

**IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK
PREDIKSI PENJUALAN PRODUK TOKO RETAIL**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

**THOMAS DITO RIGORASTIO
00000061313**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK
PREDIKSI PENJUALAN PRODUK TOKO RETAIL**



THOMAS DITO RIGORASTIO

00000061313

UMN

UNIVERSITAS

MULTIMEDIA

NUSANTARA

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA**

TANGERANG

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Thomas Dito Rigorastio
Nomor Induk Mahasiswa : 00000061313
Program Studi : Informatika

Skripsi dengan judul:

Implementasi Algoritma Random Forest Untuk Prediksi Penjualan Produk Toko Retail

merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan **TIDAK LULUS** untuk mata kuliah yang telah saya tempuh.

Tangerang, 4 Juli 2025



(Thomas Dito Rigorastio)

HALAMAN PENGESAHAN

Skripsi dengan judul

IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK PREDIKSI PENJUALAN PRODUK TOKO RETAIL

oleh

Nama : Thomas Dito Rigorastio
NIM : 00000061313
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Kamis, 17 Juli 2025

Pukul 13.00 s/s 15.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang



(Alexander Waworuntu, S.Kom., M.T.I.)

NIDN: 0309068503

Penguji



(Dr. Ir. Winarno, M.Kom.)

NIDN: 0330106002

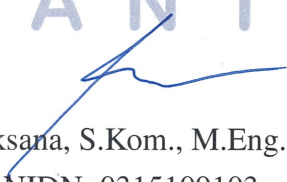
Pembimbing



(Dennis Gunawan, S.Kom., M.Sc.)

NIDN: 0320059001

Ketua Program Studi Informatika,



(Arya Wicaksana, S.Kom., M.Eng.Sc., OCA)

NIDN: 0315109103

**HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK
KEPENTINGAN AKADEMIS**

Yang bertanda tangan di bawah ini:

Nama : Thomas Dito Rigorastio
NIM : 00000061313
Program Studi : Informatika
Jenjang : S1
Judul Karya Ilmiah : Implementasi Algoritma Random
Forest Untuk Prediksi Penjualan
Produk Toko Retail

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
- Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - Embargo publikasi karya ilmiah dalam kurun waktu tiga tahun.

Tangerang, 4 Juli 2025

Yang menyatakan



Thomas Dito Rigorastio

**Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

HALAMAN PERSEMBAHAN / MOTTO

"A good name is to be more desired than great wealth, Favor is better than silver and gold."

Proverbs 22:1 (NASB)



KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas berkat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir yang berjudul "Implementasi Algoritma Random Forest untuk Prediksi Penjualan Produk Toko Retail". Laporan ini disusun sebagai salah satu syarat dalam menyelesaikan program studi dan memperoleh gelar Sarjana Pada Program Studi Informatika, Fakultas Teknik dan Informatika, Universitas Multimedia Nusantara. Penulis menyadari bahwa penyelesaian skripsi ini tidak akan mungkin tercapai tanpa adanya dukungan, bimbingan, dan bantuan dari berbagai pihak sejak masa perkuliahan hingga proses penyusunan skripsi.

Penulis Juga mengucapkan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, M.Sc., selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Bapak Arya Wicaksana, S.Kom., M.Eng.Sc., OCA, selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
4. Bapak Dennis Gunawan, S.Kom., M.Sc., sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Orang Tua dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tesis ini.
6. Pihak- pihak dan teman lain yang telah membantu dalam pengerjaan proses skripsi secara moral maupun saran.

