



SKRIPSI

**OPTIMASI XGBOOST DENGAN HYPERPARAMETER
GRIDSEARCH DAN RANDOMSEARCH PADA
KLASIFIKASI GIZI BALITA**

AVI OKTAVIANI

NPM 21081010227

DOSEN PEMBIMBING

Dr. Basuki Rahmat, S. Si. MT.

Dr. Ir. Kartini, S.Kom., MT.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**



SKRIPSI

**OPTIMASI XGBOOST DENGAN HYPERPARAMETER
GRIDSEARCH DAN RANDOMSEARCH PADA
KLASIFIKASI STATUS GIZI BALITA**

Avi Oktaviani

NPM 21081010227

DOSEN PEMBIMBING

Dr. Basuki Rahmat, S. Si. M.T.

Dr. Ir.Kartini, S.Kom, M.T.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
SURABAYA
2025**

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

OPTIMASI XGBOOST DENGAN HYPERPARAMETER GRIDSEARCH DAN RANDOMSEARCH PADA KLASIFIKASI STATUS GIZI BALITA

Oleh :

AVI OKTAVIANI

NPM. 21081010227

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional "Veteran" Jawa Timur Pada tanggal 14 November 2025

Menyutujui

Dr. Basuki Rahmat, S. Si, M.T.

NIP 19690723 2021211002

: (Pembimbing I)

Dr. Ir. Kartini, S.Kom, M.T.

NIP 19611110 199103 2 001

: (Pembimbing II)

Dr. Eng. Ir. Anggraini Puspita Sari, ST., M.T.

NPT 222198 60 816400

: (Ketua Penguji)

Afina Lina Nurlaili, S.Kom., M.Kom.

NIP 1993121 3202203 2 010

: (Penguji II)

Mengetahui

Dekan Fakultas Ilmu Komputer

Prof. Dr. Ir. Novirina Hendrasarie, MT.

NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

LEMBAR PERSETUJUAN

**OPTIMASI XGBOOST DENGAN HYPERPARAMETER GRIDSEARCH DAN
RANDOMSEARCH PADA KLASIFIKASI STATUS GIZI BALITA**

Oleh :

AVI OKTAVIANI

NPM. 21081010227



Menyetujui,

Koordinator Program Studi Informatika
Fakultas Ilmu Komputer

Fetty Tri Anggraeny, S.Kom. M.Kom

NIP. 19820211 2021212 005

SURAT PERNYATAAN BEBAS PLAGIASI

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini :

Nama : Avi Oktaviani
NPM : 21081010227
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya mengatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiasi pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksa dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 3 Desember 2025

Yang membuat pernyataan



Avi Oktaviani

NPM. 21081010227

ABSTRAK

Nama Mahasiswa /NPM : Avi Oktaviani / 21081010227
Judul Skripsi : Optimasi XGBoost dengan *Hyperparameter*
GridSearch dan RandomSearch pada Klasifikasi
Status Gizi Balita
Dosen Pembimbing : 1. Dr. Basuki Rahmat, S. Si, M.T.
2. Dr. Ir. Kartini, S.Kom., M.T.

Penelitian ini bertujuan untuk membandingkan kinerja Grid Search dan Random Search dalam mengklasifikasikan status gizi pada balita dengan 6 kelas, yaitu gizi buruk, gizi kurang, gizi baik, berisiko gizi lebih, gizi lebih, dan obesitas. Tahapan awal penelitian meliputi proses pre-processing data, yang mencakup penghapusan kolom tidak relevan, penghapusan data duplikat dan missing value, serta transformasi data. Untuk mengatasi ketidakseimbangan kelas, dilakukan proses smote pada data pelatihan. Eksperimen dilakukan dengan menguji berbagai skema proporsi pembagian data 80:20, 75:25, dan 70:30, serta beberapa parameter utama dari kedua model, seperti *learning rate* 0.05, 0.1, dan 0.2, *max depth* 4, 6, dan 8, dan *n_estimator* 50, 100, dan 200, *subsample* 0.7, 0.9, dan 1.0. Hasil pengujian menunjukkan bahwa kombinasi parameter terbaik untuk kedua model adalah proporsi data 80:20, *learning rate* 0.1, *max depth* 6, *n_estimator* 200, dan *subsample* 0.7. Model kemudian dievaluasi menggunakan pendekatan optimasi GridSearch dan RandomSearch, evaluasi metrik kinerja akurasi, presisi, *recall*, dan F1-Score, serta *cross-validation* 3, 5 dan 10 fold. Hasil pengujian menunjukkan bahwa model XGBoost tanpa optimasi menghasilkan akurasi 0,869, presisi 0,802, *recall* 0,831, dan F1-score 0,814. Setelah dilakukan optimasi dengan Grid Search, diperoleh akurasi 0,877, presisi 0,817, *recall* 0,813, dan F1-score 0,812, sedangkan Random Search memberikan akurasi 0,869, presisi 0,796, *recall* 0,807, dan F1-score 0,812. Dari sisi efisiensi, Grid Search membutuhkan waktu 388,40 detik, sementara Random Search hanya 123,73 detik dengan kinerja yang hampir setara. Dengan demikian, Optimasi XGBoost terbukti memberikan peningkatan kinerja yang moderat, di mana Grid Search sedikit lebih unggul pada akurasi dan presisi, sedangkan Random Search lebih efisien secara waktu dan lebih sesuai ketika sumber daya komputasi terbatas. Hasil model kemudian diintegrasikan ke dalam aplikasi web sebagai sistem bantu prediksi status gizi balita.

