



**Analisis Segmentasi dan Prediksi Penjualan Barang pada Toko Retail
Menggunakan Metode K-Means dan Random Forest**

**TUGAS AKHIR
SKRIPSI**

AZZAHRA ANANDA PUTRI

41822010057

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2026



**Analisis Segmentasi dan Prediksi Penjualan Barang pada Toko Retail
Menggunakan Metode K-Means dan Random Forest**

TUGAS AKHIR

SKRIPSI

AZZAHRA ANANDA PUTRI

41822010057

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2026

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Azzahra Ananda Putri

NIM : 41822010057

Program Studi : Sistem Informasi

Menyatakan dengan sebenar-benarnya bahwa Tugas Akhir berjudul:

"Analisis Segmentasi dan Prediksi Penjualan Barang pada Toko Retail

Menggunakan Metode K-Means dan Random Forest" adalah hasil karya saya sendiri, tidak mengandung unsur plagiarisme, pelanggaran hak cipta, atau konten ilegal dalam bentuk apapun dan tidak melanggar hukum atau hak pihak manapun.

Apabila di kemudian hari ditemukan pelanggaran terhadap pernyataan ini, saya bersedia menanggung seluruh konsekuensi hukum dan membebaskan Universitas Mercu Buana dari segala bentuk tuntutan hukum dan saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana. Demikian pernyataan ini saya buat dengan sebenar-benarnya untuk digunakan sebagaimana mestinya.

Jakarta, 30 Desember 2025


Azzahra Ananda Putri

HALAMAN SURAT KETERANGAN UJI TURNITIN



072.423.4.03.01

PERNYATAAN SIMILARITY CHECK */SIMILARITY CHECK STATEMENT*

Saya yang bertanda tangan dibawah ini menyatakan, bahwa karya ilmiah yang ditulis oleh

/The undersigned, hereby declare that the scientific work written by

Nama /Name : Azzahra Ananda Putri
NIM /Student ID Number : 41822010057
Program Studi /Study Program : Sistem Informasi

Dengan Judul Tugas Akhir

/The title:

“Analisis Segmentasi dan Prediksi Penjualan Barang pada Toko Retail Menggunakan Metode K-Means dan Random Forest”

telah dilakukan pengecekan *similarity* dengan sistem Turnitin pada tanggal:

/Similarity checks have been carried out with the Turnitin system on the date:

30 Desember 2025

dengan nilai persentase sebesar :

/with a percentage value of:

8%

dinyatakan memenuhi standar sesuai dengan ketentuan berlaku di **Fakultas Ilmu Komputer** Universitas Mercu Buana. */declared to meet standards in accordance with applicable regulations at the Faculty of Computer Science, Universitas Mercu Buana.*

File hasil cek similarity turnitin:

/Turnitin similarity report file

https://drive.google.com/file/d/1EIC4fZSpXa3s_nXWIKRfhgeY3Jby-SU/view?usp=drive_link



Jakarta, 30 Desember 2025
Admin Turnitin Fasilkom UMB

Agung Prawoto, S.Kom., B.Sc
NIK : 322970503

Fakultas Ilmu Komputer

KAMPUS MENARA BHAKTI

Jl. Raya Meruya Selatan No. 1 Kembangan, Jakarta Barat 11650

Telp. 021-5840816 (Hunting), Psw : 5700 Fax. 021-5840813

<http://www.mercubuana.ac.id>, e-mail : fasilkom@mercubuana.ac.id

HALAMAN PENGESAHAN

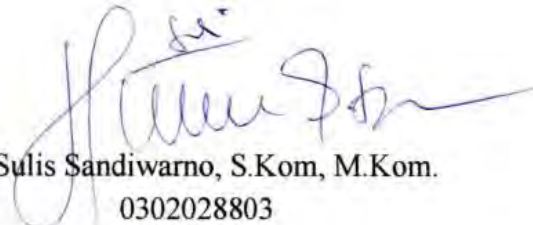
Laporan Skripsi ini diajukan oleh:

