

Hybrid Cost-Sensitive Rejection Sampling Approach for Improving Stunting Classification Performance on Imbalanced Data

Ardiyansyah^{1*}, Windi Irmayani², dan Lisnawanty²

¹Departemen Sistem Informasi Kampus Kota Pontianak, Universitas Bina Sarana Informatika, Kota Pontianak, Indonesia

²Departemen Informatika Kampus Kota Pontianak, Universitas Bina Sarana Informatika, Kota Pontianak, Indonesia

*Corresponding author: ardiyansyah.arq@bsi.ac.id

Submitted : 24 Mei 2026

Revised : 01 September 2026

Accepted : 16 September 2026

ABSTRACT

Stunting remains a critical child health problem in Indonesia, with a national prevalence of 21.6% based on the 2022 Indonesian Nutritional Status Survey (SSGI). Machine learning-based classification models face a class imbalance challenge, further complicated when the imbalance direction in the dataset (79.55% stunting versus 20.45% non-stunting) deviates from the general epidemiological pattern. This study proposes a hybrid Cost-Sensitive Rejection Sampling (CSRS) approach, integrating probability-based rejection sampling with example-dependent cost-sensitive learning, evaluated across six decision tree algorithms (J48Consolidate, CDT, C4.5, LADTree, LMT, Hoeffding Tree) and compared against five sampling techniques. Experimental results demonstrate that CSRS combined with C4.5 achieved the best performance, with Precision of 0.963, Recall of 0.970, F-Measure of 0.960, and Accuracy of 0.9703, outperforming SMOTE by an 11.8% relative improvement in F-Measure. These findings confirm that combining data-level and algorithm-level imbalance handling yields superior stunting classification performance while preserving clinically realistic data.

Keywords: C4.5, Cost-Sensitive Rejection Sampling, Class Imbalance, Decision Tree, Stunting Classification.



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

I. PENDAHULUAN

Stunting merupakan salah satu permasalahan kesehatan anak yang paling kritis di Indonesia, ditandai dengan kondisi gagal tumbuh akibat kekurangan gizi kronis terutama pada 1.000 hari pertama kehidupan. *World Health Organization* (WHO) mendefinisikan *stunting* sebagai kondisi di mana tinggi badan menurut umur (TB/U) berada di bawah minus dua standar deviasi dari median pertumbuhan referensi anak WHO (Organization, 2006). Dampak stunting tidak hanya terbatas pada aspek fisik, tetapi juga mencakup gangguan perkembangan kognitif, penurunan kapasitas belajar, serta peningkatan risiko penyakit kronis pada usia dewasa yang pada akhirnya berkontribusi terhadap hilangnya potensi sumber daya manusia suatu bangsa (Black et al., 2013).

Indonesia masih menghadapi tantangan serius dalam penurunan angka stunting. Berdasarkan hasil Survei Status Gizi Indonesia (SSGI) tahun 2022, prevalensi stunting nasional berada pada angka 21,6%, dan pemerintah menargetkan penurunan hingga 14% pada tahun 2024 melalui Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020–2024 (PPN/Bappenas, 2020). Tingginya angka tersebut mendorong kebutuhan mendesak akan sistem deteksi dini yang lebih akurat, efisien, dan dapat dioperasikan secara massal di fasilitas kesehatan tingkat primer seperti Puskesmas dan Posyandu.

Perkembangan teknologi kecerdasan buatan, khususnya *machine learning*, membuka peluang besar dalam mendukung deteksi dan klasifikasi *stunting* secara otomatis. Berbagai algoritma klasifikasi berbasis pohon keputusan telah banyak diaplikasikan dalam domain kesehatan anak karena kemampuannya menghasilkan model yang interpretatif dan mudah dipahami oleh tenaga kesehatan (Menzies et al., 2007). Dalam penelitian ini, enam algoritma klasifikasi digunakan, yaitu J48Consolidate (Ibarguren et al., 2015), Credal Decision Tree (CDT) (Abellan & Moral, 2003), C4.5 (Quinlan, 1993), LADTree (Holmes et al., 2002), Logistic Model Tree (LMT) (Landwehr et al., 2005), dan Hoeffding Tree (Hulten et al., 2001). Algoritma-algoritma tersebut dipilih karena merepresentasikan ragam

pendekatan pohon keputusan, mulai dari pohon klasik (C4.5), pohon logistik (LMT), pohon berbasis boosting (LADTree), pohon konsensus (J48Consolidate), hingga pohon yang dirancang untuk data aliran besar (Hoeffding Tree), sehingga memungkinkan evaluasi perbandingan yang komprehensif pada konteks data stunting (Witten et al., 2016).

Namun demikian, tantangan mendasar yang dihadapi dalam pengembangan model klasifikasi *stunting* adalah masalah ketidakseimbangan kelas (*class imbalance*) (He & Ma, 2013). Penelitian-penelitian sebelumnya tentang klasifikasi *stunting* berbasis machine learning umumnya hanya menggunakan satu pendekatan penanganan ketidakseimbangan kelas dan terbatas pada satu atau dua algoritma klasifikasi (Nurjanah et al., 2022)(Prasetyo et al., 2022). Selain itu, penelitian sebelumnya ada juga yang melakukan pendekatan pembagian data untuk menghitung laju kesalahan (*error rate*) dalam mengatasi ketidakseimbangan kelas dengan menggunakan satu algoritma klasifikasi seperti pada penelitian (Sinaga et al., 2023) menggunakan pendekatan *Binary Logistic Regression*, penelitian (Zullani et al., 2023) menggunakan pendekatan CART dan penelitian (Ilanda et al., 2023) menggunakan pendekatan C4.5. Berbagai teknik telah dikembangkan untuk mengatasi masalah ketidakseimbangan kelas, yang secara umum dibagi menjadi tiga kategori: (1) pendekatan tingkat data (*data-level*) seperti SMOTE, oversampling dan undersampling; (2) pendekatan tingkat algoritma (*algorithm-level*) seperti cost-sensitive learning; dan (3) pendekatan *hybrid* yang menggabungkan keduanya (Danaei et al., 2016).

Ketidakseimbangan kelas pada data kesehatan anak dapat terjadi pada kedua arah yaitu pada populasi umum, kasus *stunting* umumnya menjadi kelas minoritas dibandingkan anak dengan status gizi normal, namun pada dataset yang digunakan dalam penelitian ini justru terjadi pola sebaliknya, di mana kelas *non-stunting* menjadi minoritas. Kondisi ini menyebabkan algoritma klasifikasi standar tetap cenderung bias terhadap kelas mayoritas (El-muslih et al., 2023), sehingga menghasilkan nilai sensitivitas yang rendah dalam mendeteksi kelas minoritas. Pada konteks penelitian ini, hal tersebut berisiko menurunkan kemampuan model dalam mengenali anak *non-stunting* secara akurat, sekaligus meningkatkan potensi kesalahan klasifikasi berupa anak sehat yang salah diprediksi sebagai *stunting* (*false positive*). Di sisi lain, kegagalan mendeteksi anak yang sesungguhnya mengalami *stunting* (*false negative*) tetap merupakan kesalahan paling kritis dalam konteks klinis, karena dapat mengakibatkan keterlambatan intervensi gizi yang bersifat ireversibel (Fernández et al., 2018).

Oleh karena itu, dibutuhkan pendekatan penanganan ketidakseimbangan kelas yang tidak hanya menyeimbangkan distribusi data secara acak, tetapi juga mempertimbangkan bobot biaya kesalahan klasifikasi secara asimetris pada kedua kelas (He & Garcia, 2009). Cost-Sensitive Oversampling menggabungkan prinsip pemberian bobot biaya yang lebih tinggi pada kesalahan kelas minoritas dengan mekanisme oversampling, sehingga proses pembelajaran tidak hanya mempertimbangkan distribusi data tetapi juga konsekuensi asimetris dari kesalahan klasifikasi (Elkan, 2001). Pendekatan ini menghasilkan peningkatan yang lebih konsisten dibandingkan oversampling konvensional, namun masih memiliki keterbatasan dalam mengendalikan noise yang diperkenalkan oleh sampel sintetis. Penelitian ini mengusulkan pendekatan *hybrid Cost-Sensitive Rejection Sampling*, yang menggabungkan mekanisme *rejection sampling* berbasis probabilitas dengan skema pembobotan biaya (*cost-sensitive learning*), guna meningkatkan sensitivitas deteksi kelas minoritas tanpa mengorbankan akurasi keseluruhan model klasifikasi *stunting*. Penelitian ini menggunakan Evaluasi yang komprehensif yang membandingkan berbagai teknik sampling, yaitu SMOTE, undersampling, oversampling, Cost-Sensitive Oversampling, dan Cost-Sensitive Rejection Sampling, secara bersamaan dalam kombinasi dengan beragam algoritma pohon keputusan belum pernah dilakukan secara sistematis pada konteks data *stunting*. Kesenjangan penelitian inilah yang menjadi motivasi utama dari penelitian ini. Sebagai pengembangan lebih lanjut, penelitian ini mengeksplorasi kombinasi antara teknik *sampling* dan pendekatan *cost-sensitive learning* dalam kerangka *hybrid*.

