



**PERBANDINGAN ALGORITMA C4.5 DAN ALGORITMA NAIVE
BAYES DALAM DATA MINING UNTUK MEMPREDIKSI PENJUALAN
KOSMETIK DI TOKO JOLITA**

LAPORAN TUGAS AKHIR

HENDI SAPUTRA

41518010020

**UNIVERSITAS
MERCU BUANA**

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2024



**PERBANDINGAN ALGORITMA C4.5 DAN ALGORITMA NAIVE
BAYES DALAM DATA MINING UNTUK MEMPREDISKI PENJUALAN
KOMESTIK DI TOKO JOLITA**

LAPORAN TUGAS AKHIR

HENDI SAPUTRA

41518010020

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**UNIVERSITAS
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2024**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Hendi Saputra
NIM : 41518010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Algoritma C4.5 dan Algoritma Naive Bayes dalam Data Mining untuk Memprediksi Penjualan Kosmetik di Toko Jolita

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.



Jakarta, 8 Agustus 2024

A handwritten signature in black ink is written over a 1000 Rupiah postage stamp. The stamp features the Garuda Pancasila emblem and the text 'REPUBLIK INDONESIA', '1000', 'PERANGKAT TEMPEL', and the serial number 'B1524ALX247153994'.
Saputra.

UNIVERSITAS
MERCU BUANA

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Hendi Saputra
NIM : 41518010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Algoritma C4.5 dan Algoritma Naive Bayes dalam Data Mining untuk Memprediksi Penjualan Kosmetik di Toko Jolita.

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 Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing	:	Yustika Erliani, SE, MMSI	()
NIDN	:	0305057809	
Ketua Sidang	:	Harni Kusniyati, S.T., M.Kom.	()
NIDN	:	0324068101	
Penguji 1	:	Dr. Afiyati, S.Si., M.T.	()
NIDN	:	0316106908	
Penguji 2	:	Ida Farida, S.T., M.Kom.	()
NIDN	:	0324018301	

UNIVERSITAS
MERCU BUANA
Jakarta, 8 Agustus 2024
Mengetahui,

Dekan



Dr. Bambang Jekonowo, S.Si., MTI
NIDN : 0320037002

Ketua Program Studi



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadirat Tuhan yang Maha Esa, atas segala rahmat dan ridhanya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Ibu Dosen Pembimbing Yustika Erliani, SE, MMSI. selaku dosen pembimbing tugas akhir yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan tugas akhir ini terjadwal dengan baik.
5. Bapak Sabar Rudiarto, M.Kom selaku dosen pembimbing akademik
6. Kedua Orang Tua saya yang selalu mensupport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
7. Sahabat dan teman – teman saya yang telah membantu, mensupport, serta memberikan masukan dalam menyelesaikan Tugas Akhir.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 8 Agustus 2024

Hendi Saputra

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Hendi Saputra
NIM : 41518010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Algoritma C4.5 dan Algoritma Naive Bayes dalam Data Mining untuk Memprediksi Penjualan Kosmetik di Toko Jolita.

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 Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 8 Agustus 2024

Yang menyatakan,

UNIVERSITAS
MERCU BUANA



Hendi Saputra

ABSTRAK

Nama : Hendi Saputra
NIM : 41518010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Perbandingan Algoritma C4.5 dan Algoritma Naive Bayes dalam Data Mining untuk Memprediksi Penjualan Kosmetik di Toko Jolita
Dosen Pembimbing : Yustika Erliani, SE, MMSI

Abstrak – Penelitian ini bertujuan untuk membantu memprediksi permasalahan yang terjadi pada toko Jolita dimana toko ini menjual produk kecantikan. Sering kali toko ini kehabisan stok produk yang membuat pelanggan tidak mendapatkan produk yang dicarinya, maka peneliti melakukan penelitian untuk membantu mencari tahu produk mana saja yang paling banyak diminati pelanggan dengan menggunakan dua algoritma yaitu algoritma C4.5 dan algoritma Naive Bayes. Evaluasi kinerja menunjukkan bahwa produk wardah yang paling banyak diminati dimana data penjualan toko Jolita menggunakan algoritma C4.5 dan Naive Bayes didapatkan hasil confusion matrix pengujian C4.5 akurasi yang di dapatkan yaitu 0.99 model ini bekerja sangat baik pada data pelatihan, dengan kesalahan prediksi yang sangat kecil dan kemampuan yang sangat baik untuk menjelaskan variasi data. Namun, penting untuk juga memeriksa kinerja model pada data pengujian untuk memastikan model tidak overfitting, yaitu performa yang sangat baik pada data pelatihan tetapi kurang baik pada data yang belum dilihat sebelumnya.. Dari dua algoritma jika dilihat dari hasil akurasi maka C4.5 lebih baik dibandingkan dengan algoritma Naive Bayes. Disamping itu algoritma Naive Bayes tidak cocok digunakan pada penelitian ini karna tidak bekerja dengan baik pada data pelatihan. Kesalahan (MAE dan MSE) relatif tinggi, dan R^2 Score 0,00 menunjukkan bahwa model tidak memiliki kemampuan untuk memprediksi atau menjelaskan variasi data target. Dengan kata lain, model gagal dalam belajar dari data pelatihan dan prediksinya tidak lebih baik dari sekadar rata-rata sederhana. Ini bisa terjadi karena model terlalu sederhana (underfitting) atau ada masalah dalam pemilihan fitur atau model.