Nama : Azzahra Ananda Putri
NIM : 41822010057
Program Studi : Sistem Informasi
Judul Laporan Skripsi : Analisis Segmentasi dan Prediksi Penjualan Barang
pada Toko Retail Menggunakan Metode K-Means
dan Random Forest

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing,



Sulis Sandiwarno, S.Kom, M.Kom.
0302028803

Jakarta, 25 Januari 2026

Mengetahui,

Dekan Fakultas Ilmu Komputer



Dr. Bambang Jokonowo, S.Si., M.T.I.
NIDN: 0320037002

Ketua Program Studi
Sistem Informasi



Wawan Gunawan, S.Kom., M.T., M.Kom.
NIDN: 0424108104

KATA PENGANTAR

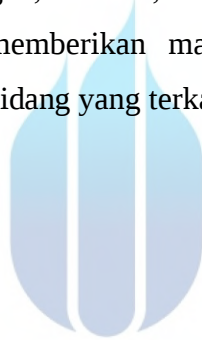
Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan tugas akhir ini dengan baik.

Penulis menyadari bahwa penyusunan tugas akhir ini tidak akan berjalan lancar tanpa bantuan, dukungan, dan arahan dari berbagai pihak. Untuk itu, penulis mengucapkan terima kasih dan penghargaan yang sebesar-besarnya kepada:

1. Allah SWT yang senantiasa memberikan kesehatan, kemudahan, dan kelancaran kepada penulis selama proses penyusunan tugas akhir ini;
2. Orangtua dan seluruh keluarga yang telah mendukung serta mempercayai proses perjuangan penulis hingga skripsi ini terselesaikan;
3. Prof. Dr. Ir. Andi Adriansyah, M. Eng selaku Rektor Universitas Mercu Buana Jakarta
4. Prof. Dr. Bambang Jokonowo, S.Si., M.T.I selaku Dekan Fakultas Ilmu Komputer (Fasilkom) Universitas Mercu Buana;
5. Wawan Gunawan, S. Kom, MT., M. Kom. selaku Ketua Program Studi Sistem Informasi;
6. Bapak Sulis Sandiwarno, S. Kom., M. Kom. selaku dosen pembimbing yang telah memberikan bimbingan dengan penuh kesabaran serta saran-saran berharga selama proses pengerjaan skripsi ini;
7. Ibu/Bapak Dosen dan staf atas segala dukungan akademik dan fasilitas yang diberikan selama menempuh studi di program ini;
8. Universitas Mercu Buana atas dukungan dan kesempatan melalui pemberian beasiswa penuh, sehingga penulis dapat menempuh pendidikan hingga tahap penyelesaian tugas akhir ini;
9. Zaidan Adzka Muhammad selaku partner penulis sejak menempuh pendidikan di bangku SMA yang setia kebersamai setiap proses, memberi semangat, dan dukungan hingga tugas akhir ini terselesaikan;
10. Hizkia, Fira, Nadhira, Thania, Bulan dan Sheela yang telah memberikan semangat, saling membantu dan kebersamaan selama masa perkuliahan dan pengerjaan tugas akhir;
11. Teman-teman Marketing Menteng yang telah kebersamai penulis selama proses magang;

12. Teman-teman Sistem Informasi 2022 Menteng–Meruya–Warbun, Sistem Informasi 2021 dan 2023 Meruya, serta HIMSISFO UMB periode 2024/2025 dan periode 2023/2024, atas pengalaman, kebersamaan, dan pembelajaran yang sangat berharga, yang telah membentuk penulis hingga menjadi seperti saat ini;
13. Teknologi kecerdasan buatan, khususnya ChatGPT dan Claude, yang telah membantu penulis dalam mempermudah proses pemahaman, eksplorasi ide, serta penyusunan materi selama pengerjaan tugas akhir;
14. Terakhir penulis mengucapkan terimakasih kepada diri sendiri karena telah mampu menyelesaikan perkuliahan dengan baik dan tepat waktu.

Sebagai penutup, penulis mengucapkan syukur kepada Tuhan Yang Maha Esa atas terselesaikannya Tugas Akhir ini. Penulis juga menyampaikan terima kasih kepada seluruh pihak yang telah memberikan dukungan, bantuan, dan kontribusi selama proses penyusunan. Semoga Tugas Akhir ini dapat memberikan manfaat serta menjadi referensi dalam pengembangan ilmu pengetahuan di bidang yang terkait.



Jakarta, 28 Desember 2025

UNIVERSITAS
MERCU BUANA

Penulis

HALAMAN PENYATAAN PERSTUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama Mahasiswa : Azzahra Ananda Putri

NIM 41822010057

Fakultas/ Program studi : Ilmu Komputer / Sistem Informasi

Judul Tugas Akhir : Analisis Segmentasi dan Prediksi Penjualan
Barang pada Toko Retail Menggunakan Metode
K-Means dan Random Forest

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 25 Desember 2025

Yang menvatakan.



Azzahra Ananda Putri

ANALISIS SEGMENTASI DAN PREDIKSI PENJUALAN BARANG PADA TOKO RETAIL MENGGUNAKAN METODE K-MEANS DAN RANDOM FOREST

AZZAHRA ANANDA PUTRI

ABSTRAK

Perkembangan analisis data telah mendukung berbagai aktivitas bisnis, termasuk di sektor ritel dan distribusi. Pada penelitian sebelumnya sudah dilakukan evaluasi terhadap sektor ritel, dimana, faktor-faktor tersebut dapat memengaruhi harga jual dan volume penjualan. Penelitian tersebut mengadopsi beberapa pendekatan machine learning seperti *DBSCAN* dan *Linier Regression*. Akan tetapi metode tersebut memiliki sejumlah kekurangan, seperti mudah terpengaruh outlier, tidak stabil pada pola penjualan yang tidak konsisten, serta kurang efektif dalam memodelkan karakteristik penjualan yang berubah secara berkala. Kondisi ini membuat hasil segmentasi maupun prediksi penjualan tidak selalu konsisten. Oleh karena itu, dalam penelitian ini kami mengusulkan sebuah pendekatan baru dengan yang dinamakan *Numerical K-Means Retail Analytics and Forecasting (N-KRAF)*, dimana model tersebut menggabungkan K-means dan Random Forest. *K-Means* digunakan untuk mengelompokkan produk berdasarkan jejak penjualan historis seperti frekuensi transaksi dan jumlah barang yang terjual. Sementara itu, *Random Forest* dimanfaatkan sebagai salah satu variabel dalam model untuk memprediksi jumlah penjualan, bersama dengan harga jual dan kategori produk. Dari hasil pengujian, model *N-KRAF* memberikan segmentasi produk sekaligus prediksi penjualan yang lebih akurat dibandingkan pendekatan sebelumnya. Hal ini menunjukkan bahwa penggabungan kedua metode tersebut dapat membantu toko ritel dalam memahami karakteristik produknya dan membuat keputusan penjualan yang lebih terarah.

Kata Kunci : *K-Means*, *Random Forest*, *N-KRAF*, Segmentasi Produk, Prediksi Penjualan, Analisis Ritel, *Data Mining*.

SEGMENTATION ANALYSIS AND SALES PREDICTION OF PRODUCTS IN A RETAIL STORE USING K-MEANS AND RANDOM FOREST METHODS

AZZAHRA ANANDA PUTRI

ABSTRACT

The development of data analysis has supported various business activities, including those in the retail and distribution sectors. Previous studies have evaluated the retail sector, identifying factors that can influence selling prices and sales volumes. These studies adopted several machine learning approaches, such as DBSCAN and Linear Regression. However, these methods have several limitations, including high sensitivity to outliers, instability when dealing with inconsistent sales patterns, and limited effectiveness in modeling sales characteristics that change periodically. These conditions cause the results of segmentation and sales prediction to be inconsistent. Therefore, this study proposes a new approach called Numerical K-Means Retail Analytics and Forecasting (N-KRAF), which combines K-Means and Random Forest methods. K-Means is used to cluster products based on historical sales traces, such as transaction frequency and the number of items sold. Meanwhile, Random Forest is utilized as one of the variables in the model to predict sales quantity, along with selling price and product category. Based on the experimental results, the N-KRAF model provides both product segmentation and sales prediction with higher accuracy compared to previous approaches. This indicates that the integration of these two methods can assist retail stores in understanding product characteristics and making more targeted sales decisions.

