

**PENERAPAN *DATA MINING* UNTUK PREDIKSI  
BARANG ELEKTRONIK PALING DIMINATI  
MENGUNAKAN METODE NAÏVE BAYES DAN C4.5**

**SKRIPSI SARJANA SISTEM INFORMASI**

**Disusun Oleh :**

Threva Veronika Boru Manalu

197006416177



**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNOLOGI KOMUNIKASI &  
INFORMATIKA  
UNIVERSITAS NASIONAL**

**2025**

**PENERAPAN DATA MINING UNTUK PREDIKSI  
BARANG ELEKTRONIK PALING DIMINATI  
MENGUNAKAN METODE NAÏVE BAYES DAN C4.5  
SKRIPSI SARJANA SISTEM INFORMASI**

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer  
dari Fakultas Teknologi Komunikasi dan Informatika.

**Disusun Oleh :**

Threva Veronika Boru Manalu

197006416177



**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS TEKNOLOGI KOMUNIKASI &  
INFORMATIKA**

**UNIVERSITAS NASIONAL**

**2025**

HALAMAN PENGESAHAN

TUGAS AKHIR

**PENERAPAN DATA MINING UNTUK PREDIKSI BARANG  
ELEKTRONIK PALING DIMINATI MENGGUNAKAN METODE  
NAÏVE BAYES DAN C4.5**



Threva Veronika Boru Manalu  
197006416177

Dosen Pembimbing

A handwritten signature in black ink, likely belonging to Dr. Agung Triayudi, is placed below the text 'Dosen Pembimbing'.

(Dr. Agung Triayudi, S.Kom, M.Kom)

## PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

### **PENERAPAN DATA MINING UNTUK PREDIKSI BARANG ELEKTRONIK PALING DIMINATI MENGGUNAKAN METODE NAÏVE BAYES DAN C4.5.**

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 01 September 2025



Threva Veronika Boru Manalu

197006416177

## LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

### **Penerapan Data Mining untuk Prediksi Barang Elektronik Paling Diminati Menggunakan Metode Naïve Bayes dan C4.5**

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diajukan pada Sidang Review Akhir Semester Genap 2024-2025 pada tanggal 27 Agustus Tahun 2025.

Dosen Pembimbing I

  
Dr. Agung Triayudi, S.Kom.,

M.Kom

NIDN : 0419068604

Ketua Program Studi

  
Dr. Andri Pratiwi, S.Kom.,

MMSI

NIDN : 0303097902

**LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI**

Nama : Threva Veronika Boru Manalu

NPM : 197006416177

Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika

Program Studi : Sistem Informasi

Tanggal Sidang : Rabu, 27 Agustus 2025

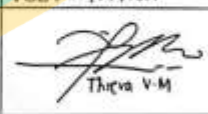
JUDUL DALAM BAHASA INDONESIA :

PENERAPAN DATA MINING UNTUK PREDIKSI BARANG ELEKTRONIK  
PALING DIMINATI MENGGUNAKAN METODE NAIVE BAYES DAN C4.5.

JUDUL DALAM BAHASA INGGRIS :

IMPLEMENTATION OF DATA MINING FOR PREDICTING THE MOST  
POPULAR ELECTRONIC PRODUCTS USING THE NAÏVE BAYES METHOD  
AND C4.5.

**TANDA TANGAN DAN TANGGAL**

| Pembimbing                                                                          | Prodi                                                                               | Mahasiswa                                                                            |
|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| TGL : 27/08/2025                                                                    | TGL : 27/08/2025                                                                    | TGL : 05/08/2025                                                                     |
|  |  |  |

## KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Maha Esa yang telah memberikan berkat dan penyertaan-Nya sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul **“Penerapan Data Mining Untuk Prediksi Barang Elektronik Paling Diminati Menggunakan Metode Naïve Bayes dan C4.5”** sebagai salah satu syarat untuk menyelesaikan studi jenjang strata satu (S-1) Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional.

Penelitian dan penulisan skripsi ini tidak terlepas dari bantuan berbagai pihak, oleh karena itu penulis menyampaikan banyak terima kasih kepada dosen pembimbing Tugas Akhir, Dr. Agung Triayudi, S.Kom., M.Kom., yang telah meluangkan banyak waktu, kesabaran, bimbingan, arahan serta memaklumi segala kekurangan penulis selama penelitian tugas akhir dan penyusunan skripsi. Penulis juga mengucapkan banyak terima kasih kepada :

1. Orang tua penulis yang telah memberikan doa dan dukungan.
2. PT Air Mas Perkasa yang telah memberikan bantuan selama penelitian dalam bentuk data.
3. Seluruh dosen pengajar di Program Studi Sistem Informasi FTKI.
4. Terima kasih kepada semua pihak yang tidak bisa disebutkan satu persatu yang telah memberikan banyak doa dan bantuan dalam proses penyusunan skripsi ini.

Semoga Tuhan Yang Maha Esa membalas segala bantuan dan kebaikan yang telah diberikan dengan hal yang lebih baik. Penulis mengharapkan kritik dan saran yang bersifat membangun dan semoga skripsi ini juga dapat memberikan manfaat dalam bidang Sistem Informasi.

