



SKRIPSI

OPTIMASI PREDIKSI PENJUALAN PRODUK FASHION DI TIKTOKSHOP MENGGUNAKAN XGBOOST DAN BAYESIAN OPTIMIZATION

FIRHAN ALI SOFI

21081010326

DOSEN PEMBIMBING

Dr. I Gede Susrama Mas Diyasa, ST., MT.

Achmad Junaidi, 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

OPTIMASI PREDIKSI PENJUALAN PRODUK FASHION DI TIKTOKSHOP MENGGUNAKAN XGBOOST DAN BAYESIAN OPTIMIZATION

FIRHAN ALI SOFI

NPM 21081010326

DOSEN PEMBIMBING

Dr. I Gede Susrama Mas Diyasa, ST., MT.

Achmad Junaidi, S.Kom., M.Kom.

**KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN
TEKNOLOGI UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN
JAWA TIMUR FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA SURABAYA**

2025

Halaman ini sengaja dikosongkan

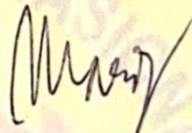
LEMBAR PENGESAHAN

OPTIMASI PREDIKSI PENJUALAN PRODUK FASHION DI TIKTOKSHOP MENGUNAKAN XGBOOST DAN BAYESIAN OPTIMIZATION

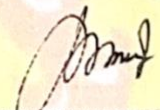
Oleh :
FIRHAN ALI SOFI
NPM. 21081010326

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

Dr. Ir. I Gede Susrama Mas, ST., MT. IPU.
NIP. 19700619 2021211 009


..... (Pembimbing I)

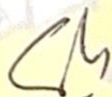
Achmad Junaidi, S.Kom., M.Kom.
NPT. 3 7811 04 0199 1


..... (Pembimbing II)

Budi Nugroho, S.Kom., M.Kom.
NIP. 19800907 202121 1 005


..... (Ketua Penguji)

Eva Yulia Puspaningrum, S.Kom., M.Kom.
NIP. 19890705 202121 2 002


..... (Anggota Penguji)

Mengetahui,
Dekan Fakultas Ilmu Komputer


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

Halaman ini sengaja dikosongkan

LEMBAR PERSETUJUAN

**OPTIMASI PREDIKSI PENJUALAN PRODUK FASHION DI TIKTOKSHOP
MENGUNAKAN XGBOOST DAN BAYESIAN OPTIMIZATION**

Oleh :
FIRHAN ALI SOFI
NPM. 21081010326



Menyetujui,
Koordinator Program Studi Informatika
Fakultas Ilmu Komputer

Fetty Tri Anggraeny, S.Kom. M.Kom
NIP. 19820211 2021212 005

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIAT

Yang bertandatangan di bawah ini:

Nama : Firhan Ali Sofi
NPM : 21081010326
Program : Sarjana (S1)
Program Studi : Informatika
Fakultas : Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Tugas Akhir/Skripsi/Tesis/Disertasi 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 plagiarasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi/Tesis/Disertasi 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, 22 Mei 2025
Yang Membuat Pernyataan,



FIRHAN ALI SOFI
NPM. 21081010326

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM	Firhan Ali Sofi / 21081010326
Judul Skripsi	Optimasi Prediksi Penjualan Produk Fashion di TiktokShop Menggunakan XGBoost dan Bayesian Optimization
Dosen Pembimbing	1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU. 2. Achmad Junaidi, S.Kom. M.Kom

Prediksi penjualan merupakan aspek penting dalam pengelolaan stok dan perencanaan strategi bisnis. Penelitian ini bertujuan untuk mengembangkan model prediksi penjualan produk fashion menggunakan algoritma XGBoost yang dioptimasi dengan Bayesian Optimization. Data yang digunakan dalam penelitian ini adalah data transaksi penjualan dari bisnis fashion D'Busana dari periode 2022 hingga maret 2025, dengan fokus pada platform e-commerce TikTok Shop. Model dikembangkan untuk memprediksi jumlah penjualan berdasarkan faktor-faktor seperti siklus hidup produk, pola musiman, dan ketersediaan stok.

