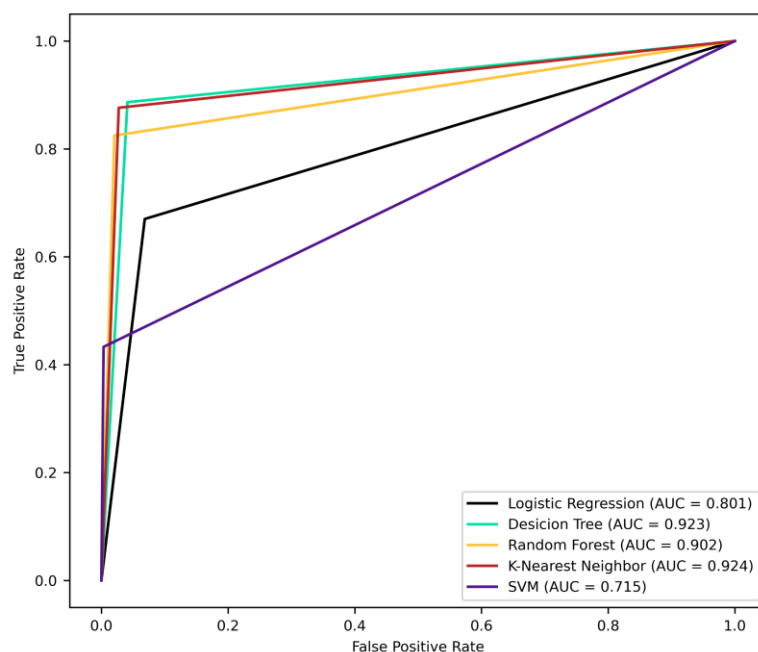
Gambar 6. Presisi, recal, dan f1-score

Berdasarkan Gambar 6, evaluasi performa lima algoritma klasifikasi, LR, DT, RF, KNN, SVM untuk deteksi stunting pada balita memberikan wawasan tentang kemampuan masing-masing algoritma dalam mengatasi tugas klasifikasi yang sensitif ini.

LR menghasilkan presisi 76,47%, yang mengindikasikan bahwa model ini mampu mengklasifikasikan stunting dengan benar dalam sebagian besar kasus. Recall 67,01% menunjukkan kemampuan model dalam mengidentifikasi stunting pada balita yang terindikasi dengan cukup baik. F1-score 71,42% mencerminkan keseimbangan antara presisi dan recall yang baik. DT mencapai presisi 87,12% dan recall 90,72%,



mengindikasikan kemampuannya dalam mengidentifikasi stunting dengan tingkat keakuratan yang tinggi. F1-score sebesar 88,88% menunjukkan kinerja yang seimbang antara presisi dan recall. RF memperoleh presisi tinggi sebesar 93,10%, tetapi recall yang lebih rendah sebesar 83,50%. F1-score 88,04% menunjukkan kemampuan model untuk menjaga keseimbangan antara presisi dan recall. KNN mencapai presisi 91,39% dan recall 87,62%, dengan F1-score sebesar 89,47%, mengindikasikan kinerja yang baik dalam mengklasifikasikan stunting dengan keseimbangan yang baik antara presisi dan recall. SVM memiliki presisi yang sangat tinggi sebesar 97,67%, tetapi recall yang rendah sebesar 43,29%. F1-score sebesar 60,00% menunjukkan bahwa model ini mungkin mengalami kesulitan dalam mengidentifikasi sebagian besar kasus stunting yang sebenarnya terjadi. KNN menonjol sebagai model terbaik dengan F1-score tertinggi sebesar 89,47%. Meskipun SVM memiliki presisi tinggi, recall yang rendah membuat F1-score-nya lebih rendah. Hasil ini menunjukkan bahwa, dalam konteks deteksi stunting pada balita, keberimbangan antara kemampuan model untuk mengidentifikasi kasus stunting dan menjaga tingkat akurasi secara keseluruhan adalah kunci untuk pemilihan model terbaik. KNN muncul sebagai pilihan terbaik dengan performa yang baik dan seimbang dalam mengatasi masalah ini.



