



SKRIPSI

PERBANDINGAN KINERJA ALGORITMA XGBOOST DAN LIGHTGBM DALAM KLASIFIKASI STATUS GIZI PADA BALITA DENGAN OPTIMASI *HYPERPARAMETER* GRIDSEARCHCV

MUHAMMAD YUSTAF LANA BADRIUL HEGAR
NPM 21081010289

DOSEN PEMBIMBING

Dr. Rr. Ani Dijah Rahajoe, ST., M.Cs.

Muhammad Muharrom Al Haromainy, S.Kom., M.Kom.

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



SKRIPSI

**PERBANDINGAN KINERJA ALGORITMA
XGBOOST DAN LIGHTGBM DALAM
KLASIFIKASI STATUS GIZI PADA BALITA
DENGAN OPTIMASI *HYPERPARAMETER*
GRIDSEARCHCV**

MUHAMMAD YUSTAF LANA BADRIUL HEGAR
NPM 21081010289

DOSEN PEMBIMBING

Dr. Rr. Ani Dijah Rahajoe, ST., M.Cs.

Muhammad Muharrom Al Haromainy, S.Kom, M.Kom

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

LEMBAR PENGESAHAN SKRIPSI

PERBANDINGAN KINERJA ALGORITMA XGBOOST DAN LIGHTGBM DALAM KLASIFIKASI STATUS GIZI PADA BALITA DENGAN OPTIMASI HYPERPARAMETER GRIDSEARCHCV

Oleh :

MUHAMMAD YUSTAF LANA BADRIUL HEGAR

NPM. 21081010289

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. Rr Ani Dijah Rahajoe, ST, M.Cs.
NIP 19730512 200501 2 003

: (Pembimbing I)

Muhammad Muharrom Al Haromainy,
S.Kom., M.Kom.
NIP 19950601 202203 1 006

: (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

: (Anggota Penguji)

Mengetahui

Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

**PERBANDINGAN KINERJA ALGORITMA XGBOOST DAN LIGHTGBM
DALAM KLASIFIKASI STATUS GIZI PADA BALITA DENGAN
OPTIMASI HYPERPARAMETER GRIDSEARCHCV**

Oleh :

MUHAMMAD YUSTAF LANA BADRIUL HEGAR

NPM. 21081010289



**Menyetujui,
Koordinator Program Studi Informatika
Fakultas Ilmu Komputer**

Fetty Tri Anggraeny, S.Kom. M.Kom
NIP. 19820211 202121 2 005

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan dibawah ini :

Nama : Muhammad Yustaf Lana Badriul Hegar
NPM : 21081010289
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 menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

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

Surabaya, 1 Desember 2025

Yang Membuat Pernyataan,



Muhammad Yustaf Lana Badriul Hegar
NPM. 21081010289

ABSTRAK

Nama Mahasiswa / NPM : Muhammad Yustaf Lana Badriul Hegar /
21081010289

Judul Skripsi : Perbandingan kinerja Algoritma XGBosst dan
LightGBM dalam Klasifikasi Status Gizi Pada
Balita dengan Optimasi *Hyperparameter*
GridSearchCV

Dosen Pembimbing : 1. Dr. Rr. Ani Dijah Rahajoe, ST., M.,Cs.
2. Muhammad Muharrom Al Haromainy, S.Kom.,
M.Kom.