Tahap pengolahan data mencakup preprocessing, pembuatan fitur time-series, serta eksperimen dengan teknik feature engineering untuk meningkatkan performa model. Data dibagi dengan rasio 80:10:10 untuk training, validation, dan testing. Konfigurasi terbaik diperoleh dengan menggunakan teknik Bayesian Optimization dengan rentang $n_estimators$ 50-500 dan 5 fold dalam cross-validation. Parameter yang dioptimasi meliputi max_depth , $learning_rate$, $n_estimators$, $colsample_bytree$, dan $subsample$, dengan rentang nilai yang disamakan. Hasil eksperimen menunjukkan bahwa model XGBoost yang dioptimasi menggunakan Bayesian Optimization memberikan performa terbaik dibandingkan dengan model baseline, dengan nilai MSE sebesar 0.0001, RMSE sebesar 0.0105, MAE sebesar 0.0013, dan R^2 sebesar 0.9591. Dalam penelitian ini, R^2 dijadikan sebagai metrik utama untuk menilai apakah performa model mengalami peningkatan atau tidak. Evaluasi model menunjukkan bahwa model ini dapat memprediksi penjualan dengan akurasi yang tinggi, meskipun terdapat tantangan terkait fluktuasi musiman dan ketersediaan stok.

Hasil penelitian ini menunjukkan bahwa optimasi menggunakan Bayesian Optimization efektif dalam meningkatkan performa model prediksi penjualan, dan dapat diimplementasikan untuk mendukung perencanaan stok dan strategi penjualan yang lebih efisien di platform e-commerce.

Kata kunci : Prediksi Penjualan, Tiktok Shop, XGBoost, Bayesian Optimization, Machine Learning.

Halaman ini sengaja dikosongkan

ABSTRACT

Nama Mahasiswa / NPM	Firhan Ali Sofi / 21081010326
Judul Skripsi	Sales Prediction Optimization for Fashion Products on TikTokShop Using XGBoost and Bayesian Optimization
Dosen Pembimbing	1. Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU 2. Achmad Juna, S.Kom. M.Kom

Sales prediction is a crucial aspect of inventory management and business strategy planning. This study aims to develop a sales prediction model for fashion products using the XGBoost algorithm optimized with Bayesian Optimization. The data used in this study is sales transaction data from the fashion business D'Busana for the period of 2022 to 2025, with a focus on the e-commerce platform TikTok Shop. The model is developed to predict sales quantity based on factors such as product lifecycle, seasonal patterns, and stock availability.

The data processing phase includes preprocessing, time-series feature creation, as well as experimentation with feature engineering techniques to improve model performance. The data is split with an 80:10:10 ratio for training, validation, and testing. The best configuration is obtained by using Bayesian Optimization with a range of $n_estimators$ from 50 to 500 and 5-fold cross-validation. The hyperparameters optimized in this study include max_depth , $learning_rate$, $n_estimators$, $colsample_bytree$, and $subsample$, all of which were explored within identical value ranges. The experimental results show that the XGBoost model optimized using Bayesian Optimization delivers the best performance compared to the baseline model, achieving an MSE of 0.0001, RMSE of 0.0105, MAE of 0.0013, and R^2 of 0.9591. In this study, R^2 is used as the primary metric to assess whether the model's performance has improved. The model evaluation indicates that it can predict sales with high accuracy, although challenges remain in accounting for seasonal fluctuations and stock availability.

The results of this study indicate that optimization using Bayesian Optimization is effective in improving the sales prediction model's performance and can be implemented to support more efficient stock planning and sales strategies on e-commerce platforms.

Keyword: Sales Prediction, Tiktok Shop, XGBoost, Bayesian Optimization, Machine Learning.

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat, hidayah, dan karunia-Nya, sehingga skripsi yang berjudul “Optimasi Prediksi Penjualan Produk Fashion di TikTok Shop Menggunakan XGBoost dan Bayesian Optimization” ini akhirnya dapat diselesaikan dengan baik. Penyusunan skripsi ini merupakan salah satu bentuk tanggung jawab akademik penulis dalam menyelesaikan studi di Program Studi Informatika. Tentunya, dalam proses penyusunan skripsi ini, penulis tidak bekerja sendiri. Oleh karena itu, penulis ingin mengucapkan terima kasih sebesar-besarnya kepada:

1. **Bapak Dr. I Gede Susrama Mas Diyasa, S.T., M.T., IPU**, selaku Dosen Pembimbing Utama yang dengan penuh kesabaran telah membimbing, mengarahkan, dan memberikan masukan yang sangat berarti bagi penyempurnaan skripsi ini.
2. **Bapak Achmad Junaidi, S.Kom., M.Kom**, selaku Dosen Pembimbing Kedua atas bimbingan, arahan, serta dukungan yang sangat membantu selama proses penelitian berlangsung.
3. Seluruh Dosen Program Studi Informatika yang telah memberikan ilmu, wawasan, dan pengalaman berharga selama masa perkuliahan serta ujian-ujian tak terlupakan yang membentuk karakter dan kedisiplinan penulis.
4. Teman-teman seperjuangan, khususnya **Varrel, Iqbal, Faizal, dan Kesya**, yang senantiasa mendampingi penulis dalam berbagai suka dan duka, lengkap dengan informasi “ngopi” nya yang selalu membuat penulis bersemangat.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Akhir kata, semoga skripsi ini dapat memberikan manfaat, meskipun sederhana, dan menjadi pengingat bahwa segala sesuatu akan selesai jika dikerjakan dengan konsisten.

Surabaya, 16 Mei 2025

Firhan Ali Sofi
NPM 21081010326

Halaman ini sengaja dikosongkan

DAFTAR ISI

LEMBAR PENGESAHAN.....	ii
LEMBAR PERSETUJUAN.....	iv
SURAT PERNYATAAN BEBAS PLAGIAT	vi
ABSTRAK.....	viii
ABSTRACT	x
KATA PENGANTAR	xii
DAFTAR ISI	xiv
DAFTAR GAMBAR.....	xvi
DAFTAR TABEL	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	5
1.5 Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terdahulu	7
2.2 Landasan Teori.....	14
2.1.1 Machine Learning.....	14
2.1.2 Gradient Boosting	15
2.1.3 Feature Engineering	16
2.1.4 Extreme Gradient Boosting (XGBoost).....	19
2.1.5 Hyperparameter Tuning	23
2.1.6 Evaluation Metrik	25
2.1.7 Penggunaan XGBoost Dalam Bidang Lain	27
BAB III DESAIN DAN IMPLEMENTASI SISTEM.....	29
3.1 Tahapan Penelitian	29
3.2 Studi Pustaka	30
3.3 Pengumpulan Dataset	30
3.4 Preprocessing Data.....	36
3.4.1 Data Cleaning.....	37

3.4.2 Feature Engineering	38
3.4.3 Pembagian Dataset	39
3.4.4 Feature Selection	40
3.5 Model Development	41
3.6 Hyperparameter Tuning	45
3.7 Proses Training	48
3.8 Evaluasi Analisis Hasil	49
BAB IV PENGUJIAN DAN ANALISA.....	54
4.1 Persiapan Data	54
4.1.1 Library	55
4.1.2 Pemuatan Dataset	57
4.2 Preprocessing Data	59
4.2.1 Data Cleaning	59
4.2.2 Data Cleaning (Mean Fill)	60
4.2.3 Feature Engineering	61
4.2.2.1 Time Features	61
4.2.2.2 Time Series Features	64
4.2.4 Pembagian Data	66
4.2.5 Feature Selection	67
4.3 Training Model Baseline	72
4.4 Skenario Pengujian	75
4.4.1 Konfigurasi Parameter Tuning	75
4.4.2 Pengujian Bayesian Optimization	78
4.4.3 Pengujian Random Search	81
4.4.4 Pengujian Optuna	84
4.4.5 Konfigurasi Nilai Terbaik	87
4.4.6 Evaluasi Model Terbaik	88
4.5 Implementasi Model Prediksi	91
BAB V PENUTUP	96
5.1 Kesimpulan	96
5.2 Saran	97
DAFTAR PUSTAKA.....	99