Semoga hasil penelitian skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 4 Juli 2025



Thomas Dito Rigorastio

IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK PREDIKSI PENJUALAN PRODUK TOKO RETAIL

Thomas Dito Rigorastio

ABSTRAK

Prediksi penjualan yang akurat penting dalam industri *retail* untuk mendukung pengelolaan inventaris dan pengambilan keputusan bisnis. Penelitian ini bertujuan membangun model prediksi jumlah penjualan produk toko *retail* menggunakan algoritma *Random Forest*. Dataset yang digunakan berasal dari Kaggle dan mencakup data penjualan, harga, diskon, dan faktor eksternal lainnya. Proses preprocessing mencakup pembersihan data, *encoding* fitur kategorik, serta penghapusan *outlier* dengan metode IQR. Model dibangun menggunakan *Random Forest Regressor* dan dilakukan tuning *hyperparameter* menggunakan *RandomizedSearchCV*. Evaluasi dilakukan dengan metrik MAE, RMSE, dan MAPE. Hasil pengujian menunjukkan model memberikan performa yang cukup baik dengan nilai MAE sebesar 7.27, RMSE sebesar 8.49, dan MAPE sebesar 17,79% pada data uji. Temuan ini membuktikan bahwa algoritma *Random Forest* efektif digunakan dalam prediksi penjualan pada sistem berbasis data historis.

Kata kunci: *Machine Learning*, Prediksi Penjualan, *Random Forest*, *RandomizedSearchCV*, Regresi, Retail.



IMPLEMENTATION OF RANDOM FOREST ALGORITHM IN RETAIL STORE PRODUCT SALES PREDICTION SYSTEM

Thomas Dito Rigorastio

ABSTRACT

Accurate sales prediction is essential in the retail industry to support inventory management and business decision-making. This study aims to build a model for predicting the number of product sales in retail stores using the Random Forest algorithm. The dataset used was obtained from Kaggle and includes sales data, pricing, discounts, and other external factors. The preprocessing process involved data cleaning, categorical feature encoding, and outlier removal using the IQR method. The model was built using Random Forest Regressor and tuned with RandomizedSearchCV. Evaluation was conducted using MAE, RMSE, and MAPE metrics. The testing results showed that the model performed reasonably well, achieving a MAE of 7.27, RMSE of 8.49, and MAPE of 17.79% on the test data. These findings demonstrate that the Random Forest algorithm is effective for sales prediction in systems based on historical data.

Keywords: Machine Learning, Sales Forecasting, Random Forest, RandomizedSearchCV, Regression, Retail



DAFTAR ISI

HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
HALAMAN PERSEMBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR KODE	xiii
DAFTAR RUMUS	xiv
DAFTAR LAMPIRAN	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	4
1.3 Batasan Permasalahan	4
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
1.6 Sistematika Penulisan	5
BAB 2 LANDASAN TEORI	7
2.1 Penjualan Produk Toko Retail	7
2.2 Machine Learning	7
2.3 Supervised Learning	8
2.4 Interquartile Range (IQR)	8
2.5 Transformasi Logaritmik	10
2.6 Deployment dengan Streamlit	12
2.7 Random Forest	13
2.7.1 Hyperparameter	16
2.7.2 Keunggulan dan kekurangan	17
2.8 Metriks Evaluasi	17
2.8.1 Root Mean Squared Error (RMSE)	18
2.8.2 Mean Absolute Error (MAE)	18
2.8.3 Mean Absolute Percentage Error (MAPE)	19
BAB 3 METODOLOGI PENELITIAN	20
3.1 Studi Literatur	21
3.2 Pengumpulan Data	22
3.3 Pembersihan Data	23
3.3.1 Pembersihan Missing Value	23
3.3.2 Pembersihan Duplikasi Data	23
3.3.3 Pembersihan Outlier	24
3.4 Preprocessing	25
3.4.1 Pembuatan Fitur Baru	26
3.4.2 Encoding	26
3.4.3 Pemisahan Data	27
3.5 Pembangunan Model	27
3.6 Deployment	29