Penelitian ini bertujuan untuk membandingkan kinerja algoritma XGBoost dan LightGBM dalam mengklasifikasikan status gizi pada balita dengan empat kelas, yaitu gizi baik, gizi kurang, gizi lebih, dan gizi buruk. Tahapan awal penelitian meliputi proses *pre-processing* data, yang mencakup penghapusan kolom tidak relevan, penghapusan data duplikat, penanganan *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, 70:30, dan 60:40, serta beberapa parameter utama dari kedua model, seperti *learning rate* 0.01, 0.05, 0.1, dan 0.15, *max depth* 3, 4, 5, dan 6, dan *n_estimators* 50, 100, 150, dan 200. Hasil pengujian menunjukkan bahwa kombinasi parameter terbaik untuk kedua model adalah proporsi data 70:30, *learning rate* 0.15, *max depth* 6, dan *n_estimators* 200. Model kemudian dievaluasi menggunakan tiga pendekatan, yaitu optimasi GridSearchCV, evaluasi metrik kinerja akurasi, presisi, *recall*, dan *F1-Score*, serta *cross-validation* 1–10 fold. Hasil penelitian menunjukkan bahwa algoritma LightGBM memberikan performa terbaik pada proporsi data 70:30 dengan nilai akurasi *weighted average*, *presisi*, *recall*, dan *F1-Score* masing-masing sebesar 92%. Sementara itu, algoritma XGBoost pada proporsi yang sama memperoleh akurasi 88%, presisi 89%, *recall* 88%, dan *F1-Score* 88%. Dari sisi waktu komputasi, LightGBM membutuhkan 304 detik, sedangkan XGBoost lebih efisien dengan waktu 156 detik. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa LightGBM unggul dalam performa klasifikasi, sedangkan XGBoost lebih efisien dalam waktu komputasi.

Kata kunci : XGboost, LightGBM, GridSearchCV, Klasifikasi, Status Gizi Balita

ABSTRACT

Student Name / NPM : Muhammad Yustaf Lana Badriul Hegar /
21081010289
Thesis Title : Comparison of the Performance of the XGBosst
and LightGBM Algorithms in Classifying
Nutritional Status in Toddlers with GridSearchCV
Hyperparameter Optimization
Advisor : 1. Dr. Rr. Ani Dijah Rahajoe, ST., M.,Cs.
2. Muhammad Muharrom Al Haromainy, S.Kom.,
M.Kom.

This study aims to compare the performance of the XGBoost and LightGBM algorithms in classifying nutritional status in toddlers into four classes: well-nourished, undernourished, overnourished, and severely malnourished. The initial stages of the study included data preprocessing, which included removing irrelevant columns, removing duplicate data, handling missing values, and data transformation. To address class imbalance, a Smote process was performed on the training data. Experiments were conducted by testing various data split schemes of 80:20, 70:30, and 60:40, as well as several key parameters of both models, such as learning rates of 0.01, 0.05, 0.1, and 0.15, max depths of 3, 4, 5, and 6, and n_estimators of 50, 100, 150, and 200. The test results showed that the best parameter combination for both models was a data split of 70:30, a learning rate of 0.15, a max depth of 6, and n_estimators of 200. The models were then evaluated using three approaches, namely GridSearchCV optimization, evaluation of performance metrics of accuracy, precision, recall, and F1-Score, and 1–10-fold cross-validation. The results showed that the LightGBM algorithm provided the best performance at a data split of 70:30 with accuracy, precision, recall, and F1-Score values of 92% each. Meanwhile, the XGBoost algorithm, with the same proportions, achieved 88% accuracy, 89% precision, 88% recall, and an F1-score of 88%. In terms of computational time, LightGBM required 304 seconds, while XGBoost was more efficient with a time of 156 seconds. Overall, the results of this study indicate that LightGBM excels in classification performance, while XGBoost is more efficient in computational time.

Keywords : XGBoost, LightGBM, GridSearchCV, Classification, Toddler nutrition status

KATA PENGANTAR

Segala puji dan Syukur penulis panjatkan ke hadirat Allah SWT atas limpahan Rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul **“Perbandingan kinerja Algoritma XGBosst dan LightGBM dalam Klasifikasi Status Gizi Pada Balita dengan Optimasi Hyperparameter GridSearchCV”** dengan baik.

Penulis menyampaikan rasa terimakasih yang sebesar-besarnya kepada :

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, M.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 Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa timur.
4. Ibu Dr. Rr. Ani Dijah Rahajoe, ST., M.Cs. selaku Dosen Pembimbing I dan Bapak Muhammad Muharrom Al Haromainy, S.Kom., M.Kom. selaku Dosen Pembimbing 2 yang telah meluangkan waktunya untuk memberikan kesempatan bagi penulis untuk melakukan penelitian, dengan sabar membimbing, memberikan arahan, serta motivasi selama penyusunan skripsi ini.
5. Dosen-dosen Program Studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur yang berperan penting dalam masa studi saya sebagai mahasiswa.
6. Kedua orang tua penulis yang selalu menjadi sumber semangat, doa, dan dukungan tanpa henti, sehingga penulis dapat menyelesaikan penelitian ini dengan lancar dan penuh makna.
7. Kayla Alfaida Asti Kirana dan Kaina Mayang Mahatma Sesha, selaku saudara yang selalu memberi semangat dan menginspirasi penulis agar menjadi orang yang hebat.
8. Pendamping yang selalu memberikan semangat dan dukungan dari awal sampai akhir, untuk penulis, Agatha Diani Putri Saka.

9. Rengga, Irsyad, Aldi, Surya, Ammar, Safi, Feno, Alex, Richard, Dodek, Gany, Dicky, selaku sahabat terbaik penulis yang selalu hadir dalam suka dan duka, memberi semangat tanpa pamrih, serta menjadi tempat serta menjadi tempat berbagi tawa, cerita, dan energi positif di setiap fase perjalanan ini. Terima kasih sudah menjadi bagian dari kisah hidup yang tidak akan pernah penulis lupakan. Kalian bukan cuma teman, tapi rumah dalam bentuk manusia.

Penulis menyadari bahwa penyusunan skripsi ini masih memiliki kekurangan. Untuk itu, masukan dan saran yang bersifat membangun sangat diharapkan guna penyempurnaan di masa yang akan datang. Penulis juga berharap semoga karya ini dapat memberikan manfaat bagi pembaca secara luas, serta menjadi pengalaman berharga bagi penulis sendiri.

Surabaya, 1 Desember 2025

Muhammad Yustaf Lana Badriul Hegar

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI	ii
LEMBAR PERSETUJUAN	iii
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xviii
DAFTAR GAMBAR	xix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA.....	7
2.1. Penelitian Terdahulu.....	7
2.2. Klasifikasi	12
2.3 Gizi Balita	14
2.3. Posyandu	16
2.4. <i>Body mass index</i> (BMI)	16
2.5. Algoritma <i>Light Gradient Boosting Machine</i> (LightGBM).....	17
2.6. Algoritma <i>Extreme Gradient Boosting</i> (XGBoost).....	18
2.7. Optimasi	20
2.8. GridSearchCV	21
2.9. <i>Confusion Matrix</i>	22
2.10. Komparasi.....	24
2.11. <i>Graphical User Interface</i>	25
2.12. <i>Flask</i>	25
BAB III DESAIN DAN IMPLEMENTASI SISTEM	26
3.1 Metodologi Penelitian	26

3.2 Studi Literatur	27
3.3 Pengumpulan Data	27
3.4 <i>Exploratory Data Analysis</i> (EDA)	31
3.5 <i>Data preprocessing</i>	33
3.5.1 <i>Data Cleaning</i>	35
3.5.1 <i>Data Transformation</i>	38
3.5.2 <i>Data Split</i>	39
3.6 Model	40
3.6.2 LightGBM dengan GridSearchCV	41
3.6.1 XGBoost dengan GridSearchCV	45
3.7 Evaluasi Model	50
3.7.1 Evaluasi Model LightGBM	51
3.7.2 Evaluasi Model XGBoost	52
3.8 Skenario Pengujian	53
3.9 Implementasi <i>Website</i>	54
BAB IV PENGUJIAN DAN ANALISIS	56
4.1 <i>Data preprocessing</i>	56
4.1.1 <i>Exploratory Data Analysis</i> (EDA)	56
4.1.2 <i>Dataset Cleaning</i>	59
4.1.3 <i>Data Tranformation</i>	62
4.1.4 <i>Dataset splitting</i>	64
4.1.5 <i>Balancing Data</i>	66
4.2 Pembuatan Model	70
4.2.1 Model <i>Extreme Gradient Boosting</i> (XGBoost)	70
4.2.2 Model <i>Light gradient boosting machine</i> (LightGBM)	71
4.2.3 Model XGboost dengan GridSearchCV	72
4.2.4 Model LightGBM dengan GridSearchCV	73
4.3 Pengujian Model LightGBM	74
4.3.1 Pengujian Proporsi Data	76
4.3.2 Parameter <i>Learning Rate</i>	83
4.3.3 Parameter <i>Max Depth</i>	91
4.3.4 Parameter <i>N Estimators</i>	97

4.3.5 Pengujian Model XGBoost	105
4.3.6 Hasil Perbandingan Model Terbaik.....	106
4.4 Analisis Hasil	109
4.4.1 Model LightGBM Tanpa GridSearchCV	109
4.4.2 Model XGBoost Tanpa GridSearchCV	110
4.4.3 Hasil Perbandingan Model.....	112
4.5 Implementasi <i>Website</i>	113
BAB V KESIMPULAN.....	119
5.1 Kesimpulan	119
5.2 Saran Pengembangan	120
DAFTAR PUSTAKA	121
LAMPIRAN.....	124

DAFTAR TABEL

Tabel 2. 1 Kategori dan Ambang Batas Status Gizi Anak.....	15
Tabel 2. 2 Struktur Confusion Matrix	22
Tabel 3. 1 Atribut Dataset.....	28
Tabel 3. 2 Contoh Dataset status gizi balita	30
Tabel 3. 3 Lima Data Awal Perhitungan Manual	31
Tabel 3. 4 Data Hasil EDA.....	35
Tabel 3. 5 Dataset Sebelum Imputasi	36
Tabel 3. 6 Dataset Setelah Imputasi	37
Tabel 3. 7 Sesudah Tranformasi Data	39
Tabel 3. 8 Lima data Pertama.....	42
Tabel 3. 9 Ringkasan PrediksiModel LightGBM.....	45
Tabel 3. 10 Perhitungan Lima Data Pertama	46
Tabel 3. 11 Ringkasan Hasil Prediksi XGBoost	50
Tabel 3. 12 Skenario Pengujian Model	53
Tabel 4. 1 Hasil Proporsi Data	83
Tabel 4. 2 Hasil Pengujian Learning rate	90
Tabel 4. 3 Hasil Pengujian Max depth	97
Tabel 4. 4 Hasil Pengujian N estimators	104
Tabel 4. 5 Perbandingan Model Terbaik	106
Tabel 4. 6 Hasil Analisis Perbandingan.....	112