Kata kunci: Toko Jolita, C4.5, Naive Bayes

ABSTRACT

Nama : Hendi Saputra
NIM : 41518010020
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Comparison of C4.5 Algorithm and Naive Bayes Algorithm in Data Mining to Predict Cosmetic Sales at Jolita Store
Dosen Pembimbing : Yustika Erliani, SE, MMSI

***Abstract** - This research aims to help predict problems that occur at Jolita stores where this store sells beauty products. Often times this store runs out of product stock which makes customers not get the product they are looking for, so researchers conducted research to help find out which products are most in demand by using two algorithms, namely the C4.5 algorithm and the Naive Bayes algorithm. Performance evaluation shows that Wardah products are the most in demand where Jolita store sales data using the C4.5 algorithm and Naive Bayes obtained confusion matrix results of C4.5 testing the accuracy obtained is 0.99 this model works very well on training data, with very small prediction errors and excellent ability to explain data variations. However, it is important to also check the model's performance on test data to ensure the model is not overfitting, i.e. performing very well on training data but less well on data that has not been seen before. Of the two algorithms when viewed from the accuracy results, C4.5 is better than the Naive Bayes algorithm. In addition, the Naive Bayes algorithm is not suitable for use in this study because it does not work well on training data. The errors (MAE and MSE) are relatively high, and the R^2 Score of 0.00 indicates that the model has no ability to predict or explain variations in the target data. In other words, the model fails to learn from the training data and its predictions are no better than a simple average. This can happen because the model is too simple (underfitting) or there is a problem in feature or model selection.*

Keyword: Store Jolita, C4.5, Naive Bayes

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT.....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	2
1.5 Batasan Masalah	3
BAB II TINJAUAN PUSTAKA	4
2.1 Penelitian Terdahulu	4
2.2 Teori Pendukung.....	12
BAB III METODE PENELITIAN	14
3.1 Jenis Penelitian.....	14
3.2 Tahapan Penelitian.....	14
BAB IV hasil DAN PEMBAHASAN.....	17
BAB V KESIMPULAN	31
5.1 Kesimpulan	31
5.2 Saran	31
DAFTAR PUSTAKA.....	32
LAMPIRAN.....	34

Lampiran 1 Kartu Asistensi	34
Lampiran 2 Lampiran Naskah Artikel Jurnal	35
Lampiran 3 Curriculum Vitae	39
Lampiran 4 Surat Pernyataan Haki	40
Lampiran 5 Sertifikat BNSP	42



DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait.....	4
------------------------------------	---



DAFTAR GAMBAR

Gambar 4. 1 Pemanggilan Dataset	17
Gambar 4. 2 Menampilkan dataset	17
Gambar 4. 3 Hasil Pre-processing.....	18
Gambar 4. 4 Menyiapkan Data	19
Gambar 4. 5 Grafik Penjualan Perhari	20
Gambar 4. 6 Perbandingan Jumlah Penjualan Barang.....	21
Gambar 4. 7 Barang dengan Total Pendapatan Tertinggi.....	22
Gambar 4. 8 Grafik Perbandingan Total Pendapatan Barang	22
Gambar 4. 9 Nilai Predikasi untuk Data Testing	24
Gambar 4. 10 Hasil Algoritma C4.5	27
Gambar 4. 11 Pohon keputusan	28
Gambar 4. 12 Hasil Algoritma Naive Bayes.....	29
Gambar 4. 13 Visualisasi Hasil Prediksi.....	30



DAFTAR LAMPIRAN

Lampiran 2 Lampiran Naskah Artikel Jurnal	35
Lampiran 3 Curriculum Vitae	39
Lampiran 4 Surat Pernyataan Haki	40
Lampiran 5 Sertifikat BNSP	42