Jakarta, 27 Agustus 2025

Threva Veronika Boru Manalu

## ABSTRAK

Perkembangan teknologi informasi mendorong meningkatnya kebutuhan masyarakat terhadap produk elektronik, terutama produk elektronik seperti smartphone, laptop maupun peralatan rumah tangga. PT Air Mas Perkasa sebagai salah satu perusahaan distribusi elektronik terkemuka di Indonesia memiliki data penjualan yang cukup besar yang dapat dimanfaatkan untuk menggali informasi produk elektronik yang paling diminati oleh pelanggan. Penelitian ini bertujuan untuk melakukan prediksi produk elektronik yang paling diminati dengan memanfaatkan penerapan data mining menggunakan metode Naïve Bayes dan Algoritma C4.5. Data yang digunakan merupakan data penjualan elektronik tahun 2023 dan 2024 dengan atribut tahun, deskripsi barang, kategori, *brand*, harga, jumlah serta status penjualan sebagai label. Tahapan penelitian menggunakan proses *Knowledge Discovery in Database* (KDD) yang meliputi seleksi data, pra pengolahan data, transformasi data hingga pembentukan model klasifikasi. Hasil pengujian dilakukan menggunakan *tools* RapidMiner menunjukkan bahwa barang elektronik paling diminati adalah ACER Notebook Travelmate P214 (i5-1135G7/8GB/512GB/Win11/UMA/14"HD/Black)[TRAVELMATEP214(TMP214/0034)] yang dijual dalam jumlah sedang maupun banyak serta memiliki harga menengah. Nilai akurasi Algoritma C4.5 lebih baik dibandingkan dengan mode Naïve Bayes. Penelitian ini diharapkan dapat memberikan kontribusi positif untuk membantu perusahaan dalam merumuskan strategi pengelolaan stock dan perencanaan pemasaran yang lebih baik.

**Kata Kunci :** *Data Mining*, Naïve Bayes, Algoritma C4.5, Klasifikasi, Barang Elektronik, RapidMiner, Prediksi Penjualan.

## ABSTRACT

The development of information technology has encouraged the increasing need of people for electronic products, especially electronic products such as smartphones, laptops and household appliances. PT Air Mas Perkasa as one of the leading electronic distribution companies in Indonesia has a large amount of sales data that can be used to explore information on electronic products that are most in demand by customers. This researcher aims to predict the most popular electronic products by utilizing the application of data mining using the Naïve Bayes method and the C4.5 Algorithm. The data used is electronic sales data for 2023 and 2024 with year attributes, product descriptions, categories, brands, prices, quantities, and sales status as labels. The research stage uses the Knowledge Discovery in Database (KDD) process which includes data selection, data pre-processing, data transformation to the formation of a classification model. The test results were carried out using the RapidMiner tool showing that the most popular electronic item is the ACER Notebook Travelmate P214 (i5-1135G7/8GB/512GB/Win11/UMA/14"HD/Black)[TRAVELMATEP214(TMP214/0034)] which is sold in medium and large quantities and has a medium price. The accuracy value of the C4.5 Algorithm is better compared to the Naïve Bayes mode. This research is expected to make a positive contribution to assist companies in formulating better stock management strategies and marketing planning.

**Keywords:** Data Mining, Naïve Bayes, C4.5 Algorithm, Classification, Electronics, RapidMiner, Sales Prediction.

## DAFTAR ISI

|                                                |             |
|------------------------------------------------|-------------|
| <b>KATA PENGANTAR.....</b>                     | <b>vii</b>  |
| <b>ABSTRAK.....</b>                            | <b>viii</b> |
| <b>ABSTRACT .....</b>                          | <b>ix</b>   |
| <b>DAFTAR ISI.....</b>                         | <b>x</b>    |
| <b>DAFTAR GAMBAR .....</b>                     | <b>xii</b>  |
| <b>DAFTAR TABEL .....</b>                      | <b>xiii</b> |
| <b>BAB I PENDAHULUAN.....</b>                  | <b>1</b>    |
| 1.1    Latar Belakang .....                    | 1           |
| 1.2    Identifikasi Masalah .....              | 2           |
| 1.3    Tujuan Penelitian.....                  | 2           |
| 1.4    Batasan Masalah.....                    | 2           |
| 1.5    Manfaat.....                            | 3           |
| <b>BAB II LANDASAN TEORI .....</b>             | <b>4</b>    |
| 2.1    Studi Literatur .....                   | 4           |
| 2.2    Landasan Teori .....                    | 19          |
| 2.2.1    Data Mining.....                      | 19          |
| 2.2.2    Klasifikasi Data Mining .....         | 19          |
| 2.2.3    Knowledge Discovery in Database ..... | 19          |
| 2.2.4    Naïve Bayes.....                      | 21          |
| 2.2.5    Agoritma C4.5 .....                   | 22          |
| 2.2.6    RapidMiner.....                       | 24          |
| 2.3    Penelitian Terdahulu .....              | 24          |
| <b>BAB III METODOLOGI PENELITIAN .....</b>     | <b>27</b>   |
| 3.1    Metodologi Penelitian .....             | 27          |
| 3.2    Metode Pengumpulan Data .....           | 27          |
| 3.3    Desain Penelitian .....                 | 28          |
| 3.4    Pengolahan Data Mining .....            | 29          |