Keywords: K-Means, Random Forest, N-KRAF, Product Segmentation, Sales Prediction, Retail Analysis, Data Mining.

DAFTAR ISI

HALAMAN PERNYATAAN KARYA SENDIRI	iii
HALAMAN SURAT KETERANGAN UJI TURNITIN	iv
HALAMAN PENGESAHAN.....	v
KATA PENGANTAR.....	vi
HALAMAN PENYATAAN PERSTUJUAN PUBLIKASI TUGAS AKHIR DI REPOSITORI UMB	viii
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI.....	xi
DAFTAR TABEL	xiv
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian.....	3
1.5 Batasan Penelitian.....	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	5
2.1 Segmentasi dan Prediksi Penjualan	5
2.2 Analisis Machine Learning dalam Segmentasi Penjualan	6
2.2.1 K-Means.....	7
2.2.2 DBSCAN	8
2.3 Analisis Machine Learning dalam Prediksi Penjualan	10
2.3.1 Regresi Linear	10
2.3.2 Random Forest	11
2.6 Penelitian Terdahulu.....	12

2.7 Analisis Literature Review	25
BAB III METODE PENELITIAN	27
3.1 Jenis Penelitian.....	27
3.2 Deskripsi Sumber Data.....	27
3.3 Tahapan Penelitian.....	28
3.3.1 Dataset.....	28
3.3.2 Data Preprocessing	28
3.3.2.1 Data Cleansing	29
3.3.2.2 Data Transformation	29
3.3.3 N-KRAF Model	29
3.3.3.1 K-Means + Dim. Reduction PCA	30
3.3.3.2 Random Forest.....	33
3.3.3.3 Concatenation	34
3.3.4 Output	36
BAB IV PEMBAHASAN	38
4.1 Experiment Set Up.....	38
4.2 Dataset.....	38
4.3 Parameter Set Up.....	39
4.3.1 Parameter Feature Engineering	39
4.3.2 Parameter StandardScaler.....	39
4.3.3 Parameter PCA	39
4.3.4 Parameter K-Means.....	40
4.3.5 Parameter Random Forest + RFE.....	40
4.4 Hasil.....	41
4.4.1 Hasil K-Means.....	41
4.4.1.1 Karakteristik Cluster	42
4.4.1.2 Evaluasi Model Clustering	44
4.4.2 Hasil Random Forest + RFE	47
4.4.2.1 Perbandingan Model (Linear Regression vs Random Forest).....	47
4.4.2.2 Feature Engineering Time Series.....	49
4.4.2.3 Recursive Feature Elimination (RFE)	50
4.4.2.4 Feature Importance.....	50
4.4.3 <i>Concatenation</i> dan Hasil <i>Forecasting</i>	52

4.5 Diskusi.....	54
BAB V PENUTUP.....	58
5.1 Kesimpulan.....	58
5.2 Saran	58
DAFTAR PUSTAKA	60
LAMPIRAN.....	64



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	12
Tabel 4. 1 Fitur PCA	51
Tabel 4. 2 Hasil Forecasting	52
Tabel 4. 3 Analisis Year-over-Year	52
Tabel 4. 4 Diskusi	56



DAFTAR GAMBAR

Gambar 2. 1 Segmentasi.....	6
Gambar 3. 1 Alur Penelitian	28
Gambar 3. 2 Alur Kerja K-Means.....	32
Gambar 3. 3 Concatenation	34
Gambar 4. 1 Hasil Clustering	42
Gambar 4. 2 Perbandingan Revenue per Cluster	44
Gambar 4. 3 Lee Mineral 300 ML.....	53
Gambar 4. 4 Lee Mineral 600 ML.....	53
Gambar 4. 5 Avion 240 ML.....	54
Gambar 4. 6 Graph Perbandingan.....	54



DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi.....	64
Lampiran 2. Curriculum Vitae.....	66
Lampiran 3. BNSP	67
Lampiran 4. Turnitin.....	68