3.7	Pengujian Model	31
3.8	Dokumentasi	32
BAB 4	HASIL DAN DISKUSI	33
4.1	Hasil Implementasi Pembelajaran Mesin	33
4.1.1	Data Gathering	33
4.1.2	Assessing & Cleaning Data	34
4.1.3	Preprocessing	41
4.1.4	Pembangunan Model	47
4.2	Implementasi Web-App dan Deployment	52
4.2.1	Penyimpanan Model	52
4.2.2	Pembuatan Aplikasi Web dengan Streamlit	53
4.2.3	Deployment ke Streamlit Cloud	54
4.3	Pengujian Model	55
4.3.1	Skenario Pengujian	56
4.3.2	Hasil Uji Coba	57
4.4	Diskusi	67
BAB 5	SIMPULAN DAN SARAN	70
5.1	Simpulan	70
5.2	Saran	70
DAFTAR PUSTAKA		72



DAFTAR TABEL

Tabel 4.1	Hasil Evaluasi Model Random Forest (Baseline)	49
Tabel 4.2	Hasil Evaluasi Dengan Best Params Hasil Tuning	52
Tabel 4.3	Hyperparameter terbaik berdasarkan hasil <i>best_params</i> dari <i>RandomizedSearchCV</i>	56
Tabel 4.4	Tabel Skenario Uji Coba Hyperparameter Random Forest .	56
Tabel 4.5	Hasil Evaluasi pada Variasi Hyperparameter <i>n_estimators</i> .	58
Tabel 4.6	Hasil Evaluasi pada Variasi Hyperparameter <i>max_depth</i> . .	60
Tabel 4.7	Hasil Evaluasi Variasi <i>min_samples_leaf</i> pada Model Random Forest	63
Tabel 4.8	Hasil Evaluasi Model terhadap Variasi <i>min_samples_split</i> .	65



DAFTAR GAMBAR

Gambar 2.1	Alur Kerja Random Forest	15
Gambar 3.1	Flow Metodologi Penelitian	20
Gambar 3.2	Gambar Alur Penelitian	21
Gambar 3.3	Retail Store Dataset Kaggle	22
Gambar 3.4	Flowchart Pembersihan Data	23
Gambar 3.5	Flowchart IQR	25
Gambar 3.6	Flowchart Preprocessing	26
Gambar 3.7	Flowchart Pembangunan Model	27
Gambar 3.8	Flowchart Random Forest	29
Gambar 3.9	Diagram alur proses deployment model ke aplikasi web	31
Gambar 3.10	Flowchart Pengujian Model	32
Gambar 4.1	Import Dataset	33
Gambar 4.2	Dataframe	33
Gambar 4.3	Informasi Data	34
Gambar 4.4	Informasi Data	35
Gambar 4.5	Info Missing Value	36
Gambar 4.6	Duplikasi Data dan Ringkasan Statistik	36
Gambar 4.7	Distribusi Nilai 0 units Sold pada Category	37
Gambar 4.8	Distribusi Nilai 0 Demand pada Category	38
Gambar 4.9	Hasil Filter nilai Units Sold dan Demand yang di bawah nol	38
Gambar 4.10	Box Plot Check	40
Gambar 4.11	Hasil Filter nilai Units Sold dan Demand yang di bawah nol	41
Gambar 4.12	Box Plot Cross Check	41
Gambar 4.13	Pembuatan Fitur Baru	43
Gambar 4.14	Label Encoding Seasonality	44
Gambar 4.15	Encoding Fitur Category	45
Gambar 4.16	Pembuatan Fitur Baru	46
Gambar 4.17	Transformasi Logaritmik Pada Target	48
Gambar 4.18	Potongan kode baseline model	48
Gambar 4.19	Potongan kode baseline model	48
Gambar 4.20	Kombinasi hyperparameter	50
Gambar 4.21	Potongan kode tuning dengan randomized search	51
Gambar 4.22	Tampilan awal aplikasi prediksi penjualan produk retail	54
Gambar 4.23	Tampilan bagian bawah aplikasi dan hasil prediksi	55
Gambar 4.24	Pengaruh n_estimators terhadap MAPE	59
Gambar 4.25	Pengaruh n_estimators terhadap MAPE pada set data train dan test	59
Gambar 4.26	Pengaruh max_depth terhadap MAPE	61
Gambar 4.27	Pengaruh max_depth terhadap MAPE pada set data train dan test	62
Gambar 4.28	Pengaruh min_samples_leaf terhadap MAPE	63
Gambar 4.29	Pengaruh min_samples_leaf terhadap MAPE pada set data train dan test	64
Gambar 4.30	Pengaruh min_samples_split terhadap MAPE	66
Gambar 4.31	Pengaruh min_samples_split terhadap MAPE pada set data train dan test	66
Gambar 4.32	Learning Curve	68