|                                         |                                                               |           |
|-----------------------------------------|---------------------------------------------------------------|-----------|
| 3.4.1                                   | <i>Selection</i> .....                                        | 29        |
| 3.4.2                                   | <i>Preprocessing</i> .....                                    | 29        |
| 3.4.3                                   | <i>Transformation</i> .....                                   | 29        |
| 3.4.4                                   | Data Mining .....                                             | 30        |
| 3.5                                     | Interpretation/Evaluasi .....                                 | 30        |
| <b>BAB IV HASIL DAN PEMBAHASAN.....</b> |                                                               | <b>31</b> |
| 4.1                                     | Gambaran Umum Lokasi Penelitian .....                         | 31        |
| 4.1.1                                   | Sejarah Singkat Perusahaan .....                              | 31        |
| 4.1.2                                   | Visi Perusahaan .....                                         | 32        |
| 4.1.3                                   | Misi Perusahaan .....                                         | 32        |
| 4.1.4                                   | Struktur Organisasi PT Air Mas Perkasa .....                  | 32        |
| 4.2                                     | Hasil .....                                                   | 33        |
| 4.3                                     | Pembahasan .....                                              | 33        |
| 4.3.1                                   | <i>Data Selection</i> .....                                   | 33        |
| 4.3.2                                   | Pra-pengolahan Data .....                                     | 36        |
| 4.3.3                                   | Data Transformation .....                                     | 40        |
| 4.3.4                                   | Algoritma C4.5 .....                                          | 52        |
| 4.3.5                                   | Naïve Bayes .....                                             | 56        |
| 4.3.6                                   | Implementasi dan Hasil Pengujian Menggunakan RapidMiner ..... | 58        |
| <b>BAB V PENUTUP .....</b>              |                                                               | <b>63</b> |
| 5.1                                     | Kesimpulan .....                                              | 63        |
| 5.2                                     | Saran .....                                                   | 64        |
| <b>DAFTAR PUSTAKA.....</b>              |                                                               | <b>65</b> |
| <b>LAMPIRAN.....</b>                    |                                                               | <b>66</b> |

## DAFTAR GAMBAR

|                                                                 |    |
|-----------------------------------------------------------------|----|
| Gambar 2. 1 Tahapan Knowledge Discovery in Database .....       | 20 |
| Gambar 2. 2 Rumus Perhitungan Entropy .....                     | 23 |
| Gambar 2. 3 Rumus Perhitungan Nilai Gain.....                   | 23 |
| Gambar 3. 1 Desain Penelitian .....                             | 28 |
| Gambar 4. 1 Struktur Organisasi PT Air Mas Perkasa .....        | 32 |
| Gambar 4. 2 <b>Pohon Keputusan</b> .....                        | 56 |
| Gambar 4. 3 Operator Pengujian Naïve Bayes .....                | 59 |
| Gambar 4. 4 Jumlah Nilai Label Status Penjualan.....            | 59 |
| Gambar 4. 5 Jumlah Nilai Barang Elektronik paling diminati..... | 60 |
| Gambar 4. 6 Confussion Matrix Naïve Bayes.....                  | 60 |
| Gambar 4. 7 Operator Pengujian C4.5 .....                       | 61 |
| Gambar 4. 8 <b>Pohon Keputusan</b> Algoritma C4.5.....          | 61 |
| Gambar 4. 9 Jumlah Nilai Barang Elektronik paling diminati..... | 62 |
| Gambar 4. 10 Confussion Matrix Algoritma C4.5.....              | 62 |



## DAFTAR TABEL

|                                                                                  |    |
|----------------------------------------------------------------------------------|----|
| Tabel 2. 1 Penelitian Terdahulu .....                                            | 18 |
| Tabel 4. 1 Tabel Data Selection .....                                            | 36 |
| Tabel 4. 2 Tabel Pra-pengolahan Data .....                                       | 40 |
| Tabel 4. 3 Tabel Data Transformation pada 74 sampel data .....                   | 51 |
| Tabel 4. 4 Kategori Atribut Jumlah .....                                         | 51 |
| Tabel 4. 5 Kategori Atribut Harga .....                                          | 52 |
| Tabel 4. 6 Kombinasi nilai atribut dan kelas atribut jumlah pada 74 data sampel. | 52 |
| Tabel 4. 7 Model Perhitungan C4.5 pada 74 data sampel .....                      | 55 |
| Tabel 4. 8 Tabel Perhitungan klasifikasi (Prior) .....                           | 56 |
| Tabel 4. 9 Tabel Perhitungan Probabilitas Kategori .....                         | 57 |
| Tabel 4. 10 Tabel Perhitungan Probabilitas Brand .....                           | 57 |
| Tabel 4. 11 Tabel Perhitungan Probabilitas Harga .....                           | 57 |
| Tabel 4. 12 Tabel Perhitungan Probabilitas Jumlah .....                          | 58 |