Di antara seluruh metode sampling yang diuji, Cost-Sensitive Rejection Sampling (CSRS) tampil sebagai pendekatan yang paling dominan dan memberikan kinerja terbaik dalam penelitian ini. CSRS bekerja dengan mekanisme probabilistik yang secara selektif menolak sampel kelas mayoritas berdasarkan fungsi kepadatan probabilitas yang diadaptasi terhadap distribusi kelas, sekaligus mengintegrasikan matriks biaya asimetris ke dalam fungsi objektif algoritma klasifikasi (Robert & Casella, 2004). Berbeda dengan SMOTE yang menciptakan data baru, CSRS hanya bekerja pada sampel nyata yang telah tersedia, sehingga menghindari risiko introduksi pola sintetis yang tidak representatif terhadap kondisi klinis aktual. Integrasi simultan antara *rejection sampling* dan *cost-sensitive learning* menciptakan sinergi yang menghasilkan *trade-off optimal* antara sensitivitas pendeteksian *stunting* dan spesifisitas klasifikasi secara keseluruhan (Haixiang et al., 2017).

Berdasarkan uraian di atas, penelitian ini bertujuan untuk: (1) mengevaluasi kinerja enam algoritma klasifikasi berbasis pohon keputusan, yaitu J48Consolidate, CDT, C4.5, LADTree, LMT, dan Hoeffding Tree, pada data *stunting* yang tidak seimbang; (2) membandingkan efektivitas lima teknik penanganan ketidakseimbangan kelas, yaitu

SMOTE, undersampling, oversampling, Cost-Sensitive Oversampling, dan Cost-Sensitive Rejection Sampling; (3) mengidentifikasi kombinasi metode sampling dan algoritma klasifikasi yang memberikan kinerja terbaik; serta (4) menganalisis kontribusi dominan Cost-Sensitive Rejection Sampling sebagai pendekatan *hybrid* yang paling efektif dalam meningkatkan akurasi deteksi kasus stunting pada data tidak seimbang.

II. METODE PENELITIAN

Penelitian ini dilaksanakan melalui serangkaian tahapan sistematis yang mencakup pengumpulan dataset, preprocessing data, penerapan metode sampling, implementasi algoritma klasifikasi, dan evaluasi kinerja model. Alur keseluruhan tahapan penelitian disajikan pada uraian berikut.

A. Dataset

Dataset yang digunakan dalam penelitian ini bersumber dari platform publik *Kaggle* yang dapat diakses melalui tautan <https://www.kaggle.com/datasets/harnelia/faktor-stunting/data>. Dalam penelitian ini, label kelas *stunting* pada data yang digunakan terdapat dua kategori biner yaitu *yes* dan *no*. Total keseluruhan data berjumlah 10.000 rekaman balita dengan distribusi kelas yang sangat tidak seimbang, di mana proporsi kelas *non-stunting* jauh lebih kecil dibandingkan kelas *stunting*. Rincian distribusi kelas sebelum *preprocessing* disajikan pada Tabel 1.

Tabel 1. Distribusi Kelas Dataset Stunting

Sample Size	Features	Imbalance Class Majority : Stunting (yes)	Imbalance Class Minority : Non-Stunting (no)	Imbalance Ratio
10,000	8	7,955	2,045	3,89 : 1

Rasio ketidakseimbangan (*Imbalance Ratio*/IR) dataset ini tergolong dalam kategori ketidakseimbangan moderat yang artinya jumlah data anak yang mengalami stunting hampir 4 kali lebih banyak di bandingkan data anak yang non-stunting sehingga berpotensi menurunkan kinerja algoritma klasifikasi standar secara signifikan apabila tidak ditangani (Yulianti et al., 2023).

B. Preprocessing Data

Tahapan ini mencakup encoding variabel kategorikal menjadi bentuk numerik dan melakukan pembagian data dengan rasio 70:30 (Joseph, 2022) yaitu 70% data *training* agar model memperoleh pola data yang memadai, sedangkan 30% data *testing* digunakan sebagai data pengujian untuk menghasilkan kinerja model yang stabil. Penanganan *class imbalance* diterapkan pada data *training* (70%) dengan beberapa skenario yaitu tanpa sampling (*baseline*), Undersampling, Oversampling, SMOTE, Cost-Sensitive Oversampling dan Cost-Sensitive Rejection Sampling (CSRS).

C. Cost-Sensitive Rejection Sampling

Metode Cost-Sensitive Rejection Sampling yang diimplementasikan dalam penelitian ini mengacu pada formulasi yang dikembangkan oleh Alejandro Correa Bahnsen dalam penelitiannya tentang *example-dependent cost-sensitive learning* (Correa Bahnsen et al., 2015). Berbeda dengan pendekatan *cost-sensitive* konvensional yang menetapkan biaya seragam per kelas, Bahnsen mengusulkan kerangka di mana setiap sampel individu memiliki biaya misklasifikasi tersendiri yang bergantung pada karakteristik sampel tersebut (*example-dependent costs*).

1. Kerangka Biaya *Example-Dependent*

Diberikan dataset $D = \{(x_i, y_i)\}$ dengan $i = 1, 2, \dots, N$, di mana $x_i \in \mathbb{R}^d$ adalah vektor fitur dan $y_i \in \{0, 1\}$ adalah label kelas (1 = stunting, 0 = non-stunting). Untuk setiap sampel x_i , didefinisikan matriks biaya *example-dependent* sebagai (Correa Bahnsen et al., 2015):

$$C_i = \begin{bmatrix} C_{TN_i} & C_{FP_i} \\ C_{FN_i} & C_{TP_i} \end{bmatrix} \quad (1)$$

di mana:

C_{TP_i} = biaya ketika sampel positif (*stunting*) diklasifikasikan dengan benar sebagai positif

C_{TN_i} = biaya ketika sampel negatif (*non-stunting*) diklasifikasikan dengan benar sebagai negatif

C_{FP_i} = biaya ketika sampel negatif salah diklasifikasikan sebagai positif (*false positive*)

C_{FN_i} = biaya ketika sampel positif salah diklasifikasikan sebagai negatif (*false negative*)

Dalam konteks stunting, biaya *false negative* (C_{FN}) ditetapkan jauh lebih besar dari pada C_{FP} karena kegagalan mendeteksi anak yang sesungguhnya mengalami *stunting* memiliki konsekuensi klinis yang lebih serius dibandingkan salah mendiagnosis anak sehat sebagai *stunting* (Black et al., 2013).

2. Total Biaya Klasifikasi

Total biaya yang dihasilkan oleh sebuah pengklasifikasi f pada dataset D didefinisikan oleh (Correa Bahnsen et al., 2015) sebagai:

$$\text{Cost}(f, D) = \sum_{i=1}^N [y_i \cdot \hat{y}_i \cdot C_{TP_i} + y_i \cdot (1 - \hat{y}_i) \cdot C_{FN_i} + (1 - y_i) \cdot \hat{y}_i \cdot C_{FP_i} + (1 - y_i) \cdot (1 - \hat{y}_i) \cdot C_{TN_i}] \quad (2)$$

di mana $\hat{y}_i \in \{0, 1\}$ adalah prediksi model untuk sampel ke- i . Tujuan optimasi adalah meminimalkan total biaya ini:

$$f^* = \arg \min_f \text{Cost}(f, D) \quad (3)$$

3. Mekanisme *Rejection Sampling* Berbasis Biaya

Mekanisme inti dari Cost-Sensitive *Rejection Sampling* adalah pengambilan sampel ulang dari distribusi mayoritas berdasarkan probabilitas penerimaan yang diturunkan dari rasio biaya setiap sampel (Correa Bahnsen et al., 2015). Bahnsen mendefinisikan probabilitas penerimaan sampel x_i sebagai:

$$q(x_i) = \frac{w(x_i)}{\max_j w(x_j)} \quad (4)$$

di mana $w(x_i)$ adalah fungsi bobot biaya untuk sampel x_i yang didefinisikan sebagai:

$$w(x_i) = C_{FN_i} \cdot P(y_i = 1) + C_{FP_i} \cdot P(y_i = 0) \quad (5)$$

dengan $P(y_i = 1)$ dan $P(y_i = 0)$ adalah probabilitas posterior kelas positif dan negatif untuk sampel x_i yang diestimasi menggunakan model awal (misalnya J48Consolidate, Credal Decision Tree, C4.5, LADTree, Logistic Model Tree (LMT) dan Hoeffding Tree) (Correa Bahnsen et al., 2015). Sampel x_i kemudian diterima ke dalam dataset latih akhir jika:

$$u \leq q(x_i), u \sim \text{Uniform}(0,1) \quad (6)$$

dan ditolak (rejected) apabila sebaliknya. Proses ini secara efektif mengurangi sampel mayoritas yang memiliki bobot biaya rendah, yaitu sampel yang mudah diklasifikasikan dan jauh dari batas keputusan, sambil mempertahankan sampel yang berada di wilayah kritis dekat batas keputusan antarkelas (Correa Bahnsen et al., 2015).

4. Fungsi Objektif Terintegrasi

Setelah proses rejection sampling menghasilkan dataset yang lebih seimbang D' , algoritma klasifikasi dilatih menggunakan fungsi objektif yang dimodifikasi dengan bobot biaya (Correa Bahnsen et al., 2015):

$$\mathcal{L}_{CSRS} = \sum_{i \in D'} w(x_i) \cdot \ell(y_i, \hat{y}_i) \quad (7)$$

di mana $\ell(y_i, \hat{y}_i)$ adalah fungsi kerugian dasar (misalnya *log-loss* atau *hinge loss* tergantung algoritma yang digunakan). Integrasi simultan antara rejection sampling pada level data dan pembobotan biaya pada level algoritma inilah yang menjadikan CSRS unggul dibandingkan pendekatan tunggal (Correa Bahnsen et al., 2015).

D. Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan empat metrik standar yang diturunkan dari Confusion Matrix seperti *precision, recall, accuracy, F-Measure* (Fawcett, 2006). Namun *F-Measure* menjadi metrik utama dalam pengambilan keputusan untuk menentukan metode sampling dan algoritma terbaik karena mengingat sensitivitasnya terhadap ketidakseimbangan kelas dan secara eksplisit sebagai penyeimbang antara *precision* dan *recall* serta tidak bias terhadap kelas mayoritas (Correa Bahnsen et al., 2015). Sedangkan metrik *accuracy* tetap digunakan sebagai metrik pelengkap dan untuk memberikan gambaran terhadap kinerja semua model (Sokolova & Lapalme, 2009).

III. HASIL DAN PEMBAHASAN

A. Hasil Eksperimen

Eksperimen dilakukan dengan menguji enam algoritma klasifikasi berbasis pohon keputusan, yaitu J48Consolidate, CDT, C4.5, LADTree, LMT, dan Hoeffding Tree, masing-masing dikombinasikan dengan enam metode sampling yaitu Cost-Sensitive Rejection Sampling, Cost-Sensitive Oversampling, SMOTE, Oversampling, Undersampling, dan tanpa sampling (*baseline*). Evaluasi dilakukan menggunakan empat metrik, yaitu Precision, Recall, F-Measure, dan Accuracy. Keseluruhan hasil eksperimen disajikan dalam bentuk tabel secara lengkap. Eksperimen dimulai dengan melakukan pengujian menggunakan 6 algoritma klasifikasi tanpa *sampling* hingga melakukan eksperimen menggunakan *sampling* dan *cost sensitive sampling* yang dapat di lihat pada tabel 2 berikut:

Tabel 2. Hasil Eksperimen Keseluruhan

Algoritma	Precision	Recall	F-Measure	Accuracy
Tanpa Sampling				
J48Consolidate	0,824	0,784	0,797	0,7837
CDT	0,843	0,854	0,841	0,8543
C4.5	0,843	0,854	0,839	0,8536
LADTree	0,847	0,857	0,845	0,8569
LMT	0,846	0,855	0,839	0,8553
Hoeffding Tree	0,834	0,847	0,831	0,8470
Undersampling				
US+J48Consolidate	0,783	0,766	0,772	0,7662
US+CDT	0,799	0,807	0,797	0,8069
US+C4.5	0,806	0,812	0,800	0,8124
US+LADTree	0,810	0,816	0,803	0,8156
US+LMT	0,801	0,807	0,792	0,8068
US+Hoeffding Tree	0,800	0,801	0,780	0,8011
Oversampling				
OS+J48Consolidate	0,837	0,833	0,833	0,8332
OS+CDT	0,809	0,808	0,808	0,8081
OS+C4.5	0,837	0,833	0,833	0,8332
OS+LADTree	0,769	0,763	0,762	0,7631
OS+LMT	0,849	0,844	0,844	0,8443
OS+Hoeffding Tree	0,746	0,741	0,739	0,7407
SMOTE				
SMT+J48Consolidate	0,874	0,871	0,871	0,8712
SMT+CDT	0,872	0,869	0,869	0,8694