DAFTAR KODE

Kode 4.1	Melihat Struktur dan Informasi dataset	34
Kode 4.2	Mengubah Tipe Data	35
Kode 4.3	Pengecekan Missing Value	35
Kode 4.4	Duplikasi Data dan Ringkasan Statistik	36
Kode 4.5	Analisis Distribusi Nilai ≤ 0	37
Kode 4.6	Conditional Filtering	38
Kode 4.7	Boxplot Cek Untuk Outlier	39
Kode 4.8	Metode IQR	40
Kode 4.9	Pembuatan Fitur baru	42
Kode 4.10	Pemilihan Fitur X dan Y	46
Kode 4.11	Penyimpanan Joblib	53



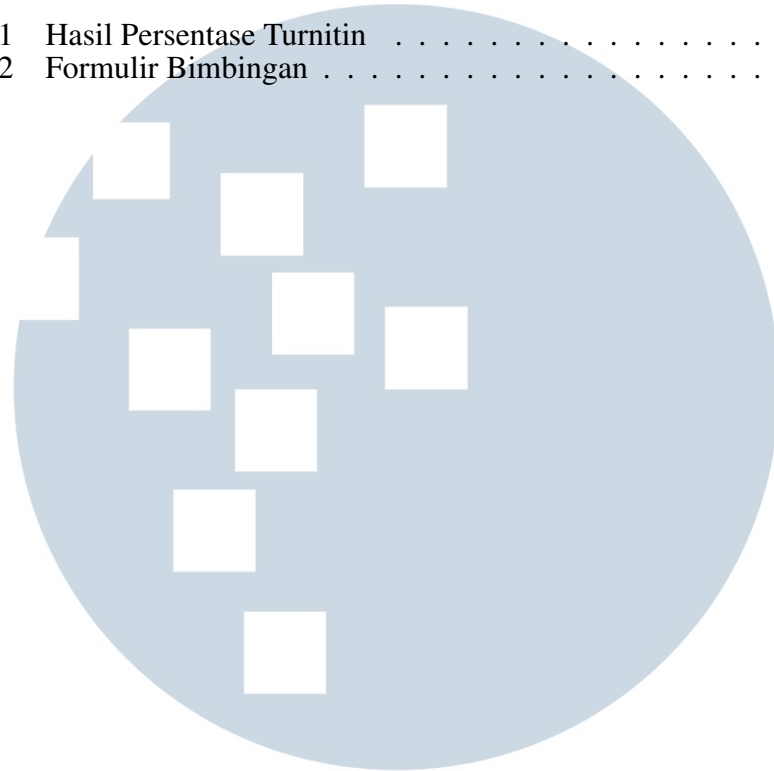
DAFTAR RUMUS

Rumus 2.1	<i>IQR</i>	8
Rumus 2.2	<i>Lower Bound</i>	9
Rumus 2.3	<i>Upper Bound</i>	9
Rumus 2.4	<i>Lower Bound</i>	10
Rumus 2.5	<i>Upper Bound</i>	10
Rumus 2.6	<i>Transformasi Logaritmik</i>	11
Rumus 2.7	<i>Transformasi Eksponensial</i>	11
Rumus 2.8	<i>RMSE</i>	18
Rumus 2.9	<i>MAE</i>	18
Rumus 2.10	<i>MAPE</i>	19



DAFTAR LAMPIRAN

Lampiran 1	Hasil Persentase Turnitin	75
Lampiran 2	Formulir Bimbingan	85



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA