

Komparasi Algoritma *Machine Learning* untuk Prediksi Risiko *Stunting* Berbasis Data P3KE Provinsi Gorontalo

Siti Andini Utiahman^{1,*}, Andi Mulawati Mas Pratama¹, Satriadi D. Ali²

¹Sistem Informasi, Fakultas Ilmu Komputer, Universitas Ichsan Gorontalo, Gorontalo, Indonesia

²Informatika, Fakultas Ilmu Komputer, Universitas Ichsan Gorontalo Utara, Gorontalo, Indonesia

Email: ^{1,*}siti_andini@unisan.ac.id, ²mulapratama@gmail.com, ³ady.stmik@gmail.com

Email Penulis Korespondensi: siti_andini@unisan.ac.id*

Submitted: 17/05/2026; Accepted: 18/06/2026; Published: 30/06/2026

Abstrak– *Stunting* merupakan permasalahan kesehatan masyarakat yang serius di Provinsi Gorontalo dengan prevalensi mencapai 23,8% pada tahun 2024, masih jauh di atas target nasional sebesar 14%. Penelitian ini bertujuan mengembangkan model prediksi risiko *Stunting* menggunakan pendekatan *Machine learning* berbasis Data Pensasaran Percepatan Penghapusan Kemiskinan Ekstrem (P3KE). Dataset yang digunakan mencakup 87.902 rumah tangga di Provinsi Gorontalo. Lima algoritma dibandingkan, yaitu *Decision Tree*, *K-Nearest Neighbors*, *Support Vector Machine*, *Random Forest*, dan *XGBoost*. Evaluasi dilakukan pada dua skenario, yaitu tanpa dan dengan penerapan *Synthetic Minority Oversampling Technique (SMOTE)*. Optimasi hyperparameter dilakukan menggunakan *RandomizedSearchCV* dengan 3-fold cross-validation. Hasil eksperimen menunjukkan bahwa *XGBoost* tanpa *SMOTE* merupakan model terbaik dengan *F1-score* sebesar 90,39%, *accuracy* 90,48%, *precision* 99,91%, *recall* 82,52%, dan *AUC-ROC* 94,09%. Penerapan *SMOTE* tidak memberikan peningkatan kinerja yang signifikan, dengan selisih *F1-score* yang hanya mencapai 0,0067 poin. Sebaliknya, teknik tersebut meningkatkan waktu komputasi hingga 65%. Temuan ini mengonfirmasi bahwa distribusi kelas yang relatif seimbang, yaitu 54,2% : 45,8%, tidak memerlukan proses *oversampling*. Analisis *SHAP* mengidentifikasi *BAB: Milik sendiri* dan *Indeks Sanitasi* sebagai fitur yang paling berpengaruh dalam prediksi risiko *Stunting*. Hasil penelitian menunjukkan bahwa data registrasi sosial P3KE tanpa melibatkan data antropometri tetap mampu menghasilkan model prediksi risiko *Stunting* dengan kinerja yang kompetitif, yang ditunjukkan oleh nilai *F1-score* di atas 90%.

Kata Kunci: *Machine learning*; Prediksi *Stunting*; *XGBoost*; *SHAP*; Data P3KE; *SMOTE*; Gorontalo

Abstract– *Stunting* remains a serious public health problem in Gorontalo Province, with a prevalence of 23.8% in 2024, which remains substantially higher than the national target of 14%. This study aims to develop a *Stunting* risk prediction model using a *Machine learning* approach based on the Extreme Poverty Eradication Acceleration Targeting Data (Data Pensasaran Percepatan Penghapusan Kemiskinan Ekstrem, P3KE). The dataset comprises 87,902 households in Gorontalo Province. Five algorithms were evaluated, namely *Decision Tree*, *K-Nearest Neighbors*, *Support Vector Machine*, *Random Forest*, and *XGBoost*. The evaluation was conducted under two scenarios: with and without the *Synthetic Minority Oversampling Technique (SMOTE)*. Hyperparameter optimization was performed using *RandomizedSearchCV* with 3-fold cross-validation. The experimental results indicate that *XGBoost* without *SMOTE* achieved the best performance, with an *F1-score* of 90.39%, *accuracy* of 90.48%, *precision* of 99.91%, *recall* of 82.52%, and an *AUC-ROC* of 94.09%. The application of *SMOTE* did not yield a significant performance improvement, with a difference in *F1-score* of only 0.0067 points. Instead, it increased computation time by up to 65%. These findings confirm that a near-balanced class distribution (54.2%:45.8%) does not require *oversampling*. *SHAP* analysis identified ownership of adequate sanitation facilities and the Sanitation Index as the most influential features in predicting *Stunting* risk. The results further indicate that P3KE social registry data, without incorporating anthropometric measurements, can effectively support *Stunting* risk prediction, achieving a competitive *F1-score* exceeding 90%.

Keywords: *Machine learning*; *Stunting* Prediction; *XGBoost*; *SHAP*; P3KE Data; *SMOTE*; Gorontalo

1. PENDAHULUAN

Stunting merupakan kondisi gagal tumbuh pada anak balita yang ditandai dengan tinggi badan di bawah standar usia akibat kekurangan gizi kronis dan infeksi berulang [1]. Kondisi ini tidak hanya berdampak pada pertumbuhan fisik, tetapi juga pada perkembangan kognitif, produktivitas jangka panjang, dan kualitas sumber daya manusia suatu bangsa [2]. Berdasarkan data Organisasi Kesehatan Dunia (WHO), lebih dari 149 juta anak di bawah usia lima tahun di seluruh dunia mengalami *Stunting*, dengan konsentrasi terbesar di kawasan Asia Selatan dan Afrika Sub-Sahara [3].

Indonesia termasuk negara dengan beban *Stunting* tinggi di Asia Tenggara. Berdasarkan Survei Status Gizi Indonesia (SSGI) 2021 yang dirilis Kementerian Kesehatan RI, prevalensi *Stunting* nasional tercatat 24,4%, turun dari 27,7% pada 2019. Meskipun menunjukkan tren penurunan konsisten, angka tersebut masih melebihi ambang batas kritis WHO sebesar 20%. Pada 2024, prevalensi *Stunting* Indonesia berhasil turun menjadi 19,8%, pertama kalinya berada di bawah ambang batas WHO [4]. Pemerintah Indonesia telah menetapkan target penurunan *Stunting* menjadi 14% pada tahun 2024 melalui Peraturan Presiden Nomor 72 Tahun 2021 tentang Percepatan Penurunan *Stunting* [5]. Namun berdasarkan SSGI 2024, angka nasional masih berada di kisaran 19,8 % [6], menunjukkan bahwa strategi intervensi yang lebih tepat sasaran dan berbasis data masih sangat diperlukan.

Provinsi Gorontalo menjadi salah satu daerah yang menghadapi tantangan serius dalam penanganan *Stunting*. Survei Kesehatan Indonesia (SKI) 2023 mencatat prevalensi *Stunting* di Gorontalo sebesar 26,9%, melampaui rata-rata nasional [7]. Pada tahun 2024, angka ini turun menjadi 23,8% berdasarkan SSGI 2024, namun masih jauh dari target nasional 14%. Yang lebih mengkhawatirkan adalah tren yang tidak stabil: dari 29,0 % pada 2021, turun ke 23,8% pada 2022, naik kembali ke 26,9% pada 2023, kemudian turun ke 23,8% pada 2024. Ketidakstabilan *tren* ini mengindikasikan bahwa intervensi yang dilakukan belum sepenuhnya menasar akar permasalahan. Disparitas antarwilayah juga sangat signifikan: Kabupaten Gorontalo mencatat prevalensi tertinggi sebesar 34,7%, sementara Boalemo relatif lebih rendah di angka 16,0% [6]. Pemerintah Provinsi Gorontalo telah menempuh berbagai program lintas sektor, mulai dari Bapak Asuh Anak *Stunting* (BAAS), Gerakan Orang Tua Asuh Cegah *Stunting* (GENTING) oleh BKKBN, hingga program unggulan *Bele Mo'osehati* yang merupakan kolaborasi IDI, BKKBN, dan Dinas Kesehatan [8], [9]. Namun efektivitas program-program tersebut sangat bergantung pada ketepatan identifikasi sasaran. Tanpa sistem deteksi dini yang akurat dan skalabel, intervensi berisiko tidak menjangkau keluarga yang paling membutuhkan.

Salah satu instrumen pendataan yang dimiliki pemerintah adalah Data Pensasaran Percepatan Penghapusan Kemiskinan Ekstrem (P3KE) yang mencakup profil sosial-ekonomi rumah tangga miskin secara komprehensif. Data P3KE Provinsi Gorontalo untuk segmen rumah tangga dengan daya listrik terpasang di bawah atau sama dengan 900 watt mencakup 119.664 baris data dengan 40 atribut, termasuk label "Resiko *Stunting*" yang telah ditetapkan oleh Kementerian Sosial. Ketersediaan label ini menjadikan data P3KE sebagai kandidat yang sangat potensial untuk pengembangan model prediksi risiko *Stunting* berbasis *Machine learning*.

Pendekatan *Machine learning* telah terbukti efektif untuk prediksi dan klasifikasi *Stunting*. Tinjauan literatur sistematis menempatkan *XGBoost* dan *Random Forest* sebagai algoritma yang paling banyak digunakan, dengan Indonesia sebagai negara penghasil penelitian *Stunting* berbasis *Machine learning* terbanyak [10], [11]. Sejumlah studi memanfaatkan data survei demografi dan kesehatan untuk membangun model prediksi berbasis faktor sosiodemografi, di mana algoritma berbasis *ensemble* seperti *Random Forest* dan *Gradient Boosting* secara konsisten menunjukkan performa terbaik dengan akurasi 80-98% [12], [13], [14]. Beberapa penelitian terbaru juga mulai mengintegrasikan *SHAP* untuk menjelaskan kontribusi fitur, dan secara konsisten mengidentifikasi faktor sosial-ekonomi seperti sanitasi, kemiskinan, dan pendidikan orang tua sebagai prediktor *Stunting* yang dominan [15], [16], [17].

Hasdina *et al.* mengusulkan kerangka *hybrid Machine learning* yang mengintegrasikan klasifikasi, prediksi, dan klusterisasi, dengan SVM kernel RBF mencapai akurasi 91,3% [18]. Novalina *et al.* melakukan *benchmarking* menggunakan kerangka CRISP-DM pada data IFLS Indonesia dengan membandingkan teknik *random undersampling* dan *SMOTE* [19]. Wicaksono *et al.* mengombinasikan *XGBoost* dan *SHAP* pada 40.071 data anak serta menunjukkan kemampuan model dalam mengidentifikasi determinan *Stunting* yang relevan untuk perumusan kebijakan [20]. Islam *et al.* menemukan bahwa fasilitas sanitasi, tingkat kekayaan, dan pendidikan ayah merupakan prediktor utama *Stunting* [21]. Maulana dan Subhiyanto membandingkan *Random Forest*, SVM, dan *XGBoost* dengan *SMOTE*, serta menemukan bahwa *XGBoost* memberikan kinerja terbaik dengan *F1-score* 88,57% [22]. Saragih *et al.* melaporkan bahwa *Random Forest* menghasilkan galat terendah dalam prediksi prevalensi *Stunting* dibandingkan SVR dan *Decision Tree* [23]. Pratama *et al.* menunjukkan bahwa *Random Forest Regression* mencapai *success rate* sebesar 90,537% [24]. Sementara itu, Handayani dan Sibuea menggunakan *K-Means* untuk pemetaan wilayah rawan *Stunting*, namun pendekatan *unsupervised* menghasilkan informasi yang lebih terbatas dibandingkan pendekatan *supervised* [25]. Joharini dan Subekti menekankan bahwa variabel sosial-ekonomi seperti sanitasi, pendapatan, dan pendidikan sering kali sulit diperoleh dari fasilitas kesehatan primer, padahal informasi tersebut tersedia secara lengkap dalam data P3KE [26]. Syahril *et al.* melakukan klasifikasi *Stunting* menggunakan 11 algoritma berbasis data antropometri, dengan SVM-RBF mencapai akurasi tertinggi sebesar 78% [27]. Yunus *et al.* membandingkan *Naive Bayes* dan *Decision Tree C4.5* pada 224 data balita, dengan *Decision Tree C4.5* mencapai akurasi 87% [28]. Temuan-temuan tersebut menunjukkan bahwa penelitian *Stunting* berbasis *Machine learning* masih didominasi oleh penggunaan data klinis dengan jumlah sampel yang relatif terbatas serta belum banyak mengintegrasikan dimensi sosial-ekonomi rumah tangga.

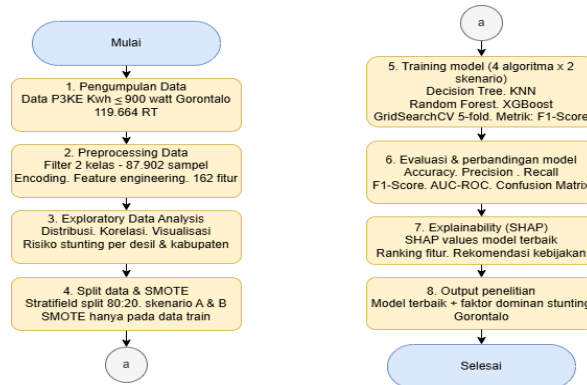
Meskipun penelitian prediksi *Stunting* berkembang pesat, masih terdapat beberapa kesenjangan. Pertama, sebagian besar penelitian bergantung pada data antropometri atau survei nasional dan belum mengintegrasikan kondisi sosial-ekonomi rumah tangga secara komprehensif. Kedua, pemanfaatan data registrasi sosial P3KE sebagai basis prediksi risiko *Stunting* masih sangat terbatas, padahal data ini mencakup kondisi hunian, akses sanitasi, penerimaan bantuan sosial, dan karakteristik kepala keluarga dalam satu dataset berskala besar. Ketiga, penelitian *Stunting* yang berfokus pada Provinsi Gorontalo masih terbatas baik dari sisi cakupan data maupun kedalaman analisis. Keempat, penerapan *SHAP* untuk menghasilkan rekomendasi kebijakan lokal yang dapat ditindaklanjuti hingga tingkat kabupaten dan kecamatan belum banyak dilakukan di Gorontalo.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan mengevaluasi performa berbagai algoritma *Machine learning* untuk prediksi risiko *Stunting* berbasis data P3KE, menganalisis pengaruh *SMOTE*, serta mengidentifikasi faktor sosial-ekonomi dominan menggunakan *SHAP* guna mendukung kebijakan intervensi berbasis data di Provinsi Gorontalo.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini menggunakan kerangka CRISP-DM yang meliputi pengumpulan dan pemahaman data, preprocessing, EDA, pembagian data, penanganan ketidakseimbangan kelas, pembangunan dan evaluasi model, serta analisis explainability dengan SHAP. Alur penelitian ditampilkan pada Gambar 1.



Gambar 1. Alur Tahapan Penelitian

2.2 Dataset

Data yang digunakan dalam penelitian ini adalah Data Pensasaran Percepatan Penghapusan Kemiskinan Ekstrem (P3KE) Provinsi Gorontalo, khususnya rumah tangga dengan daya listrik terpasang ≤ 900 watt. Data diperoleh dari Dinas PM, ESDM, dan Transmigrasi Provinsi Gorontalo, terdiri atas 119.664 *record* dan 40 atribut yang mencakup identitas, lokasi, kondisi hunian, akses layanan dasar, karakteristik kepala keluarga, serta status penerimaan bantuan sosial.

Kolom target Resiko *Stunting* merupakan penanda penyasaran administratif (*administrative targeting label*) yang ditetapkan dalam program P3KE oleh Kementerian Sosial, bukan hasil diagnosis klinis atau pengukuran antropometri. Label ini dibentuk berdasarkan kondisi sosial-ekonomi rumah tangga, sebagaimana tercermin dari struktur data yang seluruhnya terdiri atas variabel sosial-ekonomi tanpa atribut antropometri anak. Oleh karena itu, penelitian ini memodelkan risiko *Stunting* pada tingkat rumah tangga berdasarkan determinan sosial-ekonomi, bukan untuk menggantikan diagnosis klinis pada individu. Model yang dihasilkan diharapkan dapat berfungsi sebagai alat bantu penyasaran yang efisien dan mudah diinterpretasikan untuk mengidentifikasi rumah tangga prioritas.

Variabel target terdiri atas tiga kategori, yaitu Beresiko *Stunting* (47.687 sampel; 39,8%), Tidak Beresiko *Stunting* (40.215 sampel; 33,6%), dan Bukan Target Sasaran (31.762 sampel; 26,6%). Kategori Bukan Target Sasaran merupakan klasifikasi administratif di luar cakupan penyasaran sehingga tidak relevan untuk pemodelan klasifikasi biner dan dikeluarkan dari analisis. Setelah proses penyaringan, jumlah data yang digunakan sebanyak 87.902 rumah tangga. Kelompok fitur yang digunakan disajikan pada Tabel 1.

Tabel 1. Kelompok fitur data P3KE yang digunakan dalam penelitian

Kelompok Fitur	Contoh Variabel	Jumlah Fitur
Sanitasi & akses air	Sumber air minum, fasilitas sanitasi, bahan bakar memasak	3
Kondisi fisik hunian	Jenis dan kualitas atap, dinding, lantai	6
Profil kepala keluarga	Pendidikan, pekerjaan, jenis kelamin, usia	4
Status ekonomi & bansos	Desil kesejahteraan, penerima PKH/BPNT/BST/dll	8
Lokasi	Kabupaten/Kota, Kecamatan	2

2.3 Alat dan Bahan Penelitian

Penelitian ini dilaksanakan menggunakan infrastruktur komputasi dan perangkat lunak sebagaimana disajikan pada Tabel 2.

Tabel 2. Spesifikasi alat dan bahan penelitian

Komponen	Spesifikasi/Keterangan
Platform	Kaggle Notebooks (Cloud Computing)
Sistem Operasi	Linux Ubuntu
Prosesor (CPU)	Intel Xeon Processor

GPU	NVIDIA Tesla P100 / T4 (tergantung sesi <i>Kaggle</i>)
RAM	13–16 GB
Storage	±73 GB (<i>temporary storage</i>)
Bahasa Pemrograman	Python 3
Library Machine learning	Scikit-learn, XGBoost, Imbalanced-learn
Library Explainability	SHAP (<i>TreeExplainer</i>)
Library Data & Visualisasi	NumPy, Pandas, Matplotlib, Seaborn
Metode Tuning	RandomizedSearchCV dengan 3-Fold Stratified CV
Dataset	Data P3KE KWH ≤900 Watt Provinsi Gorontalo
Ukuran Dataset	119.664 baris, 40 kolom, format .xlsx

2.4 Preprocessing Data

2.4.1 Seleksi dan pembersihan data

Dataset di *filter* untuk mempertahankan hanya dua kelas target yang relevan, menghasilkan 87.902 sampel. Kolom yang tidak informatif untuk pemodelan meliputi nama kepala keluarga, NIK, alamat, ID keluarga, dan kolom dengan nilai konstan seperti Daya Listrik Terpasang dieliminasi. Nilai "Kosong" pada fitur kategorik diperlakukan sebagai kategori tersendiri dan tidak dibuang untuk menjaga integritas data.

2.4.2 Rekayasa Fitur (*Feature Engineering*)

Tiga fitur baru dikonstruksi untuk merepresentasikan dimensi penting yang tidak tersedia secara eksplisit: (a) Indeks Kualitas Hunian: skor gabungan kualitas atap, dinding, dan lantai dengan rentang 0-6; (b) Indeks Sanitasi: kombinasi *biner* tiga indikator WASH (BAB: Milik sendiri, akses air bersih, dan penggunaan bahan bakar bersih) dengan rentang 0-3; dan (c) Jumlah Bansos: total program bantuan sosial yang diterima suatu rumah tangga dari tujuh program tersedia, dengan rentang 0-7. Usia kepala keluarga juga diekstraksi dari kolom tanggal lahir menggunakan tahun referensi 2022.

2.4.3 Encoding dan Normalisasi

Fitur ordinal, seperti desil kesejahteraan dan tingkat pendidikan kepala keluarga, dikonversi menggunakan *label encoding* sesuai urutan hierarkinya. Sementara itu, fitur nominal, seperti jenis pekerjaan, jenis atap, sumber air, dan lokasi, dikonversi menggunakan *one-hot encoding*, sehingga jumlah fitur meningkat dari 40 atribut menjadi 162 fitur. Seluruh fitur kemudian dinormalisasi menggunakan *StandardScaler* ($mean = 0$, standar deviasi = 1).

Meskipun normalisasi tidak memengaruhi model berbasis pohon, seperti *Decision Tree*, *Random Forest*, dan *XGBoost*, langkah ini tetap diterapkan untuk mendukung algoritma yang sensitif terhadap skala fitur, yaitu *K-Nearest Neighbors* dan *Support Vector Machine*. Selain itu, penerapan pra-proses yang seragam memastikan *fair comparison* antar algoritma dalam satu *pipeline*. Untuk mencegah *data leakage*, *StandardScaler* di-fit hanya pada data *training* dan selanjutnya digunakan untuk mentransformasikan data *test*.

2.5 Pembagian Data dan SMOTE

Dataset dibagi menjadi data *training* (80%) dan data *test* (20%) menggunakan *stratified split*, menghasilkan 70.321 sampel *training* dan 17.581 sampel *test*. Meskipun distribusi kelas relatif seimbang (54,2%:45,8%), *SMOTE* tetap dievaluasi sebagai skenario pembandingan. Pengujian ini bertujuan memverifikasi manfaat *oversampling* pada kondisi ketidakseimbangan ringan sekaligus mengevaluasi *trade-off* antara peningkatan performa dan biaya komputasi. Dua skenario diterapkan, yaitu Skenario A menggunakan data *training* asli, sedangkan Skenario B menerapkan *SMOTE* hingga distribusi kelas menjadi seimbang (76.298 sampel). *SMOTE* menghasilkan sampel sintesis kelas minoritas melalui interpolasi linear antara sampel dan tetangga terdekatnya [29]. Untuk mencegah *data leakage*, *SMOTE* hanya diterapkan pada data *training*, sementara data *test* tetap dipertahankan dalam kondisi asli. Dengan lima algoritma dan dua skenario, total percobaan yang dilakukan adalah 10 model.

2.6 Algoritma Machine learning

Lima algoritma *Machine learning* dipilih dalam penelitian ini berdasarkan representasi komprehensif terhadap paradigma pembelajaran yang berbeda, yaitu *decision tree-based*, *instance-based*, *kernel-based*, *ensemble*, dan *gradient boosting*. Pemilihan kelima algoritma *Machine learning* dipilih berdasarkan tinjauan literatur penelitian *Stunting* sebelumnya dan relevansinya terhadap data tabular berskala besar.

2.6.1 Decision Tree (DT)

DT membangun model prediktif berupa struktur pohon hierarkis yang terdiri dari *node* internal (pengujian fitur), cabang (hasil pengujian), dan *node* daun (keputusan kelas). Pada setiap *node*, algoritma memilih fitur pemecah

terbaik berdasarkan kriteria impuritas [14]. Penelitian ini menggunakan kriteria entropi yang diformulasikan pada persamaan (1):

$$H(S) = - \sum_{i=1}^n p_i \log_2(p_i) \quad (1)$$

di mana (p_i) merupakan proporsi sampel yang termasuk dalam kelas ke- (i) pada simpul (S). Selanjutnya, *Information Gain* dihitung sebagai selisih antara nilai *entropy* sebelum dan sesudah proses *splitting* data. Algoritma *Decision Tree* digunakan sebagai model *baseline* karena mampu menghasilkan aturan keputusan yang transparan serta mudah diinterpretasikan oleh pemangku kepentingan *non-teknis*. Berdasarkan hasil proses *hyperparameter tuning*, konfigurasi terbaik diperoleh pada parameter ($max_depth = 5$) dan ($criterion = entropy$). Pengaturan ini menghasilkan pohon keputusan dengan kedalaman yang terbatas sehingga dapat mengurangi risiko *overfitting* sekaligus mempertahankan kemampuan generalisasi model.

2.6.2 K-Nearest Neighbors (KNN)

KNN merupakan algoritma berbasis *instance*, tidak membangun model eksplisit selama fase *training*. Klasifikasi dilakukan dengan menghitung jarak antara sampel baru dan seluruh data *training*, kemudian mengambil K tetangga dengan jarak terkecil. Jarak dihitung menggunakan metrik *Euclidean* pada persamaan (2) [30]:

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

Prediksi akhir ditentukan melalui mekanisme pembobotan berbasis *invers* jarak ($weights = distance$), sehingga tetangga yang memiliki jarak lebih dekat terhadap sampel uji memberikan kontribusi yang lebih besar dalam proses penentuan kelas. Berdasarkan hasil *hyperparameter tuning*, konfigurasi optimal diperoleh pada parameter ($n_neighbors = 7$) dan ($weights = distance$). Meskipun *K-Nearest Neighbor (KNN)* dikenal sebagai metode yang sederhana dan efektif untuk berbagai permasalahan klasifikasi, kompleksitas komputasi pada tahap prediksi sebesar $O(n \cdot d)$, dengan (n) menyatakan jumlah sampel pelatihan dan (d) menyatakan jumlah fitur, menyebabkan metode ini menjadi kurang efisien ketika diterapkan pada *dataset* berskala besar. Hal ini disebabkan karena setiap sampel uji harus dibandingkan dengan seluruh sampel pelatihan untuk menghitung jarak sebelum menentukan tetangga terdekat.

2.6.3 Support Vector Machine (SVM)

SVM mencari *hyperplane* optimal $wT x + b = 0$ yang memaksimalkan margin $2/\|w\|$ antara dua kelas dalam ruang fitur berdimensi tinggi. Untuk data yang tidak dapat dipisahkan secara linear, SVM menggunakan fungsi *kernel* untuk memetakan data ke ruang berdimensi lebih tinggi [15]. Penelitian ini menggunakan *kernel Radial Basis Function* dengan persamaan (3):

$$K(x_i, x_j) = \exp(-\gamma |x_i - x_j|^2) \quad (3)$$

di mana γ mengontrol lebar *kernel*. Parameter $C = 1$ yang diperoleh dari proses *tuning* mengontrol *trade-off* antara memaksimalkan *margin* dan meminimalkan kesalahan klasifikasi. Kompleksitas *training* sebesar $O(n^2)$ hingga $O(n^3)$ menjadi kendala utama pada *dataset* dengan 70.321 *sample training*, sehingga menghasilkan waktu *training* selama 2.421 detik pada skenario tanpa *SMOTE*.

2.6.4 Random Forest (RF)

RF merupakan algoritma *ensemble* yang membangun sejumlah pohon keputusan secara *independen* menggunakan dua mekanisme randomisasi: (1) *Bootstrap Aggregating (Bagging)*, setiap pohon dilatih pada *subset* acak data *training* dengan pengembalian; dan (2) *Random Feature Selection*, pada setiap pemecahan node, hanya subset acak \sqrt{f} fitur yang dipertimbangkan [13], [16]. Prediksi final dihasilkan melalui *majority voting* seperti pada persamaan (4):

$$y(x) = \text{mode } h_i(x)_{i=1}^T \quad (4)$$

di mana $h_i(x)$ adalah prediksi dari pohon ke- i dan T adalah jumlah total pohon. Nilai optimal yang diperoleh adalah $n_estimators = 200$ dan $max_depth = None$, yang mengindikasikan bahwa pohon dengan kedalaman yang tidak dibatasi memberikan performa yang lebih baik pada data P3KE yang kompleks. Mekanisme *bagging* secara *inheren* mampu mengurangi varians model dan membantu mencegah terjadinya *overfitting*.

2.6.5 XGBoost (Extreme Gradient Boosting)

XGBoost merupakan implementasi *gradient boosting* yang dioptimasi untuk efisiensi komputasi dan performa [20] [31]. Berbeda dengan *Random Forest* yang membangun pohon secara paralel, *XGBoost* membangun pohon secara sekuensial, di mana setiap pohon baru $F_i(x)$ ditambahkan untuk meminimalkan fungsi objektif, seperti ditunjukkan pada persamaan berikut (5):

$$L(\theta) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_k \lambda \Omega(f_k) \quad (5)$$

di mana l adalah fungsi *loss*, dan $\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|^2$ merupakan fungsi regularisasi L1/L2 yang digunakan untuk mencegah *overfitting*, dengan T menyatakan jumlah daun dan w menyatakan bobot daun. *XGBoost* menggunakan mekanisme *histogram-based split finding* yang mengelompokkan nilai fitur ke dalam b bin, sehingga kompleksitas pencarian *split* dapat direduksi dari $O(n \cdot f)$ menjadi $O(b \cdot f)$, dengan $b \ll n$. Nilai optimal yang diperoleh adalah $max_depth = 5$, $learning_rate = 0,1$, dan $subsample = 0,8$ (Skenario A) serta 1,0 (Skenario B). Perbedaan nilai *subsample* antara kedua skenario mengindikasikan bahwa keberadaan sampel sintesis hasil *SMOTE* memengaruhi karakteristik distribusi data yang terdeteksi selama proses *tuning* otomatis.

2.7 Hyperparameter Tuning

Pencarian *hyperparameter* optimal dilakukan menggunakan *RandomizedSearchCV* dengan 5 iterasi, *3-fold stratified cross-validation*, metrik optimasi *F1-score*, dan paralelisasi penuh ($n_jobs = -1$). *RandomizedSearchCV* dipilih karena lebih efisien secara komputasi dibandingkan *GridSearchCV*, terutama pada dataset berskala besar, tanpa menyebabkan penurunan kualitas hasil yang signifikan.

Ruang pencarian *hyperparameter* dibatasi pada rentang nilai yang umum digunakan dalam literatur untuk menjaga efisiensi komputasi, mengingat ukuran dataset yang besar (70.321 *sampel training* dan 162 fitur) serta total 10 konfigurasi model yang dievaluasi. Kombinasi ruang pencarian yang ringkas, lima iterasi pencarian acak, dan validasi silang tiga lipat memungkinkan perbandingan yang adil antar algoritma dengan waktu komputasi yang tetap terkendali. Ruang pencarian dan nilai *hyperparameter* optimal untuk masing-masing algoritma disajikan pada Tabel 3.

Tabel 3. Ruang pencarian *hyperparameter*

Algoritma	Hyperparameter	Ruang Pencarian	Nilai Opt (A)	Nilai Opt (B)
DT	<i>max_depth</i>	[5, 10]	5	5
	<i>min_samples_split</i>	[2, 5]	5	5
	<i>criterion</i>	[<i>gini</i> , <i>entropy</i>]	<i>entropy</i>	<i>entropy</i>
KNN	<i>n_neighbors</i>	[3, 5, 7]	7	7
	<i>weights</i>	[<i>uniform</i> , <i>distance</i>]	<i>distance</i>	<i>distance</i>
SVM	<i>C</i>	[1, 10]	1	1
	<i>kernel</i>	[<i>rbf</i>]	<i>rbf</i>	<i>rbf</i>
RF	<i>n_estimators</i>	[100, 200]	200	200
	<i>max_depth</i>	[10, <i>None</i>]	<i>None</i>	<i>None</i>
	<i>min_samples_split</i>	[2, 5]	5	5
XGBoost	<i>n_estimators</i>	[100]	100	100
	<i>max_depth</i>	[3, 5]	5	5
	<i>learning_rate</i>	[0.1]	0.1	0.1
	<i>subsample</i>	[0.8, 1.0]	0.8	1.0

2.8 Metrik Evaluasi

Performa model dievaluasi menggunakan lima metrik, yaitu *accuracy*, *precision*, *recall*, *F1-score*, dan AUC-ROC. *Recall* menjadi perhatian utama karena dalam konteks kesehatan masyarakat, kesalahan *false negative* (gagal mengidentifikasi rumah tangga yang berisiko *Stunting*) memiliki konsekuensi yang lebih besar dibandingkan *false positive* (mengklasifikasikan rumah tangga tidak berisiko sebagai berisiko). Persamaan (6)-(9) menunjukkan formulasi metrik evaluasi yang digunakan [14], [15], [16]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{6}$$

$$Precision = \frac{TP}{TP+FP} \tag{7}$$

$$Recall = \frac{TP}{TP+FN} \tag{8}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \tag{9}$$

dimana TP = *True Positive*, TN = *True Negative*, FP = *False Positive* dan FN= *False Negative*.

2.9 Explainability dengan SHAP

Analisis *explainability* menggunakan *SHAP* (*SHAPley Additive exPlanations*) berbasis teori permainan kooperatif *SHAPley* yang mendistribusikan kontribusi setiap fitur terhadap prediksi secara adil [21] [20]. *SHAP value* untuk fitur ke- i diformulasikan sebagai pada persamaan (10)

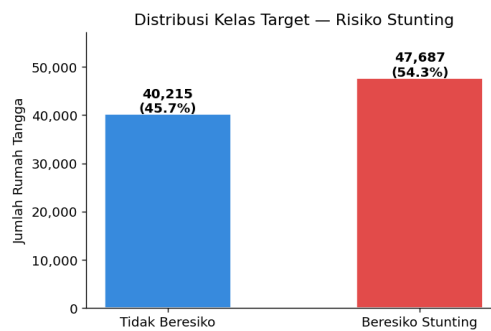
$$\phi_i \sum_{S \subseteq F \setminus i} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup i) - f(S)] \tag{10}$$

di mana ϕ_i adalah nilai kontribusi fitur ke- i (*SHAPley value*), S adalah subset fitur tanpa fitur i , F adalah himpunan seluruh fitur, $f(S)$ adalah performa model yang menggunakan subset fitur S , dan $f(S \cup i) - f(S)$ adalah kontribusi *marginal* fitur i terhadap subset S . Analisis dilakukan pada dua level, yaitu: (1) analisis global menggunakan *SHAP summary plot* dan *bar plot* untuk mengidentifikasi 20 fitur dengan nilai *Mean |SHAP|* tertinggi pada 1.000 sampel data uji; dan (2) *SHAP dependence plot* untuk tiga fitur teratas guna menganalisis hubungan antara nilai fitur dan kontribusinya terhadap prediksi model.

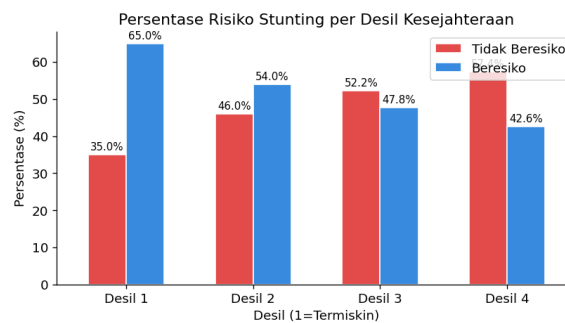
3. HASIL DAN PEMBAHASAN

3.1 Hasil Eksplorasi Data (EDA)

Analisis eksplorasi dilakukan terhadap 87.902 sampel yang telah melalui proses penyaringan. Distribusi kelas target pada Gambar 2 menunjukkan 47.687 sampel (54,2%) berkategori Beresiko *Stunting* dan 40.215 sampel (45,8%) berkategori Tidak Beresiko *Stunting*. Rasio kelas sebesar 1,19:1 mengindikasikan ketidakseimbangan yang relatif rendah (*mild imbalance*) sehingga distribusi data dapat dianggap mendekati seimbang.



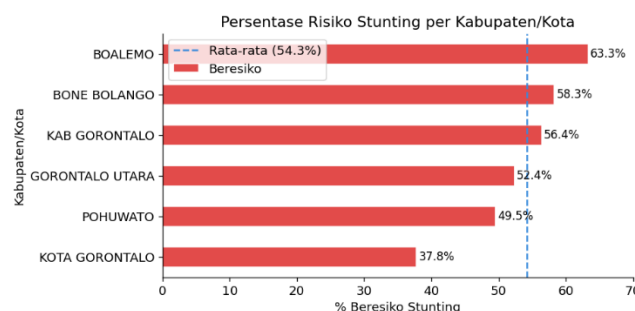
Gambar 2. Distribusi Kelas Target Risiko *Stunting*



Gambar 3. Presentasi Risiko *Stunting* Per Desil Kesejahteraan

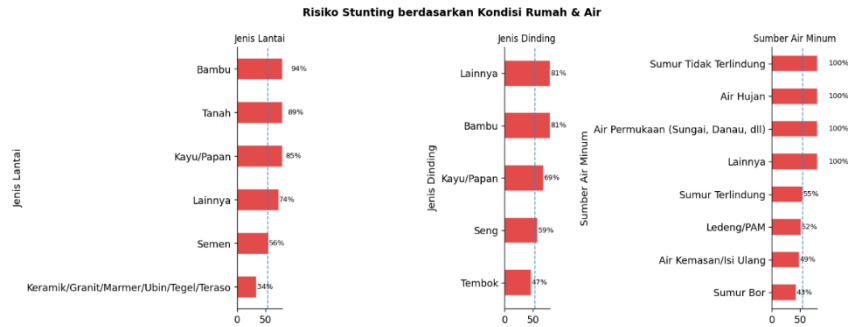
Gambar 3 menunjukkan persentase rumah tangga berisiko *Stunting* berdasarkan desil kesejahteraan. Terlihat *gradien* yang konsisten, di mana desil 1 (termiskin) memiliki proporsi risiko tertinggi, sedangkan desil 4 terendah. Temuan ini mengonfirmasi hubungan yang kuat antara kemiskinan multidimensi dan risiko *Stunting* dalam data P3KE.

Gambar 4 menyajikan distribusi rumah tangga berisiko *Stunting* per Kabupaten/Kota. Kabupaten Gorontalo memiliki proporsi tertinggi dibandingkan wilayah lainnya, sejalan dengan data SKI 2023 yang melaporkan prevalensi *Stunting* sebesar 34,7%. Temuan ini menunjukkan adanya variasi spasial yang signifikan dalam distribusi risiko *Stunting* di Gorontalo.



Gambar 4. Presentasi Risiko *Stunting* Per Kabupaten/Kota

Gambar 5 menunjukkan hubungan antara kondisi sanitasi, sumber air minum, dan risiko *Stunting*. Rumah tangga tanpa fasilitas sanitasi yang layak serta menggunakan sumber air tidak terlindungi memiliki proporsi risiko yang lebih tinggi dibandingkan rumah tangga dengan akses sanitasi dan air minum yang memadai. Temuan ini memberikan indikasi awal mengenai pentingnya faktor sanitasi dan kualitas air, yang selanjutnya dikonfirmasi melalui analisis *SHAP* pada Subbagian 3.6.



Gambar 5. Risiko *Stunting* berdasarkan kondisi sanitasi dan sumber air

3.2 Hasil *Preprocessing* dan *Feature Engineering*

Tahap *preprocessing* mentransformasi dataset dari 40 atribut menjadi 162 fitur siap model. Peningkatan jumlah fitur terutama berasal dari *one-hot encoding* pada variabel lokasi kecamatan (77 fitur) dan pekerjaan kepala keluarga (lebih dari 30 kategori). Pada Skenario B, penerapan *SMOTE* meningkatkan jumlah sampel *training* dari 70.321 menjadi 76.298 melalui penambahan 5.977 sampel sintesis pada kelas minoritas.

3.3 Hasil *Training* dan Analisis Kompleksitas Komputasi

Seluruh proses pelatihan model dilakukan pada platform *Kaggle* dengan spesifikasi *Intel Xeon Processor*, GPU *NVIDIA Tesla P100/T4*, dan RAM 13-16 GB. Tabel 4 menyajikan perbandingan nilai *CV F1-score* dan waktu pelatihan seluruh algoritma pada kedua skenario. Hasil tersebut menunjukkan adanya perbedaan kompleksitas komputasi yang signifikan antar algoritma.

Keterangan: A menunjukkan skenario tanpa *SMOTE*, sedangkan B menunjukkan skenario dengan *SMOTE*; det menyatakan satuan detik. Berdasarkan Tabel 4, *SVM* membutuhkan waktu pelatihan terlama, yaitu 2.421 detik pada Skenario A dan 3.063 detik pada Skenario B. Tingginya waktu komputasi ini disebabkan oleh kompleksitas pelatihan *SVM* yang berkisar antara $O(n^2)$ hingga $O(n^3)$, sehingga menjadi kendala utama pada dataset dengan 70.321 sampel *training*. Sebaliknya, *XGBoost* mencapai *CV F1-score* tertinggi (0,9015) dengan waktu pelatihan hanya 13 detik. Efisiensi tersebut diperoleh dari implementasi *histogram-based gradient boosting* yang menurunkan kompleksitas pencarian *split* dari $O(n.f)$ menjadi $O(b.f)$, dengan $(b < n)$.

Tabel 4. *CV F1-score* dan waktu *training* aktual

Algoritma	CV F1 (A)	Waktu A (det)	CV F1 (B)	Waktu B (det)
DT	0,9006	5	0,9006	8
KNN	0,8283	89	0,8186	110
SVM	0,8996	2421	0,8995	3063
RF	0,9016	71	0,9017	84
<i>XGBoost</i>	0,9015	13	0,9011	18

Tabel 4 menunjukkan bahwa *KNN* menghasilkan *CV F1-score* terendah (0,8283) sekaligus membutuhkan waktu pelatihan yang relatif tinggi (89 detik). Temuan ini sejalan dengan fenomena *curse of dimensionality*, di mana pada ruang fitur berdimensi tinggi (162 fitur), jarak antar sampel cenderung menjadi seragam sehingga kemampuan diskriminatif metode berbasis tetangga terdekat menurun. Sebaliknya, *DT*, *RF*, dan *XGBoost* menunjukkan kinerja yang lebih kompetitif (*CV F1-score* 0,9006-0,9016) karena model berbasis pohon mampu menangani data berdimensi tinggi melalui pemilihan fitur secara adaptif pada setiap *split*.

3.4 Hasil Evaluasi Model pada Data Test

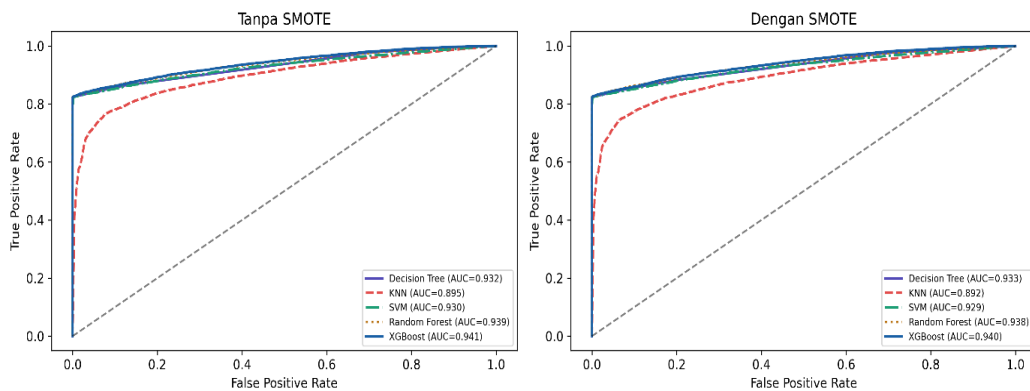
Evaluasi dilakukan pada 17.581 sampel data test yang tidak pernah digunakan selama *training* maupun *tuning*. Tabel 5 menyajikan hasil evaluasi lengkap semua model.

Tabel 5. Perbandingan performa semua model pada data test

Rank	Algoritma	Skenario	Accuracy	Precision	Recall	F1-score	AUC-ROC
------	-----------	----------	----------	-----------	--------	----------	---------

1	<i>XGBoost</i>	A	90,48%	99,91%	82,52%	90,39%	94,09%
2	<i>XGBoost</i>	B	90,47%	99,95%	82,47%	90,37%	94,01%
3	DT	B	90,45%	100,00%	82,40%	90,35%	93,25%
4	DT	A	90,45%	100,00%	82,40%	90,35%	93,20%
5	RF	B	90,39%	99,55%	82,67%	90,33%	93,77%
6	RF	A	90,35%	99,46%	82,66%	90,28%	93,87%
7	SVM	A	90,29%	99,39%	82,60%	90,22%	92,96%
8	SVM	B	90,28%	99,42%	82,56%	90,21%	92,93%
9	KNN	A	83,41%	89,70%	78,42%	83,68%	89,54%
10	KNN	B	83,15%	91,60%	75,90%	83,01%	89,23%

Berikut ini gambar Kurva ROC Semua Model sebagai berikut:



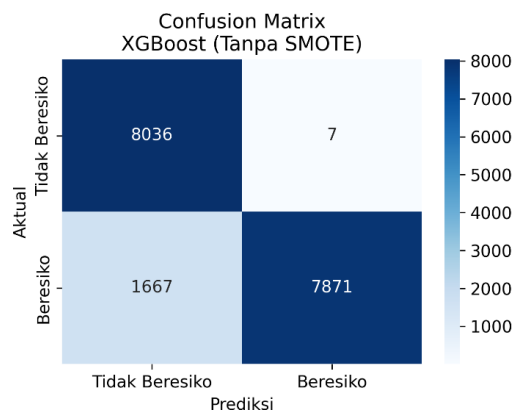
Gambar 6. Kurva ROC Semua Model

Tabel 5 menunjukkan bahwa DT, SVM, *Random Forest*, dan *XGBoost* mencapai performa yang relatif setara dengan *FI-score* pada kisaran 90,2%-90,4%, sedangkan KNN menghasilkan performa yang lebih rendah pada kisaran 83%. Gambar 6 menyajikan kurva ROC seluruh model. Kurva *XGBoost* yang paling mendekati sudut kiri atas mengindikasikan kemampuan diskriminasi terbaik dengan AUC-ROC sebesar 94,09%, diikuti oleh *Random Forest* tanpa SMOTE (93,87%) dan DT dengan SMOTE (93,25%). Sementara itu, Gambar 7 menampilkan *confusion matrix* yang menggambarkan distribusi prediksi benar dan salah pada masing-masing kelas.

3.5 Analisis dan Pembahasan Perbandingan Model

3.5.1 XGBoost sebagai Model Terbaik

Berdasarkan Tabel 5, *XGBoost* tanpa SMOTE menunjukkan performa terbaik dengan *FI-score* 90,39%, *accuracy* 90,48%, *precision* 99,91%, *recall* 82,52%, dan AUC-ROC 94,09%. Untuk menganalisis pola kesalahan klasifikasi secara lebih rinci, *confusion matrix* model terbaik tersebut disajikan pada Gambar 7.



Gambar 7. *Confusion Matrix* XGBoost (Tanpa SMOTE)

Gambar 7 menunjukkan asimetri yang menarik dalam pola prediksi model. Kelas Tidak Beresiko mencapai *recall* sempurna (1,00) dengan *precision* 0,83, sedangkan kelas Beresiko mencapai *precision* sempurna (1,00) dengan *recall* 0,82. Hasil ini menunjukkan bahwa model sangat jarang menghasilkan *false positive* pada kelas Beresiko, tetapi masih melewatkan sebagian kasus beresiko (*false negative*). Meskipun demikian, kombinasi

precision dan *recall* tersebut tetap menghasilkan *F1-score* tertinggi di antara seluruh model yang diuji. Kinerja *XGBoost* ini sejalan dengan fungsi objektif pada Persamaan (10), yang mengombinasikan minimisasi *loss* dan regularisasi L1/L2 untuk menangani data tabular heterogen berdimensi tinggi secara efektif.

3.5.2 Analisis Overfitting dan Generalisasi Model

Tabel 6 membandingkan CV *F1 training* dengan *F1-score* pada data *test* sebagai indikator *kuantitatif overfitting*. Nilai delta positif pada semua model mengindikasikan tidak adanya *overfitting*.

Tabel 6. Analisis generalisasi: CV dan F1 test (skenario A)

Algoritma	CV F1 (Train)	F1 test	Selisih (Δ)	Status
DT	0,9006	0,9035	+0,003	Generalisasi baik
KNN	0,8283	0,8368	+0,009	Generalisasi baik
SVM	0,8996	0,9022	+0,003	Generalisasi baik
RF	0,9016	0,9028	+0,001	Sangat stabil
XGBoost	0,9015	0,9039	+0,002	Sangat stabil

Berdasarkan Tabel 6, seluruh model menunjukkan nilai *F1-score* pada data *test* yang sedikit lebih tinggi dibandingkan CV *F1-score* pada data *training* ($\Delta > 0$), sehingga tidak terdapat indikasi *overfitting* yang signifikan. Nilai Δ terkecil diperoleh oleh RF (0,001) dan *XGBoost* (0,002), yang menunjukkan kemampuan generalisasi paling stabil. Temuan ini sejalan dengan mekanisme regularisasi *inheren* pada kedua algoritma, sebagaimana ditunjukkan pada Persamaan (9) dan (10). *Random Forest* mengandalkan rata-rata prediksi antar-pohon (*averaging*), sedangkan *XGBoost* menambahkan regularisasi L1/L2 yang membantu mengendalikan kompleksitas model dan mengurangi sensitivitas terhadap *noise* pada data *training*.

3.5.3 Analisis Pengaruh SMOTE

Tabel 7 menyajikan perbandingan *F1-score* dan waktu pelatihan antara skenario tanpa *SMOTE* (A) dan dengan *SMOTE* (B). Hasil menunjukkan bahwa penerapan *SMOTE* tidak memberikan peningkatan performa yang berarti pada seluruh algoritma. Perubahan *F1-score* hanya berkisar antara -0,0067 hingga +0,0005, sedangkan waktu pelatihan meningkat sebesar 18%-65%. Temuan ini mengindikasikan bahwa pada dataset P3KE dengan rasio kelas 1,19:1, yang tergolong ketidakseimbangan ringan (*mild imbalance*), penambahan sampel sintetis melalui *SMOTE* tidak menghasilkan informasi tambahan yang cukup untuk meningkatkan kinerja model. Dengan demikian, pada data dengan karakteristik serupa, penggunaan *SMOTE* cenderung menambah biaya komputasi tanpa memberikan peningkatan performa yang sepadan.

Tabel 7. Pengaruh *SMOTE* terhadap *F1-score* dan waktu *training*

Algoritma	F1 (A)	F1 (B)	Δ F1	Waktu A (det)	Waktu B (det)	Δ Waktu
DT	0,9035	0,9035	0,0000	5	8	+65%
KNN	0,8368	0,8301	-0,0067	89	110	+24%
SVM	0,9022	0,9021	-0,0001	2421	3063	+27%
RF	0,9028	0,9033	+0,0005	71	84	+18%
XGBoost	0,9039	0,9037	-0,0001	13	18	+38%

3.5.4 Analisis Dimensionalitas

Ekspansi dimensi dari 40 atribut menjadi 162 fitur setelah *preprocessing* memberikan dampak yang berbeda pada setiap algoritma. Model berbasis pohon, seperti DT, RF, dan *XGBoost*, tetap menunjukkan performa yang tinggi karena mampu melakukan seleksi fitur secara adaptif pada setiap *split*, sehingga kurang sensitif terhadap peningkatan dimensi. Sebaliknya, KNN mengalami penurunan kinerja akibat fenomena *curse of dimensionality*. Pada ruang fitur berdimensi tinggi, jarak antar sampel cenderung menjadi semakin seragam sehingga konsep tetangga terdekat kehilangan kemampuan diskriminatifnya. Kondisi ini menjelaskan mengapa *F1-score* KNN 83,41% tertinggal jauh dibandingkan model berbasis pohon yang mencapai *F1-score* di atas 90%, meskipun seluruh algoritma menggunakan data dan tahapan *preprocessing* yang sama.

3.6 Implikasi Kebijakan bagi Penanganan Stunting di Provinsi Gorontalo

Temuan analisis *SHAP* memberikan implikasi praktis bagi Pemerintah Provinsi Gorontalo. Dominasi BAB: Milik sendiri dan Indeks Sanitasi sebagai fitur paling berpengaruh menunjukkan bahwa kondisi sanitasi merupakan indikator utama dalam mengidentifikasi rumah tangga berisiko *Stunting*. Hasil ini sejalan dengan kerangka *WASH* (*Water, Sanitation, and Hygiene*) dan memperkuat bukti bahwa upaya penurunan *Stunting* perlu diintegrasikan dengan peningkatan akses sanitasi dasar.

Secara operasional, hasil penelitian ini dapat dimanfaatkan dalam tiga aspek. Pertama, indikator sanitasi dapat digunakan sebagai kriteria *screening* awal untuk memprioritaskan verifikasi lapangan oleh kader posyandu

sehingga sumber daya yang terbatas dapat difokuskan pada rumah tangga dengan risiko tertinggi. Kedua, temuan ini memberikan dasar empiris bagi penguatan program lintas sektor, seperti integrasi indikator sanitasi ke dalam sasaran Bapak Asuh Anak *Stunting* (BAAS) dan *Bele Mo'osehati*. Ketiga, kemampuan model dalam mengidentifikasi rumah tangga prioritas langsung dari data P3KE memungkinkan penyasaran intervensi tanpa menunggu pengumpulan data antropometri, serta mendukung pemetaan disparitas risiko hingga tingkat kecamatan dan desa. Dengan demikian, pendekatan yang diusulkan berpotensi menjadi instrumen pendukung pengambilan keputusan yang efisien, transparan, dan mudah diinterpretasikan dalam perumusan kebijakan penanggulangan *Stunting* yang lebih tepat sasaran.

3.7 Hasil dan Analisis SHAP

Untuk meningkatkan interpretabilitas model, analisis *SHAP* diterapkan pada model terbaik, yaitu *XGBoost* tanpa *SMOTE*, menggunakan 1.000 sampel data test. Tabel 8 menampilkan 20 fitur dengan nilai *Mean |SHAP|* tertinggi, yang menggambarkan besarnya kontribusi rata-rata setiap fitur terhadap prediksi model pada tingkat global.

Gambar 8 (*SHAP summary plot*) menunjukkan distribusi nilai *SHAP* untuk 20 fitur terpenting. Titik berwarna merah merepresentasikan nilai fitur yang tinggi, sedangkan biru menunjukkan nilai yang rendah. Posisi horizontal menggambarkan arah dan besarnya kontribusi fitur terhadap prediksi model. Terlihat bahwa dua fitur sanitasi, yaitu BAB: Milik sendiri (*Mean |SHAP|* = 3,0256) dan Indeks Sanitasi (*Mean |SHAP|* = 1,5645), mendominasi kontribusi model. Nilai BAB: Milik sendiri hampir dua kali lipat Indeks Sanitasi, dan keduanya jauh melampaui fitur peringkat ketiga, yaitu Usia Kepala Keluarga (*Mean |SHAP|* = 0,2426).

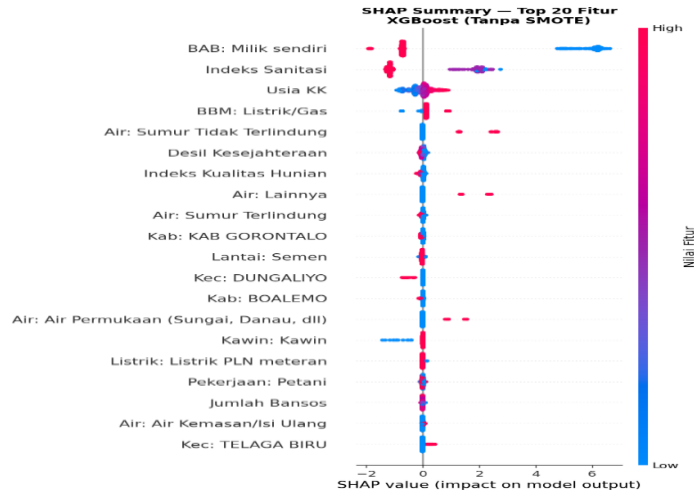
Fitur BAB: Milik sendiri menunjukkan pola yang konsisten, di mana kepemilikan fasilitas sanitasi (nilai tinggi) berkorelasi dengan nilai *SHAP* negatif yang menurunkan probabilitas kelas Beresiko, sedangkan ketiadaan fasilitas tersebut (nilai rendah) berkorelasi dengan nilai *SHAP* positif yang meningkatkan probabilitas risiko. Perlu dicatat bahwa fitur BAB: Milik sendiri juga merupakan salah satu komponen pembentuk Indeks Sanitasi. Kemunculan keduanya pada peringkat teratas menegaskan bahwa dimensi sanitasi merupakan determinan utama dalam model, baik ketika direpresentasikan sebagai indikator tunggal maupun sebagai indeks gabungan.

Distribusi nilai *Mean |SHAP|* pada Tabel 8 juga memperlihatkan pola *long-tail*, di mana dua fitur teratas memiliki kontribusi yang jauh lebih besar dibandingkan 18 fitur lainnya (0,0119-0,2426). Temuan ini menunjukkan bahwa dimensi sanitasi merupakan sumber informasi paling dominan yang digunakan model *XGBoost* dalam memprediksi risiko *Stunting*.

Munculnya fitur spasial Kabupaten Gorontalo (*Mean |SHAP|* = 0,0279; peringkat 10) dan Kecamatan Dungaliyo (*Mean |SHAP|* = 0,0203; peringkat 12) mengindikasikan adanya pengaruh kontekstual wilayah yang berhasil ditangkap model melalui representasi *one-hot encoding* lokasi. Temuan ini menunjukkan bahwa model mampu mengidentifikasi pola spasial laten yang tidak sepenuhnya dijelaskan oleh karakteristik sosial-ekonomi rumah tangga.

Tabel 8. Rangking 20 fitur terpenting berdasarkan *Mean |SHAP|*

Rank	Fitur	<i>Mean SHAP </i>
1	BAB: Milik sendiri	3,0256
2	Indeks Sanitasi	1,5645
3	Usia KK	0,2426
4	BBM: Listrik/Gas	0,1506
5	Air: Sumur Tidak Terlindung	0,0993
6	Desil Kesejahteraan	0,0486
7	Indeks Kualitas Hunian	0,0486
8	Air: Lainnya	0,0391
9	Air: Sumur Terlindung	0,0287
10	Kab: Kab. Gorontalo	0,0279
11	Lantai: Semen	0,0266
12	Kec: Dungaliyo	0,0203
13	Kab: Boalemo	0,0178
14	Air: Permukaan (Sungai/Danau, dll)	0,0176
15	Kawin: Kawin	0,0175
16	Listrik: Listrik PLN Meteran	0,0141
17	Pekerjaan: Petani	0,0137
18	Jumlah Bansos	0,0136
19	Air: Air Kemasan/Isi Ulang	0,0128
20	Kec. Telaga Biru	0,0119



Gambar 8. SHAP Summary Plot dengan Top 20 Fitur Penting

Selain itu, fitur hasil *feature engineering*, yaitu Usia Kepala Keluarga (peringkat 3) dan Indeks Kualitas Hunian (peringkat 7), memiliki kontribusi prediktif yang lebih tinggi dibandingkan banyak fitur asli. Hasil ini mengonfirmasi efektivitas proses *feature engineering* dalam meningkatkan representasi informasi pada model. Secara keseluruhan, temuan SHAP konsisten dengan penelitian Islam *et al.* [21] yang mengidentifikasi sanitasi sebagai determinan utama *Stunting*, serta Wicaksono *et al.* [20] yang menunjukkan bahwa SHAP mampu mengungkap faktor-faktor yang relevan untuk mendukung perumusan kebijakan berbasis data.

4. KESIMPULAN

Penelitian ini membangun dan membandingkan lima algoritma *Machine learning* untuk prediksi risiko *Stunting* berbasis data registrasi sosial P3KE Provinsi Gorontalo serta menginterpretasikan hasil prediksi menggunakan SHAP. Hasil penelitian menunjukkan bahwa XGBoost tanpa SMOTE merupakan model terbaik dengan *F1-score* 90,39%, *accuracy* 90,48%, *precision* 99,91%, *recall* 82,52%, dan AUC-ROC 94,09%, dengan waktu pelatihan sekitar 13 detik. Penerapan SMOTE tidak memberikan peningkatan performa yang berarti, tetapi meningkatkan waktu pelatihan sebesar 18%-65%, sehingga kurang efektif pada distribusi kelas yang mendekati seimbang (1,19:1). Analisis SHAP mengidentifikasi BAB: Milik sendiri (*Mean |SHAP|* = 3,0256) dan Indeks Sanitasi (*Mean |SHAP|* = 1,5645) sebagai faktor yang paling berpengaruh dalam prediksi risiko *Stunting*. Temuan ini menunjukkan bahwa data registrasi sosial P3KE tanpa data antropometri tetap mampu mendukung prediksi risiko *Stunting* dengan *F1-score* di atas 90%. Model dan profil SHAP yang dihasilkan berpotensi dimanfaatkan sebagai alat pendukung keputusan untuk membantu pemerintah daerah memprioritaskan intervensi sanitasi pada rumah tangga berisiko. Namun, penelitian ini memiliki beberapa keterbatasan. Label *Resiko Stunting* pada P3KE merupakan penanda administratif berbasis kondisi sosial-ekonomi, bukan hasil pengukuran antropometri, sehingga model memprediksi risiko pada tingkat rumah tangga dan bukan diagnosis klinis pada individu anak. Selain itu, cakupan data masih terbatas pada satu provinsi dan ruang pencarian *hyperparameter* dibatasi untuk menjaga efisiensi komputasi. Penelitian selanjutnya dapat mengintegrasikan data klinis Puskesmas dengan data P3KE, menerapkan seleksi fitur untuk mereduksi 162 fitur menjadi subset yang lebih optimal, serta mengembangkan analisis spasial yang lebih granular hingga tingkat kecamatan dan desa.

REFERENCES

- [1] N. Sari and J. Christy, "Factors Influencing Stunting among Children Aged 0–59 Months," *Indonesian Journal of Global Health Research*, vol. 7, no. 4, pp. 981–988, Aug. 2025, doi: 10.37287/ijghr.v7i4.6572.
- [2] V. Sideropoulos, A. Draper, B. A. L. Munoz-Chereau, and J.E Dockrell, "Childhood stunting and cognitive development: a meta-analysis," *J Glob Health*, Sep. 2025, doi: <https://doi.org/10.7189/jogh.15.04257>.
- [3] World Health Organization, "Joint Child Malnutrition Estimates," WHO (World Health Organization). Accessed: May 14, 2026. [Online]. Available: <https://www.who.int/data/gho/data/themes/topics/joint-child-malnutrition-estimates-unicf-who-wb>
- [4] Widyawati, "Penurunan Prevalensi Stunting tahun 2021 sebagai Model Menuju Generasi Emas Indoensia 2045," Kementerian Kesehatan Republik Indonesia. Accessed: May 17, 2026. [Online]. Available: <https://kemkes.go.id/id/penurunan-prevalensi-stunting-tahun-2021-sebagai-modal-menuju-generasi-emas-indonesia-2045>
- [5] L. S. Situmorang, "Implementasi Peraturan Presiden Nomor 72 Tahun 2021 tentang Percepatan Penurunan Stunting (Studi Kasus di Kelurahan Sicanang Kecamatan Medan Belawan)," Universitas Mdedan Area, Medan, 2023. Accessed: May 17, 2026. [Online]. Available: <https://repositori.uma.ac.id/jspui/handle/123456789/21089>