Kata kunci : XGboost, GridSearch, RandomSearch, Klasifikasi, Status Gizi Balita

ABSTRACT

Nama Mahasiswa / NPM : Avi Oktaviani / 21081010227
Judul Skripsi : Optimizing XGBoost with Grid Search and
Random Search Hyperparameter Tuning for
Classification of Nutritional Status in Children
Under Five
Dosen Pembimbing : 1. Dr. Basuki Rahmat, S. Si, M.T.
2. Dr. Ir. Kartini, S.Kom., M.T.

This study aims to compare the performance of Grid Search and Random Search in classifying the nutritional status of children under five into six classes: severe undernutrition, undernutrition, normal nutrition, at risk of overweight, overweight, and obesity. To address class imbalance, SMOTE is applied to the training data. Experiments are conducted by testing various train–test split proportions of 80:20, 75:25, and 70:30, as well as several key parameters of the model, such as *learning rates* of 0.05, 0.1, and 0.2, *max depths* of 4, 6, and 8, numbers of estimators of 50, 100, and 200, and *subsample* values of 0.7, 0.9, and 1.0. The results show that the best parameter combination for both models is a data proportion of 80:20, *learning rate* of 0.1, *max depth* of 6, 200 estimators, and *subsample* of 0.7. The model is then evaluated using Grid Search and Random Search optimization approaches, performance metrics including accuracy, *precision*, *recall*, and F1-score, as well as 3, 5 and 10-fold *cross-validation*. The experimental results indicate that the XGBoost model without optimization achieves an accuracy of 0.869, *precision* of 0.802, *recall* of 0.831, and F1-score of 0.814. After optimization with Grid Search, the model attains an accuracy of 0.877, *precision* of 0.817, *recall* of 0.813, and F1-score of 0.812, while Random Search yields an accuracy of 0.869, *precision* of 0.796, *recall* of 0.807, and F1-score of 0.812. In terms of efficiency, Grid Search requires 388.40 seconds, whereas Random Search takes only 123.73 seconds with nearly comparable performance. Thus, XGBoost optimization is proven to provide a moderate improvement in performance, with Grid Search performing slightly better in accuracy and *precision*, while Random Search is more time-efficient and more suitable when computational resources are limited. The final model is then integrated into a web application as a decision-support system for predicting the nutritional status of children under five.

Keywords: XGBoost, Grid Search, Random Search, *Classification*, Nutritional Status of Children Under Five

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan dengan judul “Optimasi XGBoost dengan Hyperparameter Gridsearch dan Randomsearch pada Klasifikasi Status Gizi Balita” dengan baik. Penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT selaku Rektor Universitas Pembangunan Nasional "Veteran" Jawa Timur.
2. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Ibu Fetty Tri Anggraeny, S.Kom. M. Kom selaku Koordinator Program Studi Informatika Fakultas Ilmu Sosial Dan Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur
4. Bapak Dr. Basuki Rahmat, S.Si, M.T., selaku dosen pembimbing pertama, dan Ibu Ir. Kartini, S.Kom, M.Kom, S.Kom., M.T.I., M.C.F., selaku dosen pembimbing kedua yang telah memberikan arahan, bimbingan, dan motivasi selama proses penyusunan tugas akhir ini.
5. Ibu Dr. Eng. Ir. Anggraini Puspita Sari, S.T., M.T. dan Ibu Afina Nurlaili, S.Kom, M.Kom., selaku dosen penguji yang telah meluangkan waktu, memberikan perhatian, serta menyampaikan masukan berharga dalam proses ujian tugas akhir.
6. Orang tua tersayang Alm. Bapak Mochammad Hani, Bapak Suryono Kristanto Ibu Anita Ikhtiarsi, atas segala usaha, doa, ikhtiar, kasih sayang, serta dukungan yang senantiasa menyertai penulis hingga berada pada titik ini. Tetap menyayangi dengan utuh meskipun penulis mengalami kegagalan. Ucapan terima kasih juga penulis sampaikan kepada ketiga saudara penulis, Adam, Yusuf, dan Hanita, yang selalu menjadi penghibur dan sumber semangat ketika penulis merasa tidak cukup kuat.
7. Anggota grup chat “2025 Menuju GG”, yaitu Pista, Depi, Jalpa, dan Nanda, yang telah menjadi tempat berbagi keluh kesah selama masa perkuliahan. Terima kasih telah menemani perjalanan penulis sejak awal hingga akhirnya dapat menyelesaikan studi bersama-sama.

8. Sahabat tersayang Azza Syahida dan Devani Mutiara, atas dukungan yang tidak pernah putus meskipun jarak memisahkan. Terima kasih telah bersedia direpotkan oleh penulis dan tetap hadir sebagai sumber kekuatan.
9. Seluruh teman-teman angkatan 2021 yang telah banyak membantu, berbagi informasi, serta memberikan dukungan selama proses penyusunan tugas akhir ini.

Surabaya, 3 Desember 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	vii
ABSTRAK	ix
KATA PENGANTAR.....	xiii
DAFTAR TABEL	xix
DAFTAR GAMBAR.....	xxi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2. Rumusan Masalah.....	4
1.3. Batasan Masalah	4
1.4. Tujuan.....	4
1.5. Manfaat.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1. Penelitian terdahulu	7
2.2. Status Gizi Balita	12
2.3 Klasifikasi	14
2.4 <i>Preprocessing data</i>	15
2.4.1. Data Cleaning	15
2.4.2. Encoding.....	15
2.4.3 Standarisasi.....	15
2.5 SMOTE.....	16
2.6 Boosting.....	17
2.7 XGBoost (<i>Extreme Gradient Boosting</i>).....	18
2.8 <i>Tuning Hyperparameter</i>	22
2.9 <i>Grid Search</i>	22
2.10 <i>Random Search</i>	23
2.11 <i>Confusion matrix</i>	23
2.11.1 Akurasi	24
2.11.2 <i>Precision</i>	24
2.11.3 <i>Recall</i>	24
2.11.4 <i>F1-Score</i>	25
BAB III METODOLOGI PENELITIAN	27
3.1 Tahapan Penelitian.....	27
3.2 Studi Literatur	28
3.3 Analisis Desain	28
3.4 Akuisisi Data.....	29

3.5 <i>Preprocessing data</i>	30
3.5.1 <i>Data Cleaning</i>	30
3.5.2 <i>Encoding</i>	31
3.5.3 Standarisasi.....	32
3.5.4. Pembagian data latih dan data uji	34
3.5.5. SMOTE	34
3.6 Model	35
3.6.1 XGBoost dengan Grid Search	36
3.6.2. XGBoost dengan Random Search	45
3.7 Evaluasi Model	53
3.8. Skenario Pengujian	57
3.9. Implementasi	58
BAB IV HASIL DAN PEMBAHASAN.....	61
4.1 Persiapan Data	61
4.1.2 <i>Library</i>	61
4.1.2 Membaca dataset	62
4.2 Pengolahan Data	63
4.2.2 <i>Cleaning Data</i>	63
4.2.3 <i>Encoding data</i>	67
4.2.4 Standarisasi.....	68
4.2.5 <i>Split data</i> latih dan data uji	68
4.3 SMOTE.....	70
4.4 Pelatihan Model.....	73
4.4.2 XGBoost.....	73
4.4.3 XGBoost Optimasi Grid Search	74
4.4.4 XGBoost Optimasi Random Search	76
4.5 Pengujian	77
4.5.2 Pengujian <i>Split data</i>	78
4.5.3 Pengujian Parameter	85
4.5.4 Pengujian Grid Search	108
4.5.5 Pengujian Random Search.....	113
4.5.6 Pengujian XGBoost tanpa Optimasi	116
4.6 Hasil Skenario Pengujian.....	119
4.6.2 Hasil pengujian <i>split data</i>	119
4.6.3 Hasil pengujian parameter	120
4.6.4 Perbandingan model terbaik	122
4.7 Hasil Implementasi	126
BAB V KESIMPULAN DAN SARAN	131