SMT+C4.5	0,874	0,871	0,871	0,8712
SMT+LADTree	0,818	0,816	0,815	0,8156
SMT+LMT	0,870	0,869	0,869	0,8687
SMT+Hoeffding Tree	0,744	0,740	0,739	0,7399
Cost-Sensitive Oversampling				
CSO+J48Consolidate	0,821	0,779	0,793	0,7792
CSO+CDT	0,839	0,850	0,837	0,8504
CSO+C4.5	0,845	0,855	0,840	0,8547
CSO+LADTree	0,848	0,858	0,846	0,8580
CSO+LMT	0,840	0,851	0,834	0,8511
CSO+Hoeffding Tree	0,835	0,847	0,830	0,8468
Cost-Sensitive Rejection Sampling				
CSRS+J48Consolidate	0,957	0,763	0,840	0,7625
CSRS+CDT	0,957	0,969	0,957	0,9689
CSRS+C4.5	0,963	0,970	0,960	0,9703
CSRS+LADTree	0,954	0,968	0,957	0,9679
CSRS+LMT	0,957	0,969	0,958	0,9688
CSRS+Hoeffding Tree	0,938	0,968	0,953	0,9684

B. Analisis Kinerja Cost-Sensitive Rejection Sampling

Untuk memberikan gambaran menyeluruh tentang kinerja setiap metode sampling, Tabel 3 menyajikan rekapitulasi nilai rata-rata dan *ranking* yang didasari oleh metrik evaluasi F-Measure dari seluruh algoritma klasifikasi untuk masing-masing metode *sampling*.

Tabel 3. Rekapitulasi Rata-rata Kinerja Seluruh Metode Sampling

Metode Sampling	Precision	Recall	F-Measure	Accuracy	Ranking
Cost-Sensitive Rejection Sampling	0,954	0,935	0,938	0,935	1
SMOTE	0,842	0,839	0,839	0,839	2
Tanpa Sampling	0,840	0,842	0,832	0,842	3
Cost-Sensitive Oversampling	0,838	0,840	0,830	0,840	4
Oversampling	0,808	0,804	0,803	0,804	5
Undersampling	0,800	0,802	0,791	0,802	6

Berdasarkan Tabel 3, terlihat dengan jelas bahwa **Cost-Sensitive Rejection Sampling (CSRS)** secara **konsisten mendominasi seluruh metode sampling** pada semua metrik evaluasi yang digunakan. CSRS mencapai rata-rata Precision sebesar 0,954, Recall sebesar 0,935, F-Measure sebesar 0,938, dan Accuracy sebesar 0,935. Nilai-nilai ini secara signifikan melampaui metode terbaik kedua, yaitu SMOTE, dengan selisih rata-rata F-Measure sebesar 0,099 (peningkatan relatif 11,8%) dan selisih rata-rata Accuracy sebesar 0,096 (peningkatan relatif 11,4%).

Dominasi CSRS terhadap seluruh metode pembandingan dapat dianalisis lebih mendalam dari aspek konsistensi antar algoritma. CSRS menghasilkan peningkatan F-Measure yang signifikan pada hampir semua algoritma dibandingkan kondisi baseline tanpa sampling. Jika dibandingkan dengan kondisi tanpa sampling, CSRS meningkatkan F-Measure pada: CDT dari 0,841 menjadi 0,957 (peningkatan 13,8%), C4.5 dari 0,839 menjadi 0,960 (peningkatan 14,4%), LADTree dari 0,845 menjadi 0,957 (peningkatan 13,3%), LMT dari 0,839 menjadi 0,958 (peningkatan 14,2%), dan Hoeffding Tree dari 0,831 menjadi 0,953 (peningkatan 14,7%). Pengecualian terjadi pada algoritma J48Consolidate, di mana CSRS justru menghasilkan Accuracy yang lebih rendah (0,7625) dibandingkan baseline (0,7837) meskipun nilai Precision-nya sangat tinggi (0,957). Fenomena ini disebabkan oleh nilai Recall yang rendah pada J48Consolidate dengan CSRS (0,763), mengindikasikan bahwa mekanisme konsolidasi pohon

pada J48Consolidate kurang optimal dalam memanfaatkan distribusi sampel yang dihasilkan oleh rejection sampling, sehingga model cenderung konservatif dalam memprediksi kelas minoritas.

C. Kombinasi Terbaik: C4.5 dengan Cost-Sensitive Rejection Sampling

Berdasarkan seluruh hasil eksperimen, kombinasi C4.5 dengan Cost-Sensitive Rejection Sampling merupakan kombinasi terbaik dalam penelitian ini, menghasilkan nilai Precision 0,963, Recall 0,970, F-Measure 0,960, dan Accuracy 0,9703. Tabel 4 menyajikan perbandingan kinerja C4.5 di bawah seluruh metode sampling untuk mempertegas superioritas CSRS.

Tabel 4. Perbandingan Kinerja Algoritma C4.5 pada Berbagai Metode Sampling

Metode Sampling	Precision	Recall	F-Measure	Accuracy
C4.5+CS Rejection Sampling	0,963	0,970	0,960	0,9703
C4.5+SMOTE	0,874	0,871	0,871	0,8712
C4.5+CS Oversampling	0,845	0,855	0,840	0,8547
C4.5 (Baseline)	0,843	0,854	0,839	0,8536
C4.5+Oversampling	0,837	0,833	0,833	0,8332
C4.5+Undersampling	0,806	0,812	0,800	0,8124

Berdasarkan tabel di atas, keseimbangan antara *Precision* dan *Recall* yang sangat baik ini tercermin dari nilai *F-Measure* sebesar 0,960, yang merupakan nilai tertinggi dari seluruh 36 kombinasi algoritma dan metode sampling yang diuji dalam penelitian ini. Sedangkan nilai *Accuracy* sebesar 0,9703 pada kombinasi C4.5 + CSRS berarti bahwa dari setiap 100 balita yang diperiksa, model berhasil mengklasifikasikan 97 di antaranya dengan benar. Lebih penting lagi, nilai Recall sebesar 0,970 menunjukkan bahwa model mampu mendeteksi 97% kasus stunting yang sesungguhnya ada dalam data, yang merupakan capaian sangat krusial dalam konteks klinis di mana kegagalan deteksi stunting berdampak langsung pada keterlambatan intervensi gizi (Black et al., 2013). Nilai Precision sebesar 0,963 juga mengindikasikan bahwa dari seluruh prediksi stunting yang dibuat model, 96,3% di antaranya adalah kasus stunting yang benar-benar nyata, sehingga meminimalkan beban pemeriksaan lanjutan yang tidak perlu bagi tenaga kesehatan.

IV. KESIMPULAN

Hasil penelitian ini secara empiris mengkonfirmasi hipotesis bahwa pendekatan *hybrid* yang mengintegrasikan penanganan ketidakseimbangan pada level data (*rejection sampling*) dan level algoritma (*cost-sensitive learning*) secara simultan menghasilkan kinerja yang lebih unggul dibandingkan pendekatan tunggal mana pun. Temuan ini konsisten dengan argumen teoritis Bahnsen bahwa penetapan biaya yang bersifat *example-dependent* mampu menangkap heterogenitas risiko misklasifikasi yang tidak dapat ditangkap oleh pendekatan berbasis kelas yang seragam. Keunggulan CSRS atas SMOTE juga relevan secara praktis. SMOTE menghasilkan sampel sintesis melalui interpolasi linier antara sampel minoritas yang ada, yang pada dataset stunting dapat menghasilkan profil balita yang secara klinis tidak realistis, misalnya kombinasi nilai tinggi badan, berat badan, dan usia yang tidak lazim ditemukan secara biologis. CSRS menghindari masalah ini sepenuhnya karena hanya bekerja dengan sampel nyata yang sudah tersedia dalam dataset.

DAFTAR PUSTAKA

- Abellan, J., & Moral, S. (2003). Building Classification Trees Using the Total Uncertainty Criterion. *INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS*, 18, 1215–1225. <https://doi.org/10.1002/int.10143>
- Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. A., Christian, P., de Onis, M., Ezzati, M., Grantham-McGregor, S., Katz, J., Martorell, R., & Uauy, R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, 382(9890), 427–451. [https://doi.org/10.1016/S0140-6736\(13\)60937-X](https://doi.org/10.1016/S0140-6736(13)60937-X)
- Correa Bahnsen, A., Aouada, D., & Ottersten, B. (2015). *Ensemble of Example-Dependent Cost-Sensitive Decision Trees*. 1–13.

- Danaei, M., Dehghani Firouzabadi, M., Allahverdipour, Y., Henry, K. D., MacIntyre, S., & Maleta, M. (2016). Risk factors for childhood stunting in 137 developing countries: A comparative risk assessment analysis at global, regional, and country levels. *PLOS Medicine*, 13(11), e1002164. <https://doi.org/10.1371/journal.pmed.1002164>
- El-muslih, S. A., Vionanda, D., Amalita, N., & Salma, A. (2023). Comparison of Error Rate Prediction Methods in Classification Modeling with the CHAID Method for Imbalanced Data. *UNP Journal Of Statistics And Data Science*, 1(4), 321–328. <https://doi.org/https://doi.org/10.24036/ujsds/vol1-iss4/81>
- Elkan, C. (2001). The foundations of cost-sensitive learning. *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI)*, 2, 973–978.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). *Learning from Imbalanced Data Sets*. Springer.
- Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications*, 73, 220–239. <https://doi.org/10.1016/j.eswa.2016.12.035>
- He, H., & Garcia, E. A. (2009). *Learning from Imbalanced Data*. 21(9), 1263–1284.
- He, H., & Ma, Y. (2013). *He - Imbalanced Learning - 2014*. 216.
- Holmes, G., Pfahringer, B., Kirkby, R., Frank, E., & Hall, M. (2002). Multiclass Alternating Decision Trees. *Machine Learning: ECML 2002*, 161–172.
- Hulten, G., Spencer, L., & Pedro, D. (2001). Mining Time-Changing Data Streams. *KDD '01: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 97–106. <https://doi.org/https://doi.org/10.1145/502512.502529>
- Ibarguren, I., Pérez, J. M., Muguerza, J., Gurrutxaga, I., & Arbelaitz, O. (2015). Coverage-based resampling: Building robust consolidated decision trees. *KNOWLEDGE-BASED SYSTEMS*, 79, 51–67. <https://doi.org/10.1016/j.knosys.2014.12.023>
- Ilanda, Y., Vionanda, D., Kurniawati, Y., & Fitria, D. (2023). Comparison of Error Rate Prediction Methods of C4 . 5 Algorithm for Imbalanced Data. *UNP Journal Of Statistics And Data Science*, 1(4), 240–247. <https://doi.org/https://doi.org/10.24036/ujsds/vol1-iss4/89>
- Joseph, V. R. (2022). *Optimal ratio for data splitting*. February, 531–538. <https://doi.org/10.1002/sam.11583>
- Landwehr, N., Hall, M., & Frank, E. (2005). Logistic Model Trees *. *Machine Learning*, 59, 161–205. <https://doi.org/https://doi.org/10.1007/s10994-005-0466-3>
- Menzies, T., Hu, Y., & Distler, J. (2007). Ask your data: Better data mining for better decisions. *IEEE Software*, 24(4), 14–17. <https://doi.org/10.1109/MS.2007.103>
- Nurjanah, S., Andreswari, R., & Hasibuan, M. A. (2022). Klasifikasi status gizi balita untuk deteksi stunting menggunakan algoritma Random Forest dengan teknik SMOTE. *Jurnal Ilmu Komputer Dan Informasi*, 15(2), 112–122. <https://doi.org/10.21609/jiki.v15i2.1043>
- Organization, W. H. (2006). *WHO Child Growth Standards: Length/Height-for-Age, Weight-for-Age, Weight-for-Length, Weight-for-Height and Body Mass Index-for-Age: Methods and Development*. WHO Press.
- PPN/Bappenas, K. (2020). *Rencana Pembangunan Jangka Menengah Nasional (RPJMN) 2020–2024*. Badan Perencanaan Pembangunan Nasional.
- Prasetyo, B., Kuswardani, D., & Fauzi, A. (2022). Deteksi dini stunting pada balita menggunakan XGBoost dengan pendekatan cost-sensitive learning. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, 11(3), 201–210. <https://doi.org/10.22146/jnteti.v11i3.3782>
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers.

- Robert, C. P., & Casella, G. (2004). *Monte Carlo Statistical Methods* (2nd ed.). Springer. <https://doi.org/10.1007/978-1-4757-4145-2>
- Sinaga, B. A., Vionanda, D., Permana, D., & Salma, A. (2023). Comparison of Error Rate Prediction Methods in Binary Logistic Regression Modeling for Imbalanced Data. *UNP Journal Of Statistics And Data Science*, 1(4), 361–368. <https://doi.org/https://doi.org/10.24036/ujsds/vol1-iss4/86>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques* (4th ed.). Morgan Kaufmann.
- Yulianti, A., Fitri, F., Amalita, N., & Vionanda, D. (2023). The SMOTE Application of CART Methods for Coping Imbalanced Data in Classifying Status Work on Labor Force in the City of Padang. *UNP Journal Of Statistics And Data Science*, 1(4), 172–179. <https://doi.org/https://doi.org/10.24036/ujsds/vol1-iss3/12>
- Zullani, L., Vionanda, D., & Fitria, D. (2023). Comparison of Error Rate Prediction in CART for Imbalanced Data. *UNP Journal Of Statistics And Data Science*, 1(5), 464–472. <https://doi.org/https://doi.org/10.24036/ujsds/vol1-iss5/117>