DAFTAR GAMBAR

Gambar 2.1 Learning tree growth XGBoost	20
Gambar 3.1 Alur Penelitian	29
Gambar 3.2 Halaman Beranda Tiktok Shop.....	31
Gambar 3.3 Halaman Kelola Pesanan Tiktok Shop	31
Gambar 3.4 Filter Semua Pesanan.....	32
Gambar 3.5 Unduh Pesanan Tiktok Shop	32
Gambar 3.6 Dataset Mentah Excel.....	33
Gambar 3.7 Alur Preprocessing data	36
Gambar 3.8 Alur Feature Selection Bawaan XGBoost	40
Gambar 3.9 Alur Proses Pelatihan Model XGBoost	41
Gambar 3.10 Alur Optimasi Pencarian Hyperparameter	45
Gambar 3.11 Alur Proses Training XGBoost dengan Bayesian Optimization	48
Gambar 4.1 Daftar Fitur dalam Dataset	58
Gambar 4.2 Jumlah Missing Values.....	59
Gambar 4.3 Fitur Baru Time Features Engineering	63
Gambar 4.4 Fitur Baru Time-Series Feature Engineering.....	65
Gambar 4.5 Visualisasi Pembagian Data	67
Gambar 4.6 Fitur Terpilih berdasarkan Feature Importance	69
Gambar 4.7 Learning Curve MSE dari Baseline Model	74
Gambar 4.8 Grafik perbandingan nilai R^2 Skenario Pengujian.....	77
Gambar 4.9 Grafik perbandingan nilai R^2 BO dengan $n_estimators$ 50 – 500	79
Gambar 4.10 Grafik perbandingan nilai R^2 BO dengan $n_estimators$ 100 – 1000	80
Gambar 4.11 Grafik perbandingan nilai R^2 RS dengan $n_estimators$ 50 – 500	82
Gambar 4.12 Grafik perbandingan nilai R^2 RS dengan $n_estimators$ 100 – 1000	83
Gambar 4.13 Grafik perbandingan nilai R^2 Optuna dengan $n_estimators$ 50 – 500	85
Gambar 4.14 Grafik perbandingan nilai R^2 Optuna dengan $n_estimators$ 100 – 1000	86
Gambar 4.15 Evaluasi Metrik MSE Model Terbaik	89
Gambar 4.16 Evaluasi Metrik RMSE Model Terbaik.....	89
Gambar 4.17 Evaluasi Metrik MAE Model Terbaik.....	90
Gambar 4.18 Evaluasi Metrik R^2 Model Terbaik.....	90
Gambar 4.19 Grafik Perbandingan Nilai Aktual dan Prediksi	92
Gambar 4.20 Grafik Prediksi Bulan Depan.....	94

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 3.1 Kolom 1 – 6.....	33
Tabel 3.2 Kolom 7 – 12.....	34
Tabel 3.3 Kolom 13 – 18	35
Tabel 3.4 Kolom 22 - 26	35
Tabel 3.5 Data Cleaning.....	37
Tabel 3.6 Tabel Pembagian Data	39
Tabel 3.7 Parameter Pada XGBoost.....	42
Tabel 3.8 Contoh Dataset Implementasi XGBoost	43
Tabel 3.9 Perhitungan Gradien dan Hessian	43
Tabel 3.10 Hasil Prediksi Baru.....	44
Tabel 3.11 Contoh Data Implementasi Bayesian Optimization	46
Tabel 3.12 Hasil Expected Improvement.....	48
Tabel 3.13 Akurasi Prediksi Metrik Evaluasi	49
Tabel 3.14 Data Evaluasi Metrik.....	50
Tabel 3.15 Perhitungan Error pada Evaluasi Metrik.....	50
Tabel 3.16 Gambaran Skenario Pengujian	53
Tabel 4.1 Tabel hasil evaluasi metrik baseline model pada data testing.....	74
Tabel 4.2 Konfigurasi Parameter Tuning	75
Tabel 4.3 Hasil Keseluruhan Skenario Pengujian	76
Tabel 4.4 Hasil pengujian BO dengan n_estimators 50 – 500	78
Tabel 4.5 Hasil pengujian BO dengan n_estimators 100 – 1000	80
Tabel 4.6 Hasil pengujian RS dengan n_estimators 50 – 500.....	81
Tabel 4.7 Hasil pengujian RS dengan n_estimators 100 – 1000.....	83
Tabel 4.8 Hasil pengujian Optuna dengan n_estimators 50 – 500	84
Tabel 4.9 Hasil pengujian Optuna dengan n_estimators 100 – 1000	86
Tabel 4.10 Hasil terbaik dari masing-masing skema pengujian.....	87
Tabel 4.11 Parameter Terbaik	88

Halaman ini sengaja dikosongkan