5.1. Kesimpulan.....	131
5.2. Saran.....	132
Daftar Pustaka	135

DAFTAR TABEL

Tabel 2.1 Kategori Dan Ambang Batas Status Gizi Anak	13
Tabel 3.1 Data sampel balita.....	29
Tabel 3.2. contoh sampel sebelum cleaning.....	30
Tabel 3.3. setelah dilakukan pembersihan missing value	31
Tabel 3.4. contoh hasil dataset setelah encoding	32
Tabel 3.5. Mean dan standar	32
Tabel 3.6. Hasil standarisasi sampel	33
Tabel 3.7. Pembagian Data Latih Dan Uji	34
Tabel 3.9 Tabel Sampel Turunan Pertama Dari Loss Function	38
Tabel 3.10 Sampel setelah dihitung Gradien	39
Tabel 3.11 hitungan hessian semua sampel	40
Tabel 3.12 Perhitungan Bobot ω_i Node Kiri	42
Tabel 3.13 perhitungan bobot ω_i Node kanan.....	42
Tabel 3.14 Prediksi Iterasi Ke-1 Grid Search	43
Tabel 3.15. hasil nilai probabilitas semua sampel iterasi 1 Grid Search.....	44
Tabel 3.16. nilai probabilitas iterasi ke-2 Grid Search	44
Tabel 3.17 Tabel Sampel Turunan Dari Loss	47
Tabel 3.18 Sampel setelah dihitung Gradien	47
Tabel 3.19 Sampel hitungan hessian.....	48
Tabel 3.20 Perhitungan Bobot ω_i Node Kiri	50
Tabel 3.21 perhitungan bobot ω_i Node kanan.....	50
Tabel 3.22 prediksi Iterasi Ke-1 Random Search	51
Tabel 3.23 hasil nilai probabilitas semua sampel iterasi 1 Random Search	52
Tabel 3.24 nilai probabilitas iterasi ke-2 Random Search	52
Tabel 3.25 Prediksi Kelas	53
Tabel 3.26 Total Prediksi Sampel.....	53
Tabel 3.27 Confusion matrix	54
Tabel 3.28 Hasil laporan klasifikasi model.....	56
Tabel 3.29 Skenario Pengujian Model.....	57
Tabel 4.1. Pengujian split data	84
Tabel 4.2 Hasil pengujian $n_estimator$	91
Tabel 4.3. Hasil pengujian max_depth	97
Tabel 4.4. Hasil pengujian $learning_rate$	102

Tabel 4.5 pengujian subsample	107
Tabel 4.10 Hasil pengujian split data.....	120
Tabel 4.11 Hasil pengujian parameter	121
Tabel 4.12 Hasil perbandingan model terbaik	122
Tabel 4.13 Hasil pengujian model	123

DAFTAR GAMBAR

Gambar 2.1 Grafik Pengukuran Gizi Balita Laki- Laki 24-60	13
Gambar 2.2 Grafik Pengukuran Gizi Balita Perempuan Usia 24-60	14
Gambar 2.3. Ilustrasi Boosting Ensemble Learning	18
Gambar 3.1. Diagram Alur Tahapan Penelitian	27
Gambar 3.2 Diagram Alur Desain Model	28
Gambar 3.3 Alur proses XGBoost	36
Gambar 3.4 Alur proses Grid Search	37
Gambar 3.5 Alur Proses Random Search	45
Gambar 3.7. wireframe sebelum prediksi	59
Gambar 3.8. Tampilan hasil prediksi	60
Gambar 4.1. Tampilan 5 Baris Pertama Dataset	63
Gambar 4.2. Missing Value Sampel	63
Gambar 4.4. Diagram Batang Sebelum Duplikat Dihapus	65
Gambar 4.5 Diagram Batang Jumlah Sampel Setelah Duplikat Dihapus	66
Gambar 4.6 Informasi Dataframe Setelah Dilakukan Cleaning Data	66
Gambar 4.7 Dataframe Setelah Encoding	67
Gambar 4.8 Hasil Akhir Data Setelah Distandarisasi	68
Gambar 4.9 Hasil pembagian data 70:30	69
Gambar 4.10 Hasil pembagian data 75:25	70
Gambar 4.11 Hasil pembagian data 80:20	70
Gambar 4.12 Hasil SMOTE 70:30	71
Gambar 4.13 Hasil SMOTE 75:25	72
Gambar 4.14 Hasil SMOTE 80:20	72
Gambar 4.15 <i>Confusion matrix</i> latih dan uji pengujian 70:30	78
Gambar 4.16 <i>Classification report</i> latih dan uji 70:30	79
Gambar 4.17 <i>Confusion matrix</i> latih dan uji pengujian 75:25	80
Gambar 4.18 <i>Classification report</i> uji pengujian 75:25	81
Gambar 4.19 <i>Confusion matrix</i> latih dan uji <i>split data</i> 80:20	82
Gambar 4.20 <i>Classification report</i> latih dan uji <i>split data</i> 80:20	83
Gambar 4.21 <i>Confusion matrix</i> uji dan latih parameter <i>n_estimator</i> 50	86
Gambar 4.22 <i>Classification report</i> parameter <i>n_estimator</i> 50	87
Gambar 4.23 <i>Confusion matrix</i> uji dan latih <i>n_estimator</i> 100	88
Gambar 4.24 <i>Classification report</i> <i>n_estimator</i> 100	89

Gambar 4.25	<i>Confusion matrix</i> uji dan latih <i>n_estimator</i> 200.....	90
Gambar 4.26	<i>Classification report</i> <i>n_estimator</i> 200.....	91
Gambar 4.27	<i>Confusion matrix</i> uji dan latih <i>max depth</i> 4	93
Gambar 4.28	<i>Classification report</i> uji dan latih <i>max depth</i> 4.....	93
Gambar 4.29	<i>Confusion matrix</i> uji dan latih <i>max depth</i> 6	94
Gambar 4.30	<i>Classification report</i> uji dan latih <i>max depth</i> 6.....	95
Gambar 4.31	<i>Confusion matrix</i> uji dan latih <i>max depth</i> 8	96
Gambar 4.32	<i>Classification report</i> uji dan latih <i>max depth</i> 8.....	96
Gambar 4.33.	<i>Confusion matrix</i> latih dan uji <i>learning_rate</i> 0.05	98
Gambar 4.34	<i>Classification report</i> latih dan uji <i>learning_rate</i> 0.05.....	99
Gambar 4.35	<i>Confusion matrix</i> latih dan uji <i>learning_rate</i> 0.1	100
Gambar 4.36	<i>Classification report</i> latih dan uji <i>learning_rate</i> 0.1.....	100
Gambar 4.37	<i>Confusion matrix</i> latih dan uji <i>learning_rate</i> 0.2.....	101
Gambar 4.38	<i>Classification report</i> latih dan uji <i>learning_rate</i> 0.2.....	102
Gambar 4.39	<i>Confusion matrix</i> latih dan uji <i>subsample</i> 0.7	103
Gambar 4.40	<i>Classification report</i> latih dan uji <i>subsample</i> 0.7.....	104
Gambar 4.41	<i>Confusion matrix</i> latih dan uji <i>subsample</i> 0.9	105
Gambar 4.42	<i>Classification</i> latih dan uji <i>subsample</i> 0.9.....	105
Gambar 4.43	<i>Confusion matrix</i> latih dan uji <i>subsample</i> 1.0	106
Gambar 4.44.	<i>Classification report</i> latih dan uji <i>subsample</i> 1.0.....	107
Gambar 4.49	<i>Confusion matrix</i> latih random search	113
Gambar 4.50	<i>Confusion matrix</i> uji random search	114
Gambar 4.51	<i>Classification report</i> latih random search	115
Gambar 4.52	<i>Classification report</i> uji random search	115
Gambar 4.53	<i>Confusion matrix</i> latih XGBoost tanpa Optimasi	116
Gambar 4.54	<i>Confusion matrix</i> uji XGBoost tanpa Optimasi	117
Gambar 4.55	<i>Classification report</i> latih tanpa Optimasi	118
Gambar 4.56	<i>Classification report</i> uji tanpa Optimasi	119