- [6] M. Butarbutar, D. P. Munthe, A. P. Ningsih, B. S. R. Panjaitan, Pendidikan Ilmu Kesehatan Masyarakat, and Penelitian, “Gambaran Kejadian *Stunting* Pada Balita Di Puskesmas Matani Kota Tomohon *Overview of Stunting Incidents in Toddlers at the Matani Community Health Center, Tomohon City,*” *Jurnal Kolaboratif Sains*, vol. 8, no. 12, pp. 8497–8502, 2025, doi: 10.56338/jks.v8i12.9706.
- [7] Y. Monoarfa, “Prevalensi *Stunting* di Gorontalo masih di Atas Nasional, Namun Alami Penurunan Signifikan,” Dinas Kesehatan Provinsi Gorontalo. Accessed: May 17, 2026. [Online]. Available: <https://dinkes.gorontaloprov.go.id/prevalensi-stunting-di-gorontalo-masih-di-atas-angka-nasional-namun-alami-penurunan-signifikan/>
- [8] N. Pembengo, “300 Dokter Anggota IDI Gorontalo Menjadi Orang Tua Asuh Keluarga Berisiko *Stunting*,” Dinas Kesehatan Provinsi Gorontalo. Accessed: May 17, 2026. [Online]. Available: <https://dinkes.gorontaloprov.go.id/300-dokter-anggota-idi-gorontalo-menjadi-orang-tua-asuh-keluarga-berisiko-stunting/>
- [9] Z. Mile and I. Mohamad, “Tekan Angka *Stunting* 28,2 Persen, Wabup Tonny Junus Luncurkan Gerakan Orang Tua Asuh ‘GENTING,’” Pemerintah Kabupaten Gorontalo. Accessed: May 17, 2026. [Online]. Available: <https://gorontalo.kab.go.id/berita/963/tekan-angka-stunting-28-2-persen-wabup-tonny-junus-luncurkan-gerakan-orang-tua-asuh-genting>
- [10] R. Rahutomo, G. N. Elwirehardja, M. Isnain, F. Asadi, and B. Pardamean, “*Machine Learning Implementations in Childhood Stunting Research: A Systematic Literature Review,*” in *2023 International Conference on Information Management and Technology (ICIMTech)*, 2023, pp. 229–234. doi: 10.1109/ICIMTech59029.2023.10277881.
- [11] I. Emilda, F. Hidayat, and B. Adiwino, “*A Systematic Literature Review on the Application of Machine Learning for Predicting Stunting Prevalence in Indonesia (2020–2024),*” *Journal SISFOKOM*, vol. 14, pp. 277–283, Jul. 2025, doi: <https://doi.org/10.32736/sisfokom.v14i3.2366>.
- [12] 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 Health*, pp. 41–49, Jan. 2023, doi: 10.3961/jpmp.22.388.
- [13] O. N. Chilyabanyama et al., “*Performance of Machine Learning Classifiers in Classifying Stunting among Under-Five Children in Zambia,*” *Children*, vol. 9, no. 7, 2022, doi: 10.3390/children9071082.
- [14] M. K. Ayele, G. A. Baye, S. H. Yesuf, A. A. Engda, and E. T. Mitiku, “*Predicting stunting status among under five children in ethiopia using ensemble machine learning algorithms,*” *Sci. Rep.*, vol. 15, no. 1, p. 27907, 2025, doi: 10.1038/s41598-025-03206-1.
- [15] T. Sugihartono, B. Wijaya, Marini, A. F. Alkayes, and H. A. Anugrah, “*Optimizing Stunting Detection through SMOTE and Machine Learning: a Comparative Study of XGBoost, Random Forest, SVM, and k-NN,*” *Journal of Applied Data Sciences*, vol. 6, no. 1, pp. 667–682, Jan. 2025, doi: 10.47738/jads.v6i1.494.
- [16] H. W. Jemil, S. W. Semayneh, A. B. Kassaw, and K. D. Gashu, “*Predicting severe stunting and its determinants among under-five in Eastern African Countries: A machine learning algorithms,*” *PLoS One*, vol. 21, no. 1, p. e0340221, Jan. 2026, doi: 10.1371/journal.pone.0340221.
- [17] P. K. Arya, K. Sur, T. Kundu, S. Dhote, and S. K. Singh, “*Unveiling predictive factors for household-level stunting in India: A machine learning approach using NFHS-5 and satellite-driven data,*” *Nutrition*, vol. 132, p. 112674, 2025, doi: <https://doi.org/10.1016/j.nut.2024.112674>.
- [18] N. Hasdyana, R. K. Dinata, Rahmi, and T. I. Fajri, “*Hybrid Machine Learning for Stunting Prevalence: A Novel Comprehensive Approach to Its Classification, Prediction, and Clustering Optimization in Aceh, Indonesia,*” *Informatics*, vol. 11, no. 4, 2024, doi: 10.3390/informatics11040089.
- [19] N. Novalina, I. A. A. Tarigan, F. K. Kameela, and M. Rizkinia, “*Benchmarking machine learning algorithm for stunting risk prediction in Indonesia,*” *Bulletin of Electrical Engineering and Informatics*, vol. 14, no. 3, pp. 2252–2263, Jun. 2025, doi: 10.11591/eei.v14i3.8997.
- [20] A. Wicaksono, D. Prasetyo, Y. Mar’atullatifah, D. U. Iswawigra, H. Mahmudah, and A. Hapsari, “*Data Analysis and Explainable Machine Learning for Stunting Prediction,*” *Journal of Artificial Intelligence and Legal Technology*, vol. 1, no. 1, pp. 35–44, Aug. 2025, Accessed: May 12, 2026. [Online]. Available: <https://journal.sah.co.id/index.php/JAILT/article/view/4>
- [21] M. Islam, N. S. J. Kibria, S. Kumar, D. C. Roy, and R. Karim, “*Prediction of undernutrition and identification of its influencing predictors among under-five children in Bangladesh using explainable machine learning algorithms,*” *PLoS One*, vol. 19, no. 12, pp. e0315393–, Dec. 2024, [Online]. Available: <https://doi.org/10.1371/journal.pone.0315393>
- [22] M. A. Hamid and R. Subhiyakt, “*Performance Comparison of Random Forest, SVM, and XGBoost Algorithms with SMOTE for Stunting Prediction,*” *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 4, pp. 1163–1169, Aug. 2025, doi: 10.30871/jaic.v9i4.9701.
- [23] V. R. Saragih, A. Arnita, Z. Indra, I. Taufik, and M. S. Sinaga, “*Comparison of supervised machine learning methods in predicting the prevalence of stunting in North Sumatra Province,*” *Journal of Soft Computing Exploration*, vol. 5, no. 4, pp. 370–379, Dec. 2024, doi: 10.52465/josce.v5i4.498.
- [24] M. A. E. Pratama, S. Hendra, H. R. Ngemba, R. Nur, R. Azhar, and R. Laila, “*Comparison of Machine Learning Algorithms for Predicting Stunting Prevalence in Indonesia,*” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 13, no. 2, pp. 200–209, Jun. 2024, doi: 10.32736/sisfokom.v13i2.2097.

- [25] M. Handayani and M. F. L. Sibuea, "Performance Analysis of Clustering Models Based on Machine Learning in Stunting Data Mapping," *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 9, no. 4, pp. 715–720, Sep. 2023, doi: 10.33330/jurteksi.v9i4.2770.
- [26] J. Joharini and A. Subekti, "A Comparative Methodological Study of Automated Machine Learning for Multiclass Stunting Prediction Using Anthropometric Data," *sinkron*, vol. 10, no. 2, pp. 991–1002, Apr. 2026, doi: 10.33395/sinkron.v10i2.15886.
- [27] S. Syahrial, R. Ilham, Z. F. Asikin, and St. S. I. Nurdin, "Stunting Classification in Children's Measurement Data Using Machine Learning Models," *Journal La Multiapp*, vol. 3, no. 2, pp. 52–60, Mar. 2022, doi: 10.37899/journallamultiapp.v3i2.614.
- [28] M. Yunus, M. K. Biddinika, and A. Fadlil, "Comparison of Machine Learning Algorithms for Stunting Classification," *Scientific Journal of Engineering Research*, vol. 1, no. 2, pp. 64–70, Apr. 2025, doi: 10.64539/sjer.v1i2.2025.9.
- [29] M. Ibnu Choldun Rachmatullah and S. Armiati, "Menerapkan SMOTE Pada Klasifikasi Data Penyakit Stroke," *Improve Jurnal Ilmiah Manajemen Informatika*, vol. 17, no. 1, pp. 9–12, Apr. 2025, Accessed: May 17, 2026. [Online]. Available: <https://ejournal.ulbi.ac.id/index.php/improve/article/view/4307>
- [30] S. A. Utiahman, and A. M. M. Pratama, "KLIK: Kajian Ilmiah Informatika dan Komputer Analisis Perbandingan KNN, SVM, Decision Tree dan Regresi Logistik Untuk Klasifikasi Obesitas Multi Kelas," *Media Online*, vol. 4, no. 6, pp. 3137–3146, 2024, doi: 10.30865/klik.v4i6.1871.
- [31] H. Joe and H.-G. Kim, "Multi-label classification with XGBoost for metabolic pathway prediction," *BMC Bioinformatics*, vol. 25, no. 1, p. 52, 2024, doi: 10.1186/s12859-024-05666-0.