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	26
Gambar 3. 2 Preprocessing Data.....	34
Gambar 3. 3 LightGBM dengan GridSearchCV	42
Gambar 3. 4 XGBoost GridSearchCV	46
Gambar 3. 5 Wireframe Website	54
Gambar 4. 1 Struktur dan isi dataset	56
Gambar 4. 2 Distribusi Jenis Kelamin	57
Gambar 4. 3 Jumlah gizi label	58
Gambar 4. 4 Hasil Hapus Kolom Tidak Relevan.....	59
Gambar 4. 5 Hasil Hapus Data Duplikat.....	60
Gambar 4. 6 Hasil Proses Missing Value	61
Gambar 4.7 Ringkasan Proses Data Cleaning	61
Gambar 4. 8 Hasil encoding Pada fitur	63
Gambar 4. 9 Hasil Label Encoding	64
Gambar 4. 10 Hasil Pembagian Data Rasio 90:10.....	65
Gambar 4. 11 Hasil Pembagian Data Rasio 80:20	65
Gambar 4. 12 Hasil Pembagian Data Rasio 70:30.....	66
Gambar 4. 13 Hasil Pembagian Data Rasio 60:40.....	66
Gambar 4. 14 Oversampling Data 90:10	68
Gambar 4.15 Oversampling Data 80:20	68
Gambar 4. 16 Oversampling Data 70:30	69
Gambar 4. 17 Oversampling Data 60:40	70
Gambar 4. 18 Confusion Matrix Proporsi Data 90:10	77
Gambar 4. 19 Clasification Report Proporsi Data 90:10	77
Gambar 4. 20 Confusion Matrix Proporsi Data 80:20	78
Gambar 4. 21 Clasification Report Proporsi Data 80:20	79
Gambar 4. 22 Confusion Matrix Proporsi Data 70:30	80
Gambar 4. 23 Clasification Report Proporsi Data 70:30	81
Gambar 4. 24 Confusion Matrix Proporsi Data 60:40	82
Gambar 4. 25 Clasification Report Proporsi Data 60:40	83

Gambar 4. 26 Confusion Matrix Learning rate 0.1	84
Gambar 4. 27 Clasification report Learning rate 0.1	85
Gambar 4. 28 Confusion Matrix Learning rate 0.01	86
Gambar 4. 29 Clasification Report Learning rate 0.01	87
Gambar 4. 30 Clasification Report Learning rate 0.05	88
Gambar 4. 31 Confusion Matrix Learning Rate 0.15.....	89
Gambar 4. 32 Clasification Report Learning rate 0.15	90
Gambar 4. 33 Confusion Matrix Max Depth 3	91
Gambar 4. 34 Clasification Report Max Depth 3	92
Gambar 4. 35 Confusion Matrix Max Depth 4	93
Gambar 4. 36 Clasification Report Max depth 4	93
Gambar 4. 37 Confusion Matrix Max Depth 5	94
Gambar 4. 38 Clasification Report Max Depth 5	95
Gambar 4. 39 Confusion Matrix Max Depth 6	96
Gambar 4. 40 Clasification report Max depth 6.....	96
Gambar 4. 41 Confusion Matrix N Estimators 50	98
Gambar 4. 42 Clasification Report N Estimators 50.....	99
Gambar 4. 43 Confusion Matrix N Estimators 100	100
Gambar 4. 44 Clasification Report N Estimators 100.....	100
Gambar 4. 45 Confusion Matrix N Estimators 150	101
Gambar 4. 46 Clasification Report N Estimators 150.....	102
Gambar 4. 47 Confusion Matrix N Estimators 200	103
Gambar 4. 48 Clasification Report N Estimators 200.....	104
Gambar 4. 49 Confusion Matrix XGBoost	105
Gambar 4. 50 Clasification Report Model XGBoost.....	106
Gambar 4. 51 Perbandingan Waktu.....	107
Gambar 4. 52 Perbandingan Hasil Model Terbaik.....	108
Gambar 4. 53 Confusion Matrix LigthGBM Tanpa GridSearchCV	109
Gambar 4. 54 Clasification Report LightGBM Tanpa GridSearchCV	110
Gambar 4. 55 Confusion Matrix XGBoost Tanpa GridSearchCV.....	111
Gambar 4. 56 Clasification Report XGBoost Tanpa GridSearchCV	111
Gambar 4. 57 Tampilan Website Hasil Prediksi Gizi Baik	114

Gambar 4. 58 Tampilan Website Hasil dari prediksi Gizi Lebih	115
Gambar 4. 59 Prediksi Tidak Relevan	116
Gambar 4. 60 Range Status Imunisasi	116
Gambar 4. 61 Range Sanitasi Lingkungan.....	117
Gambar 4. 62 Range Pendapatan Rumah Tangga	117