Gambar 7. AUC ROC curve

Dalam evaluasi performa lima algoritma klasifikasi untuk deteksi stunting pada balita, nilai Area Under the Curve (AUC) dari kurva Receiver Operating Characteristic (ROC) digunakan untuk mengukur kemampuan algoritma dalam membedakan antara kelas positif (stunting) dan kelas negatif (non-stunting). AUC adalah metrik yang berguna untuk mengukur sejauh mana model mampu mengklasifikasikan data dengan benar, terutama dalam kasus klasifikasi yang sensitif seperti ini.

Berdasarkan Gambar 7, KNN dapat dianggap sebagai model terbaik. Hal ini disebabkan oleh beberapa faktor. AUC yang tinggi menunjukkan bahwa KNN memiliki kemampuan yang kuat dalam membedakan kasus stunting dengan akurasi yang baik. Dalam kasus ini, AUC yang dihasilkan oleh masing-masing algoritma adalah sebagai berikut: LR (0,801), DT (0,923), RF (0,902), KNN (0,924), dan SVM (0,715). KNN juga telah menunjukkan performa yang seimbang dalam metrik evaluasi sebelumnya bahwa KNN memiliki presisi yang tinggi (91,39%) dan recall yang baik (87,62%). KNN memberikan keseimbangan yang baik antara kemampuan untuk mengidentifikasi stunting pada balita (ditunjukkan oleh nilai recall) dan meminimalkan kesalahan dalam mengklasifikasikan kasus non-stunting sebagai stunting (ditunjukkan oleh nilai precision). F1-score yang tinggi dan AUC yang tinggi lebih mempertegas bahwa KNN adalah model terbaik dibandingkan empat algoritma yang lain. Dengan demikian, KNN tidak hanya mencapai AUC tertinggi tetapi juga menunjukkan performa yang konsisten dan seimbang dalam tugas klasifikasi stunting pada balita, yang didukung oleh hasil evaluasi yang telah dilakukan sebelumnya.

4. KESIMPULAN

Berdasarkan hasil pengujian dan analisis perbandingan algoritma LR, DT, RF, KNN, SVM untuk deteksi stunting pada balita. KNN pada $K = 3$ menghasilkan akurasi tertinggi sebesar 94,85%, menjadikannya model terbaik dalam mengklasifikasikan stunting pada balita. Decision Tree dan Random Forest juga menghasilkan akurasi yang tinggi sebesar 94,34%. Sementara itu, Logistic Regression memiliki akurasi lebih rendah sebesar 86,63%, dan SVM



mencapai akurasi terendah sebesar 85,60%. Metrik tambahan seperti presisi, recall, dan F1-score memberikan wawasan yang lebih mendalam tentang performa algoritma. Dalam hal ini, KNN memiliki presisi yang tinggi (91,39%) dan recall yang baik (87,62%), menghasilkan F1-score tertinggi sebesar 89,47%. Penggunaan AUC dalam analisis ROC juga mendukung pemilihan KNN sebagai model terbaik, dengan AUC sebesar 0,924, menunjukkan kemampuan KNN dalam membedakan antara kelas stunting dan non-stunting. Penelitian ini mengidentifikasi KNN sebagai model terbaik dalam deteksi stunting pada balita berdasarkan evaluasi akurasi, presisi, recall, F1-score, dan AUC. KNN memiliki performa yang konsisten dan seimbang, menjadikannya pilihan unggul dalam mengatasi tugas klasifikasi yang sensitif ini. Dengan menggunakan metrik ini, penelitian ini dapat mengidentifikasi model terbaik untuk mengatasi masalah stunting pada balita secara efektif.

ACKNOWLEDGEMENTS

Penelitian ini mendapat dukungan dana dari Direktorat Riset, Teknologi, dan Pengabdian Kepada Masyarakat (DRTPM) dan Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemendikbudristek) Republik Indonesia, berdasarkan Surat Keputusan No. 0536/E5/ PG.02.00/2023 dan Perjanjian/Kontrak No. 181/E5-PG.02.00.PL/2023.

REFERENCES

- [1] TNP2K, "100 Kabupaten/Kota Prioritas untuk Intervensi Anak Kerdil (Stunting): Tim Nasional Percepatan Penanggulangan Kemiskinan," Jakarta, vol. 2, no. c, p. 287, 2021, [Online]. Available: http://www.tnp2k.go.id/images/uploads/downloads/Buku_Ringkasan_Stunting.pdf%0Awww.tnp2k.go.id
- [2] S. Lonang and D. Normawati, "Klasifikasi Status Stunting Pada Balita Menggunakan K-Nearest Neighbor Dengan Feature Selection Backward Elimination," J. MEDIA Inform. BUDIDARMA, vol. 6, no. 1, p. 49, Jan. 2022, doi: 10.30865/mib.v6i1.3312.
- [3] K. L. Perdue et al., "Using functional near-infrared spectroscopy to assess social information processing in poor urban Bangladeshi infants and toddlers," Dev. Sci., vol. 22, no. 5, 2019, doi: 10.1111/desc.12839.
- [4] S. Lonang, A. Yudhana, and M. K. Biddinika, "Rancangan Sistem Klasifikasi Kekurangan Gizi Balita Dengan Metode K-Nearest Neighbor," J. Inform. dan Rekayasa Perangkat Lunak, vol. 5, no. 1, p. 73, Mar. 2023, doi: 10.36499/jinrpl.v5i1.7834.
- [5] D. Sudigyo, A. A. Hidayat, R. Nirwantono, R. Rahutomo, J. P. Trinugroho, and B. Pardamean, "Literature study of stunting supplementation in Indonesian utilizing text mining approach," Procedia Comput. Sci., vol. 216, no. 2022, pp. 722–729, 2023, doi: 10.1016/j.procs.2022.12.189.
- [6] H. Abbad Ur Rehman, C. Y. Lin, and Z. Mushtaq, "Effective K-Nearest Neighbor Algorithms Performance Analysis of Thyroid Disease," J. Chinese Inst. Eng. Trans. Chinese Inst. Eng. A, vol. 44, no. 1, pp. 77–87, 2021, doi: 10.1080/02533839.2020.1831967.
- [7] L. A. Al Hak, "Diabetes Prediction Using Binary Grey Wolf Optimization and Decision Tree," Int. J. Comput., vol. 21, no. 4, pp. 489–494, 2022, doi: 10.47839/ijc.21.4.2785.
- [8] D. A. Salem and E. M. Hashim, "Impact of Data Pre-Processing on Covid-19 Diagnosis Using Machine Learning Algorithms," Orig. Res. Pap. Int. J. Intell. Syst. Appl. Eng. IJISAE, vol. 2023, no. 1s, pp. 164–171, 2023, [Online]. Available: www.ijisae.org
- [9] S. Lonang, A. Yudhana, and M. K. Biddinika, "Performance Analysis for Classification of Malnourished Toddlers Using K-Nearest Neighbor," Sci. J. Informatics, vol. 10, no. 3, p. 313, 2023, [Online]. Available: <http://journal.unnes.ac.id/nju/index.php/sji>
- [10] N. M. Mathkunti and S. Rangaswamy, "Machine Learning Techniques to Identify Dementia," SN Comput. Sci., vol. 1, no. 3, p. 118, May 2020, doi: 10.1007/s42979-020-0099-4.
- [11] S. Ndagijimana, I. H. Kabano, E. Masabo, and J. M. Ntaganda, "Prediction of Stunting among Under-5 Children in Rwanda Using Machine Learning Techniques," J. Prev. Med. Public Heal., vol. 56, no. 1, pp. 41–49, 2023, doi: 10.3961/jpmph.22.388.
- [12] R. Kaur, "Predicting diabetes by adopting classification approach in data mining," Int. J. Informatics Vis., vol. 3, no. 2–2, pp. 218–221, 2019, doi: 10.30630/joiv.3.2-2.229.
- [13] N. Zamri et al., "River quality classification using different distances in k-nearest neighbors algorithm," Procedia Comput. Sci., vol. 204, no. 2021, pp. 180–186, 2022, doi: 10.1016/j.procs.2022.08.022.
- [14] V. Çetin and O. Yıldız, "A comprehensive review on data preprocessing techniques in data analysis," Pamukkale Univ. J. Eng. Sci., vol. 28, no. 2, pp. 299–312, 2022, doi: 10.5505/pajes.2021.62687.
- [15] V. Werner de Vargas, J. A. Schneider Aranda, R. dos Santos Costa, P. R. da Silva Pereira, and J. L. Victória Barbosa, "Imbalanced data preprocessing techniques for machine learning: a systematic mapping study," Knowl. Inf. Syst., vol. 65, no. 1, pp. 31–57, 2023, doi: 10.1007/s10115-022-01772-8.
- [16] C. Fan, M. Chen, X. Wang, J. Wang, and B. Huang, "A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data," Front. Energy Res., vol. 9, Mar. 2021, doi: 10.3389/fenrg.2021.652801.
- [17] J. Goyal, P. Khandnor, and T. C. Aseri, "A Comparative Analysis of Machine Learning classifiers for Dysphonia-based classification of Parkinson's Disease," International Journal of Data Science and Analytics, vol. 11, no. 1, pp. 69–83, 2021, doi: 10.1007/s41060-020-00234-0.
- [18] F. D. Adhinata and A. Junaidi, "Gender Classification on Video Using FaceNet Algorithm and Supervised Machine Learning," Int. J. Comput. Digit. Syst., vol. 11, no. 1, pp. 199–208, Jan. 2022, doi: 10.12785/ijcds/110116.
- [19] N. Bin Noor, M. S. Anwar, and M. Dey, "Comparative Study Between Decision Tree, SVM and KNN to Predict Anaemic



- Condition,” in 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), IEEE, Nov. 2019, pp. 24–28. doi: 10.1109/BECITHCON48839.2019.9063188.
- [20] A. Ali, S. Naem, S. Anam, and M. M. Ahmed, “A Supervised Machine Learning Algorithms: Applications, Challenges, and Recommendations,” pp. 1–10, 2023
- [21] S. Lonang, A. Yudhana, and M. K. Biddinika, “Performance Analysis for Classification of Malnourished Toddlers Using K-Nearest Neighbor,” *Sci. J. Informatics*, vol. 10, no. 3, p. 313, 2023, doi: 10.15294/sji.v10i3.45196.
- [22] A. G. Pertiwi, N. Bachtiar, R. Kusumaningrum, I. Waspada, and A. Wibowo, “Comparison of performance of k-nearest neighbor algorithm using smote and k-nearest neighbor algorithm without smote in diagnosis of diabetes disease in balanced data,” *J. Phys. Conf. Ser.*, vol. 1524, no. 1, p. 012048, Apr. 2020, doi: 10.1088/1742-6596/1524/1/012048.
- [23] Y. Xi, M. Bi, M. Xin, C. Wang, J. Yu, and W. Hu, “Simplified SVM Equalization Algorithm Based on Single Hyperplane Training Enabled 50Gb/s PAM-4/8 With 10-G Optics in NG-PON System,” *IEEE Photonics J.*, vol. 13, no. 4, pp. 1–7, Aug. 2021, doi: 10.1109/JPHOT.2021.3103195.
- [24] A. Peryanto, A. Yudhana, and R. Umar, “Convolutional Neural Network and Support Vector Machine in Classification of Flower Images,” *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 8, no. 1, pp. 1–7, Mar. 2022, doi: 10.23917/khif.v8i1.15531.
- [25] R. A. Gulhane and S. R. Gupta, “Machine Learning Approach for Early Disease Prediction and Risk Analysis,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 10, no. 3s, pp. 27–32, 2022.
- [26] B. Sahu, S. Mohanty, and S. Rout, “A Hybrid Approach for Breast Cancer Classification and Diagnosis,” *ICST Trans. Scalable Inf. Syst.*, vol. 6, no. 20, p. 156086, Jul. 2018, doi: 10.4108/eai.19-12-2018.156086.
- [27] A. Vanacore, M. S. Pellegrino, and A. Ciardiello, “Fair evaluation of classifier predictive performance based on binary confusion matrix,” *Comput. Stat.*, Nov. 2022, doi: 10.1007/s00180-022-01301-9.
- [28] P. Singh, N. Singh, K. K. Singh, and A. Singh, “Diagnosing of disease using machine learning,” *Mach. Learn. Internet Med. Things Healthc.*, pp. 89–111, 2021, doi: 10.1016/B978-0-12-821229-5.00003-3.
- [29] J. Y. Ho et al., “Towards a time and cost effective approach to water quality index class prediction,” *J. Hydrol.*, vol. 575, pp. 148–165, 2019, doi: 10.1016/j.jhydrol.2019.05.016.
- [30] V. Kumar, G. S. Lalotra, and R. K. Kumar, “Improving performance of classifiers for diagnosis of critical diseases to prevent COVID risk,” *Comput. Electr. Eng.*, vol. 102, no. January, p. 108236, Sep. 2022, doi: 10.1016/j.compeleceng.2022.108236.