

Optimasi Model Algoritma Klasifikasi menggunakan Metode Bagging pada Stunting Balita

Gibran Nasrizal Masacgi^{1,*}, Muhammad Syaifur Rohman¹

¹ Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Indonesia

* Correspondence: 111202012516@dinus.ac.id

Copyright: © 2023 by the authors

Received: 30 Oktober 2023 | Revised: 2 November 2023 | Accepted: 6 November 2023 | Published: 20 Desember 2023

Abstrak

Stunting merupakan masalah serius dalam kesehatan balita yang mengalami kurang gizi pada masa perkembangannya. Asupan gizi yang kurang atau kurangnya pengetahuan mengenai gizi dapat menjadi faktor terjadinya stunting. Penelitian ini bertujuan untuk optimasi algoritma klasifikasi stunting pada balita yang menggunakan algoritma *K-Nearest Neighbor*, *Support Vector Machin*, *Naïve Bayes* dengan menggunakan optimasi *Bagging*. Penelitian ini menggunakan data balita berjumlah 10.000 *record*, 7 atribut, dan 2 kelas yang diambil dari Puskesmas Krobokan, Kota Semarang. Berdasarkan hasil temuan pada penelitian ini, Metode *Bagging* menunjukkan kinerja yang sangat baik dengan akurasi 89,77%, presisi 95,57%, recall 85,45%, dan F1-score 90,27%. Ini menunjukkan potensi besar dalam meningkatkan kemampuan model dalam klasifikasi data dengan tingkat ketepatan, kelengkapan, dan keseimbangan yang tinggi. Model SVM memiliki recall yang tinggi 80,59%, sementara Model *Naïve Bayes* mendapat F1-score 71,84% ini menandakan bahwa *naïve bayes* memiliki keseimbangan baik antara presisi dan recall. Secara keseluruhan, Model *Bagging* menjadi pilihan terbaik dengan kinerja yang jauh lebih unggul dibandingkan dengan model-model lainnya. Berdasarkan hasil yang diperoleh, metode *bagging* dapat memprediksi data stunting dengan akurat, menunjukkan tingkat kebenaran yang sangat baik dalam upaya prediksi stunting.

Kata kunci: *bagging; k-nearest neighbour; naive bayes; stunting; support vector machine*

Abstract

Stunting is a serious problem for the health of toddlers, who experience malnutrition during their development. Poor nutritional intake or a lack of knowledge about nutrition can be a factor in stunting. This study aims to optimize the stunting classification algorithm in toddlers using the K-Nearest Neighbor algorithm, Support Vector Machine, and Naïve Bayes by using bagging optimization. This study uses toddler data totaling 10,000 records, 7 attributes, and 2 classes taken from Krobokan Health Center, Semarang City. Based on the findings of this study, the Bagging Method shows excellent performance with 89.77% accuracy, 95.57% precision, 85.45% recall, and a 90.27% F1-score. This shows great potential for improving the model's ability to classify data with a high degree of accuracy, completeness, and balance. The SVM model has a high recall of 80.59%, while the Naïve Bayes model gets an F1-score of 71.84%, indicating that Naïve Bayes has a good balance between precision and recall. Overall, the Bagging Model is the best choice, with far superior performance compared to other models. Based on the results obtained, the Bagging Method can predict stunting data accurately, showing a very good level of truth in stunting prediction efforts.

Keywords: *bagging; k-nearest neighbour; naive bayes; stunting; support vector machine*

PENDAHULUAN

Stunting merupakan masalah serius dalam kesehatan balita yang mengalami kekurangan gizi kronis pada masa pertumbuhan dan perkembangan balita dari masa awal hidupnya



(Matdoan et al., 2022). Asupan gizi pada balita tentunya sangat penting untuk diperhatikan karena pada masa tersebut, balita membutuhkan asupan gizi yang cukup untuk masa pertumbuhan (Islam et al., 2022). Pemberian asupan gizi yang cukup tentunya sangat mempengaruhi pertumbuhan fisik dan perkembangan otak pada balita. Stunting memiliki dampak jangka panjang dan jangka pendek. Dampak jangka panjang bagi fisik, seperti pembentukan postur tubuh balita saat dewasa kemungkinan mempunyai postur fisik yang buruk dan dapat mempengaruhi prestasi akademik karena mengalami kekurangan dalam perkembangan intelektual (Ginting & Pandiangan, 2019). Sedangkan dampak jangka pendek dapat menurunkan daya tahan tubuh, gangguan perkembangan kognitif dan masalah psikologi (Nirmalasari, 2020).

Pada usia 1.000 hari pertama adalah *golden age* bayi, meskipun begitu dalam rentang usia 0-59 bulan masih mengalami permasalahan gizi (Rahmadhita, 2020). Berdasarkan data dari SSGI (Survei Status Gizi Indonesia) tahun 2022 angka prevalensi di Jawa Tengah mencapai 20,8% sedangkan di Indonesia mencapai 21,6% (Kemenkes, 2022). Dalam era teknologi saat ini, teknik data mining dapat digunakan untuk mendeteksi stunting pada balita dan menjadi bagian dari upaya untuk mengatasi masalah ini. Oleh karena itu, dalam penelitian ini penerapan data mining dilakukan dengan bantuan pemodelan pada *machine learning* yang menggunakan tiga algoritma yaitu *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* menggunakan optimasi *Bagging*. Dengan memaksimalkan hasil prediksi melalui penggunaan teknik optimasi yang dikenal sebagai *ensambel learning*, salah satu upaya untuk meningkatkan tingkat akurasi dengan menggabungkan akurasi dari beberapa model berbeda. Terdapat beberapa metode umum dalam *ensambel learning*, seperti *bagging*, *boosting*, *stacking*, dan *voting*.

Metode *Bagging* adalah salah satu metode yang digunakan untuk meningkatkan akurasi prediksi model dengan menggabungkan hasil dari beberapa model pembelajaran yang sama, biasanya dengan menggunakan teknik *bootstrap resampling* (Saputri et al., 2022). Dengan menerapkan optimasi metode *bagging* dalam model algoritma *K-Nearest Neighbor*, *Support Vector Machine*, dan *Naive Bayes* sehingga dapat meningkatkan akurasi yang dihasilkan dari model-model tersebut. *K-Nearest Neighbor* cocok untuk klasifikasi data stunting karena *fleksibel* dan dapat menyesuaikan dengan variasi data kompleks (Efriadi et al., 2022; Hidayat et al., 2021; Purwani et al., 2022; Prayoga et al., 2023). *Support Vector Machine* mampu menangani dataset dengan banyak atribut dan memisahkan kelas stunting dengan baik (Hovi et al., 2022). *Naive Bayes* efisien dalam mengatasi fitur yang saling tergantung dalam data stunting (Armansyah & Ramli 2022; Rahayu et al., 2022; Sari et al., 2020; Susanti et al., 2022).

Pada penelitian sebelumnya tentang klasifikasi data mining banyak dilakukan dengan berbagai algoritma seperti membahas prediksi stunting menggunakan algoritma *Naive Bayes* untuk prediksi stunting mendapat akurasi 64,02% (Herliansyah et al., 2021). Penelitian lain yang dilakukan oleh Ramadhan & Khoirunnisa (2021) membahas tentang klasifikasi data malaria menggunakan algoritma *Support Vector Machine*, dimana dalam penelitian ini *Support Vector Machine* mendapat akurasi 92,3%. Penelitian lainnya yang membandingkan antara algoritma *Naive Bayes* dan *K-Nearest Neighbor* dilakukan oleh Setiawan & Triayudi (2022) dalam klasifikasi gizi balita, menghasilkan akurasi 80,60% untuk *Naive Bayes* sedangkan *K-Nearest Neighbor* 91,79%. Sedangkan penelitian yang dilakukan oleh Wahyudin et al. (2023) menggunakan dataset 300 *record* untuk digunakan dalam prediksi stunting menggunakan algoritma *Naive Bayes*, pada penelitian ini menghasilkan akurasi 85,33% tanpa menggunakan optimasi.

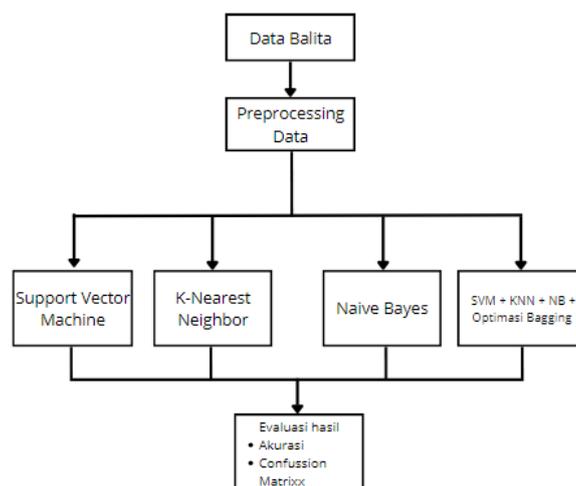
Berdasarkan hasil temuan sebelumnya masih menggunakan algoritma individu atau perbandingan saja. Meskipun telah ada upaya dari penelitian terdahulu untuk klasifikasi stunting pada balita yang menghasilkan nilai akurasi tertinggi. Namun belum ada penelitian yang menerapkan teknik *ensemble* dari model *K-Nearest Neighbor*, *Support Vector Machine*, dan *Naive Bayes*. Untuk melakukan pengembangan dari beberapa penelitian sebelumnya, pada

penelitian ini mencoba menerapkan teknik *ensemble learning* yang menggunakan penggabungan beberapa model klasifikasi yaitu *K-Nearest Neighbor*, *Support Vector Machine*, dan *Naive Bayes* untuk meningkatkan hasil akurasi, presisi, recall, dan f1-score dari model-model yang digunakan. Teknik *ensemble* pada penelitian ini menggunakan teknik metode *Bagging*. Penggunaan optimasi *bagging* dalam klasifikasi sudah dibuktikan oleh penelitian terdahulu yang dilakukan oleh Yunial (2020) dengan menggunakan optimasi *bagging* pada algoritma *decision tree* dan *neural network*, hasil akurasi *decision tree* sebelum optimasi 90,24%, setelah optimasi akurasi menjadi 90,55%, sedangkan *neural network* sebelum optimasi 88,53%, setelah optimasi akurasi menjadi 90,38%. Dapat disimpulkan bahwa penerapan optimasi *bagging* berhasil dalam meningkatkan akurasi model individu. Penelitian ini sangat penting karena meningkatkan akurasi dalam klasifikasi kasus stunting pada balita dapat membantu dalam upaya pencegahan dan penanganan lebih dini. Karena dengan menerapkan metode *bagging* dalam *machine learning* hasil dari penggabungan model dapat meningkat sehingga dapat lebih akurat dalam mengklasifikasikan stunting. Penelitian ini dapat memberikan kontribusi yang cukup baik dalam pengembangan alat deteksi stunting yang lebih handal.

Tujuan penelitian ini yaitu melakukan optimasi akurasi, presisi, recall, dan f1-score dari beberapa model klasifikasi stunting menggunakan metode *bagging*, sehingga dengan menerapkan metode *bagging* hasil akurasi, presisi, recall, dan f1-score menjadi meningkat. Dengan nilai akurasi yang tinggi, metode *bagging* dapat memberikan prediksi yang lebih akurat dan efektif dalam mendeteksi kasus stunting pada balita dengan memanfaatkan *machine learning* dalam proses klasifikasinya.

METODE

Pada penelitian ini menggunakan data balita berjumlah 10.000 *record* berasal dari Puskesmas Kecamatan Krobokan, Kota Semarang. Data tersebut akan dipilih beberapa atribut yang akan digunakan dalam pengolahan data berjumlah 7 atribut dan 2 kelas. Pada penelitian ini, skema penelitian yang digunakan seperti pada gambar 1.



Gambar 1. Skema penelitian

Berdasarkan gambar 1, sebelum melakukan pengujian dataset terdapat tahap awal yaitu preprocessing data. Pengujian dengan algoritma *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* dan optimasi *Bagging*. Setelah melakukan pengujian ada tahap evaluasi hasil untuk mengetahui sejauh mana tingkat akurat dengan presentase akurasi yang didapatkan dari model. Penelitian ini melibatkan *preprocessing* data balita, yang mencakup *mapping* untuk

mengubah data binary menjadi numeric, *class balancing* untuk menyeimbangkan jumlah sampel antar kelas, dan penggunaan *StandardScaler* untuk standarisasi data dengan range nol dan standar deviasi satu, untuk meningkatkan kinerja mesin.

Algoritma *K-Nearest Neighbor* mengukur jarak antara data yang diklasifikasikan dengan data latih terdekat sebanyak nilai K terdekat, dengan menentukan mayoritas kelas nilai K terdekat (Saputra & Primadasa, 2018). Selanjutnya klasifikasi data berdasarkan mayoritas kelas yang ada, sehingga dapat menghasilkan akurasi berdasarkan sejauh mana prediksi sesuai dengan kelas yang sebenarnya. Perhitungan jarak dilakukan pada algoritma *K-Nearest Neighbor* menggunakan rumus perhitungan jarak (Raysyah et al., 2021) :

$$d_i = \sqrt{\sum_{i=1}^p (X_{2i} - Y_{1i})^2} \quad (1)$$

Keterangan :

X = data latih

Y = data uji

i = variabel data

d = jarak

p = dimensi data

Algoritma *Support Vector Machine* mengklasifikasikan data stunting dengan mencari batas keputusan optimal untuk memisahkan kelas stunting dan tidak stunting. Dalam tahap ini, SVM mengidentifikasi vektor dukungan yang merupakan data terdekat dengan batas keputusan, dan akurasi diukur berdasarkan prediksi terhadap kelas sebenarnya (Hermanto et al., 2020).

Sementara itu, *Naive Bayes* mengklasifikasikan data stunting dengan menghitung probabilitas kelas dan atribut berdasarkan data pelatihan. Kemudian, menggunakan *Teorema Bayes*, *Naive Bayes* memprediksi kelas data baru dan mengukur akurasi dengan memeriksa sejauh mana prediksi sesuai dengan kelas sebenarnya. Adapun rumus dari *Teorema Bayes* (Wijaya & Dwiasnati, 2020) :

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \quad (2)$$

Keterangan :

X = Data class yang tidak diketahui

H = Hipotesis X data class yang spesifik

$P(H|X)$ = Hipotesis H dari kondisi X

$P(H)$ = Hipotesis H

$P(X|H)$ = Probabilitas X dari kondisi tersebut

$P(X)$ = Probabilitas X

Optimasi *bagging* dalam klasifikasi data stunting melibatkan pembagian dataset menjadi subset acak, pelatihan model pada setiap subset, dan penggabungan hasil prediksi dari beberapa model untuk meningkatkan akurasi dengan mengurangi perbedaan. Ini membantu menghasilkan prediksi yang lebih konsisten dan akurat.

Confusion matrix digunakan dalam evaluasi hasil pemodelan untuk mengukur kinerja klasifikasi. Melibatkan penghitungan *True Positive*, *True Negative*, *False Positive*, dan *False Negative* untuk menghasilkan akurasi, presisi, recall, dan F1-score, yang membantu

menganalisis sejauh mana model dapat memprediksi dengan benar dan mengidentifikasi kesalahan klasifikasi (Givari et al., 2022).

HASIL DAN PEMBAHASAN

Hasil

Pada dataset balita yang digunakan dalam penelitian ini, memiliki 2 model inputan data berupa data binary dan numerik. Hal ini dapat mempengaruhi proses mesin tidak bekerja dengan baik, maka dari itu langkah awal yang dilakukan yaitu proses *maaping* yang dilakukan untuk mengubah data binary menjadi data numerik. Pada tabel 1 adalah data asli yang belum di lakukan proses *mapping*.

Tabel 1. Data sebelum *mapping*

Jenis Kelamin	Usia	Berat Badan Lahir	Tinggi Badan Lahir	Berat Badan	Tinggi Badan	Kategori
Laki-laki	17	3,0	49	10,0	72,2	Tidak Stunting
Perempuan	11	2,9	49	2,9	65,0	Stunting

Berdasarkan data pada tabel 1, terdapat data binary yang ada pada *class* jenis kelamin dan kategori. Mengubah data binary pada *class* jenis kelamin dengan angka 1 untuk perempuan dan 0 untuk laki-laki, pada *class* kategori mengubah output stunting menjadi 1 dan tidak stunting menjadi 0. Maka dari itu proses *mapping* perlu dilakukan untuk mengubah data binary menjadi data numerik supaya dapat diproses oleh *machine learning*. Hasil data yang sudah diubah menjadi numerik, dapat dilihat pada tabel 2.

Tabel 2. Data sesudah *mapping*

Jenis Kelamin	Usia	Berat Badan Lahir	Tinggi Badan Lahir	Berat Badan	Tinggi Badan	Kategori
0	17	3,0	49	10,0	72,2	0
1	11	2,9	49	2,9	65,0	1

Pada tahap *class balancing* perlu dilakukan jika terdapat jumlah data *class* yang tidak seimbang. Data balita berjumlah 10.000 *record* dengan 2 kategori yaitu stunting dan tidak stunting. Dalam teknik ini menggunakan metode *oversampling* dengan menambahkan data pada data minoritas, yaitu data tidak stunting (0), sehingga jumlahnya setara dengan data mayoritas, yaitu data stunting (1). Data yang belum dilakukan *class balancing* dapat dilihat pada tabel 3.

Tabel 3. Data sebelum *class balancing*

	Tidak Stunting (0)	Stunting (1)
Jumlah Data	2045	7955

Berdasarkan pada tabel 3 terdapat jumlah data tidak stunting (0) dan stunting (1) yang tidak seimbang. *Class balancing* dalam *preprocessing* data digunakan untuk mengatasi ketidakseimbangan jumlah sampel antara kelas-kelas yang berbeda, meningkatkan kinerja model *machine learning* dengan memperlakukan semua kelas secara seimbang, dan menghindari bias serta *overfitting* dalam hasil prediksi. Hal ini memungkinkan model untuk lebih adil dan akurat dalam memprediksi setiap kelas dalam dataset. Pada tabel 4 adalah hasil setelah dilakukan *class balancing*.

Tahap terakhir yaitu proses standarisasi data. Pada tabel 5 terdapat data mentah yang salah satu dari nilai atributnya memiliki skala besar. Selanjutnya data tersebut akan di standarisasi data supaya tidak berdampak pada perhitungan dan menghambat kinerja mesin menjadi lebih berat. Terdapat nilai atribut yang memiliki nilai 72,2 merupakan data yang mempunyai skala paling besar dari nilai atribut yang lainnya, seperti 3,0 atau 10,0.

Tabel 4. Data sesudah *class balancing*

	Tidak Stunting (0)	Stunting (1)
Jumlah Data	7955	7955

Hal ini, nilai atribut 72.2 mungkin dapat mendominasi hasil perhitungan yang dapat mengakibatkan hasil klasifikasi tidak akurat. Pada *StandardScaler* mengubah distribusi fitur dalam dataset menjadi distribusi standar yaitu memiliki *mean (rerata)* nol dan standar deviasi satu. Dengan mengubah nilai atribut menjadi *range* 0-1 dapat memudahkan mesin dan mempercepat kinerja mesin menjadi lebih baik. Hasil data yang sudah dilakukan *StandardScaler* terdapat pada tabel 6.

Tabel 5. Data sebelum di *standardscaler*

Atribut	Nilai
Jenis Kelamin	0
Usia	17
Berat Badan Lahir	3,0
Tinggi Badan Lahir	49
Berat Badan	10,0
Tinggi Badan	72,2

Tabel 6. Data sesudah di *standardscaler*

Atribut	Nilai
Jenis Kelamin	-0,69133043
Usia	0,5630996
Berat Badan Lahir	0,96954606
Tinggi Badan Lahir	-0,44477381
Berat Badan	1,53671871
Tinggi Badan	0,42868446

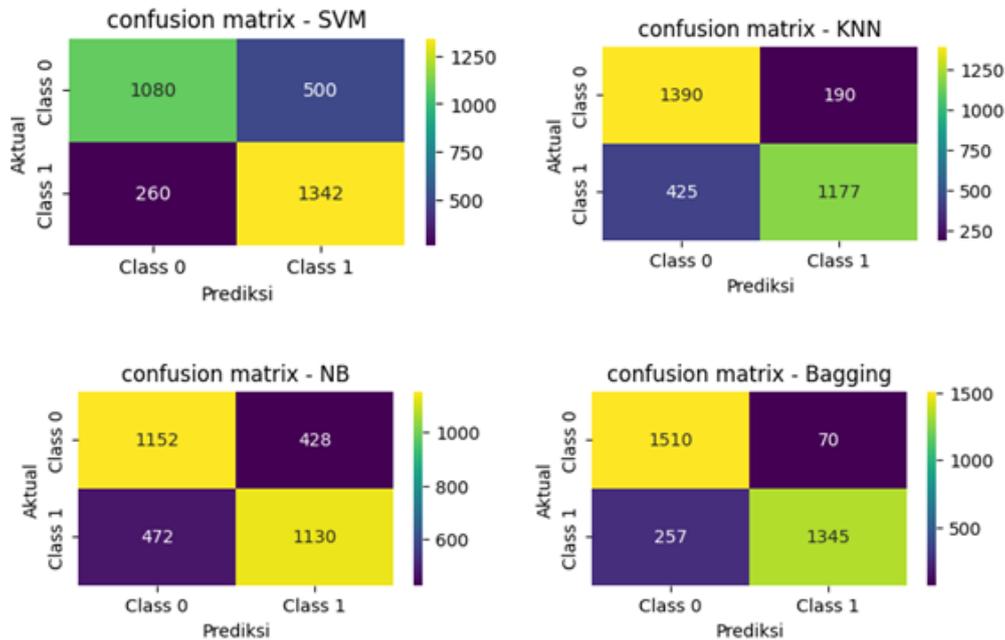
Pada penelitian ini menggunakan 3 algoritma yaitu algoritma *K-Nearest Neighbor*, *Support Vector Machine* dan *Naive Bayes* yang dioptimasi menggunakan metode *bagging*. *Bagging* digunakan untuk meningkatkan akurasi dari penggabungan ketiga model klasifikasi. Hasil klasifikasi dari ketiga algoritma tersebut serta hasil klasifikasi dari penggabungan tiga algoritma yang di optimasi dengan *bagging* memiliki nilai akurasi yang berbeda.

Pada tabel 7 adalah hasil dari algoritma yang digunakan pada penelitian ini memiliki hasil akurasi yang berbeda. Dari ketiga algoritma yang digunakan dapat disimpulkan bahwa hasil akurasi yang dihasilkan dari algoritma itu sendiri memiliki akurasi yang lebih rendah dibanding dengan penggunaan optimasi *bagging*, maka dalam penerapan optimasi *bagging* dapat meningkatkan akurasi yang lebih besar. Dengan tingkat akurasi mencapai 89.72%, Model *Bagging* mencapai hasil akurasi yang sangat signifikan. Hasil ini menunjukkan bahwa model ini memiliki kemampuan yang kuat dalam mengklasifikasikan data stunting, sehingga dapat meningkatkan akurasi dalam klasifikasi stunting secara keseluruhan.

Evaluasi hasil dengan menggunakan *confusion matrix* dari 3 algoritma dan optimasi. *Confusion matrix* digunakan untuk evaluasi hasil kinerja dari model dengan mengukur sejauh mana model dapat memberikan prediksi akurasi yang akurat. Dari 3 algoritma dan optimasi tersebut mendapat nilai akurasi, presisi, recall, dan f1-score.

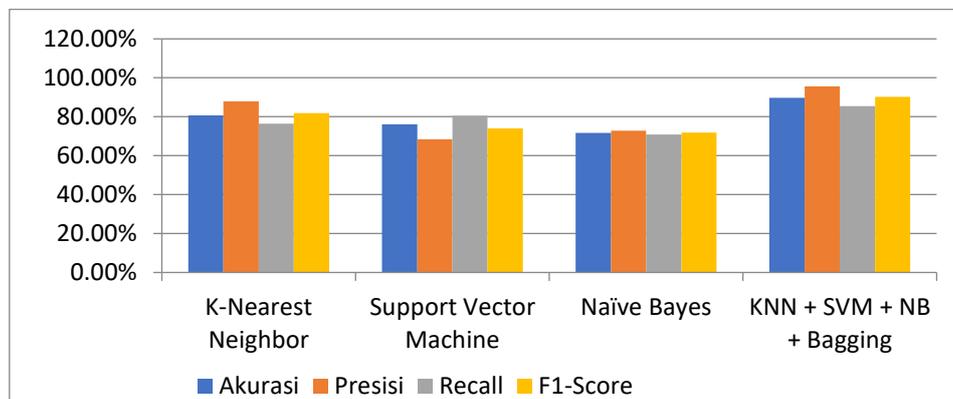
Tabel 7. Nilai akurasi

Algoritma	Akurasi
Algoritma <i>K-Nearest Neighbor</i>	80,67%
Algoritma <i>Support Vector Machine</i>	76,11%
Algoritma <i>Naïve Bayes</i>	71,71%
<i>K-Nearest Neighbor + Support Vector Machine + Naïve Bayes + Bagging</i>	89,72%



Gambar 2. Hasil *confusion matrix*

Gambar 2 menunjukkan hasil *confusion matrix*, algoritma *K-Nearest Neighbor* mendapatkan akurasi 80,71%, presisi 87,97%, recall 76,41%, f1-score 81,79%. *Support Vector Machine* akurasi 76,13%, presisi 68,35%, recall 80,59%, f1-score 74,05%. *Naïve Bayes* akurasi 71,70%, presisi 72,78%, recall 70,92%, f1-score 71,84%. *Bagging* akurasi 89,77%, presisi 95,57%, recall 85,45%, f1-score 90,27%.



Gambar 3. Perbandingan akurasi, presisi, recall, dan f1-score

Dari hasil *confusion matrix*, dapat disimpulkan bahwa penggunaan metode *bagging* sangat berpengaruh terhadap kenaikan akurasi, presisi, recall, dan f1-score dari pada penggunaan model tunggal algoritma. Pada gambar 3 adalah diagram perbandingan dari semua algoritma. Dapat disimpulkan bahwa perhitungan klasifikasi semua algoritma, dengan menggabungkan tiga algoritma menggunakan optimasi *bagging* dapat mempengaruhi nilai akurasi, presisi, recall, dan f1-score dibandingkan dengan penggunaan algoritma individu.

Pembahasan

Berdasarkan dari hasil penelitian ini dengan algoritma *K-Nearest Neighbor*, *Support Vector Machine* dan *Naive Bayes* menggunakan optimasi *Bagging* mendapat hasil akurasi yang berbeda-beda, baik dari algoritma itu sendiri atau yang di optimasi. Hasil akurasi yang didapatkan oleh algoritma *K-Nearest Neighbor* mendapatkan akurasi 80,71%, presisi 8,97%, recall 76,41%, f1-score 81,79%. *Support Vector Machine* akurasi 76,13%, presisi 68,35%, recall 80,59%, f1-score 74,05%. *Naive Bayes* akurasi 71,70%, presisi 72,78%, recall 70,92%, f1-score 71,84%. *Bagging* akurasi 89,77%, presisi 95,57%, recall 85,45%, f1-score 90,27%. Dapat disimpulkan bahwa penggunaan optimasi *bagging* pada penggabungan 3 algoritma mempengaruhi nilai akurasi, presisi, recall, dan f1-score dibandingkan dengan penggunaan algoritma individu.

Menggabungkan algoritma *K-Nearest Neighbor*, *Support Vector Machine* dan *Naive Bayes* hasil akurasi yang dihasilkan menggunakan optimasi *Bagging* mendapatkan nilai akurasi tertinggi sebesar 89,72%. Karena *bagging* merupakan pendekatan *ensemble* yang meningkatkan kinerja model klasifikasi. Prosesnya melibatkan pembuatan subset acak dari data pelatihan (*bootstrap samples*), pelatihan model-model terpisah dari setiap subset, penggabungan hasil prediksi, evaluasi model, dan penyetelan parameter. Gabungan ini bertujuan mengatasi kelemahan algoritma individu, tetapi perlu memperhatikan kompleksitas model dan evaluasi yang mendalam untuk performa yang baik. Dengan ini pemahaman setiap algoritma perlu dilakukan sebelum memilih algoritma yang akan di optimasi menggunakan *bagging*.

Berdasarkan hasil yang didapat pada penelitian ini, dapat dikatakan bahwa penerapan *bagging* sangat cocok digunakan dalam meningkatkan nilai akurasi, presisi, recall, dan f1-score. Penelitian ini menggunakan tiga algoritma dan optimasi merupakan penelitian yang baru dilakukan. Dimana penelitian sebelumnya yang dilakukan oleh (Herliansyah et al., 2021) hanya menggunakan satu algoritma *Naive Bayes* dan tidak menerapkan optimasi *bagging* untuk klasifikasi stunting dengan mendapat hasil akurasi yang cukup rendah sebesar 64,02%. Maka dari itu pada penelitian ini menerapkan teknik *ensemble* dengan metode *bagging* untuk menaikkan akurasi, presisi, recall, dan f1-score dari model klasifikasi yang digunakan. Meskipun ada upaya dari penelitian terdahulu yang telah mencoba mengklasifikasikan stunting pada balita dengan berbagai model klasifikasi yang mendapat akurasi tertinggi, tetapi penelitian tentang prediksi stunting dengan menerapkan teknik *ensemble* dari *K-Nearest Neighbor*, *Support Vector Machine*, *Naive Bayes* menggunakan metode optimasi *Bagging* belum pernah ada yang melakukan. Dengan penggunaan optimasi *bagging*, hasil akurasi, presisi, recall, dan f1-score dapat meningkat, sehingga dengan mendapat hasil yang besar model dapat mengklasifikasi data stunting atau tidak stunting dengan akurat. Hasil ini memiliki dampak yang positif dalam konteks pencegahan stunting, karena model yang lebih akurat dapat membantu identifikasi balita yang berisiko mengalami stunting dengan lebih tepat. Dengan demikian, tindakan pencegahan dan upaya yang sesuai dapat diterapkan lebih efektif, memberikan dampak yang lebih besar dalam mengatasi masalah stunting pada balita.

SIMPULAN

Pada penelitian ini, gabungan tiga algoritma *K-Nearest Neighbor*, *Support Vector Machine*, dan *Naive Bayes* dengan metode *Bagging* menghasilkan akurasi 89,77%, presisi 95,57%, recall 85,45%, dan f1-score 90,27% tertinggi. Hal ini menunjukkan bahwa penerapan *Bagging* sebagai teknik *ensemble* mampu signifikan meningkatkan akurasi prediksi stunting pada balita. Penelitian ini inovatif karena menggabungkan tiga algoritma dan menerapkan metode optimasi *bagging* yang belum pernah dilakukan sebelumnya, memberikan potensi besar dalam upaya pencegahan stunting pada balita dengan identifikasi yang lebih akurat.

REFERENSI

- Armansyah, A., & Ramli, R. K. (2022). Model Prediksi Kelulusan Mahasiswa Tepat Waktu dengan Metode Naive Bayes. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 1-10. <https://doi.org/10.29408/edumatic.v6i1.4789>
- Efriadi, D., Rahmaddeni, R., Agustin, A., & Junadhi, J. (2022). Prediksi Penambahan Piutang Iuran Jaminan Sosial Ketenagakerjaan menggunakan Algoritma K-Nearest Neighbor. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 49-57. <https://doi.org/10.29408/edumatic.v6i1.5255>
- Ginting, K. P., & Pandiangan, A. (2019). Tingkat Kecerdasan Intelegensi Anak Stunting. *Jurnal Penelitian Perawat Profesional*, 1(1), 47–52. <https://doi.org/10.37287/jppp.v1i1.25>
- Givari, M. R., Sulaeman, M. R., & Umaidah, Y. (2022). Perbandingan Algoritma SVM, Random Forest Dan XGBoost Untuk Penentuan Persetujuan Pengajuan Kredit. *Nuansa Informatika*, 16(1), 141–149. <https://doi.org/10.25134/nuansa.v16i1.5406>
- Herliansyah, V. R. Latuconsina, & A. Dinimaharawati. (2021). Prediksi Stunting Pada Balita Dengan Menggunakan Algoritma Klasifikasi Naive Bayes. *E-Proceeding of Engineering*, 8(5), 6643–6649.
- Hermanto, H., Mustopa, A., & Kuntoro, A. Y. (2020). Algoritma Klasifikasi Naive Bayes Dan Support Vector Machine Dalam Layanan Komplain Mahasiswa. *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, 5(2), 211–220. <https://doi.org/10.33480/jitk.v5i2.1181>
- Hidayat, W., Utami, E., Iskandar, A. F., Hartanto, A. D., & Prasetio, A. B. (2021). Perbandingan Performansi Model pada Algoritma K-NN terhadap Klasifikasi Berita Fakta Hoaks Tentang Covid-19. *Edumatic: Jurnal Pendidikan Informatika*, 5(2), 167-176. <https://doi.org/10.29408/edumatic.v5i2.3664>
- Hovi, H. S. W., Id Hadiana, A., & Rakhmat Umbara, F. (2022). Prediksi Penyakit Diabetes Menggunakan Algoritma Support Vector Machine (SVM). *Informatics and Digital Expert (INDEX)*, 4(1), 40–45. <https://doi.org/10.36423/index.v4i1.895>
- Islam, H. I., Mulyadien, M. K., & Enri, U. (2022). Penerapan Algoritma C4.5 dalam Klasifikasi Status Gizi Balita. *Jurnal Ilmiah Wahana Pendidikan*, 8(10), 116–125.
- Kemendes. (2022). Hasil Survei Status Gizi Indonesia (SSGI) 2022. *Kemendes*, 1–7.
- Matdoan, M. Y., Matdoan, U. A., & Saleh Far-Far, M. (2022). Algoritma K-Means Untuk Klasifikasi Provinsi di Indonesia Berdasarkan Paket Pelayanan Stunting. *PANRITA Journal of Science, Technology, and Arts*, 1(2), 41–46.
- Nirmalasari, N. O. (2020). Stunting Pada Anak : Penyebab dan Faktor Risiko Stunting di Indonesia. *Qawwam: Journal For Gender Mainstreaming*, 14(1), 19–28.
- Prayoga, P. R., Purnawansyah, P., Hasanuddin, T., & Darwis, H. (2023). Klasifikasi Daun Herbal Menggunakan K-Nearest Neighbor dan Support Vector Machine dengan Fitur Fourier Descriptor. *Edumatic: Jurnal Pendidikan Informatika*, 7(1), 160–168. <https://doi.org/10.29408/edumatic.v7i1.17521>
- Purwani, F., Wahyudi, R. T., & Jaya, I. D. (2022). Penerapan Algoritma K-Nearest Neighbor dengan Euclidean Distance untuk Menentukan Kelompok Uang Kuliah Tunggal

- Mahasiswa. *Edumatic: Jurnal Pendidikan Informatika*, 6(2), 344-353. <https://doi.org/10.29408/edumatic.v6i2.6547>
- Rahayu, A. S., Fauzi, A., & Rahmat, R. (2022). Komparasi Algoritma Naïve Bayes Dan Support Vector Machine (SVM) Pada Analisis Sentimen Spotify. *Jurnal Sistem Komputer Dan Informatika (JSON)*, 4(2), 349. <https://doi.org/10.30865/json.v4i2.5398>
- Rahmadhita, K. (2020). Permasalahan Stunting dan Pencegahannya. *Jurnal Ilmiah Kesehatan Sandi Husada*, 11(1), 225–229. <https://doi.org/10.35816/jiskh.v11i1.253>
- Ramadhan, N. G., & Khoirunnisa, A. (2021). Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine. *Jurnal Media Informatika Budidarma*, 5(4), 1580-1584. <https://doi.org/10.30865/mib.v5i4.3347>
- Raysyah, S. R., Veri Arinal, & Dadang Iskandar Mulyana. (2021). Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode Knn Dan Pca. *JSiI (Jurnal Sistem Informasi)*, 8(2), 88–95. <https://doi.org/10.30656/jsii.v8i2.3638>
- Saputra, A. Y., & Primadasa, Y. (2018). Penerapan Teknik Klasifikasi Untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritma K-Nearest Neighbor. *Techno. Com*, 17(4), 395-403. <https://doi.org/10.33633/tc.v17i4.1864>
- Saputri, N. D., Khalid, K., & Rolliawati, D. (2022). Comparison of Bagging and Adaboost Methods on C4. 5 Algorithm for Stroke Prediction. *Sistemasi: Jurnal Sistem Informasi*, 11(3), 567-577. <https://doi.org/10.32520/stmsi.v11i3.1684>
- Sari, V. R., Firdausi, F., & Azhar, Y. (2020). Perbandingan Prediksi Kualitas Kopi Arabika dengan Menggunakan Algoritma SGD, Random Forest dan Naive Bayes. *Edumatic: Jurnal Pendidikan Informatika*, 4(2), 1-9. <https://doi.org/10.29408/edumatic.v4i2.2202>
- Setiawan, R., & Triayudi, A. (2022). Klasifikasi Status Gizi Balita Menggunakan Naïve Bayes dan K-Nearest Neighbor Berbasis Web. *Jurnal Media Informatika Budidarma*, 6(2), 777-785. <https://doi.org/10.30865/mib.v6i2.3566>
- Susanti, S., Sari, A. A., Anam, M. K., Jamaris, M., & Hamdani, H. (2022). Sistem Prediksi Keuntungan Influencer Pengguna E-Commerce Shopee Affiliates menggunakan Metode Naïve Bayes. *Edumatic: Jurnal Pendidikan Informatika*, 6(2), 394-403. <https://doi.org/10.29408/edumatic.v6i2.6787>
- Wahyudin, W. C., Hana, F. M., & Prihandono, A. (2023). Prediksi Stunting Pada Balita Di Rumah Sakit Kota Semarang Menggunakan Naive Bayes. *Jurnal Ilmu Komputer dan Matematika*, 4(1), 32–36.
- Wijaya, H. D., & Dwiasnati, S. (2020). Implementasi Data Mining dengan Algoritma Naïve Bayes pada Penjualan Obat. *Jurnal Informatika*, 7(1), 1–7. <https://doi.org/10.31311/ji.v7i1.6203>
- Yunial, A. H. (2020). Analisis Optimasi Algoritma Klasifikasi Support Vector Machine, Decision Trees, dan Neural Network Menggunakan Adaboost dan Bagging. *Jurnal Informatika Universitas Pamulang*, 5(3), 247. <https://doi.org/10.32493/informatika.v5i3.6609>