



**PERBANDINGAN PERFORMA ALGORITMA SVM, K-NN,  
DAN C4.5 DALAM MENGLASIFIKASIKAN PENJUALAN  
PRODUK PADA TOKO ONLINE TABASYUK**

**LAPORAN SKRIPSI**

Disusun Oleh :

**MUHAMMAD IRFAN JATNIKA SUDRAJAT**

**41519010128**

**MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS ILMU KOMPUTER**

**UNIVERSITAS MERCU BUANA**

**JAKARTA**

**2023**



**PERBANDINGAN PERFORMA ALGORITMA SVM, K-NN,  
DAN C4.5 DALAM MENGLASIFIKASIKAN PENJUALAN  
PRODUK PADA TOKO ONLINE TABASYUK**

**LAPORAN SKRIPSI**

Disusun Oleh :  
**MUHAMMAD IRFAN JATNIKA SUDRAJAT**  
**41519010128**

**Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS MERCU BUANA  
JAKARTA  
2023**

## HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Irfan Jatnika Sudrajat  
NIM : 41519010128  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma SVM, K-NN,  
dan C4.5 Dalam Mengklasifikasikan Penjualan  
Produk Pada Toko Online Tabasyuk.

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, 22 Mei 2023



Muhammad Irfan Jatnika

Sudrajat



## HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Muhammad Irfan Jatnika Sudrajat  
NIM : 41519010128  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma SVM, K-NN,  
dan C4.5 Dalam Mengklasifikasikan Penjualan  
Produk Pada Toko Online Tabasyuk.

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 : Wawan Gunawan S.Kom., MT  
NIDN : 0424108104  
Penguji 1 : Dr. Hadi Santoso, S.Kom., M.Kom  
NIDN : 0225067701  
Penguji 2 : Dwi Anindyani Rochmah, ST., M.TI  
NIDN : 0011057801



Jakarta, 14 Agustus 2023

Mengetahui,

Dekan

Ketua Program Studi

UNIVERSITAS  
MERCU BUANA

  
Dr. Bambang Jokonowo, S.Si., MTI

  
Bagus Priambodo, ST, MTI

## KATA PENGANTAR

Puji syukur kita panjatkan kehadirat Allah SWT atas segala rahmat dan karunia- Nya, dikarenakan Tugas Akhir yang berjudul “Perbandingan Akurasi Klasifikasi Algoritma SVM, KNN, C4.5 Pada Penjualan Produk Toko Online Tabasyuk” dapat diselesaikan dalam jangka waktu yang ditentukan. Laporan Tugas Akhir ini dibuat sebagai syarat untuk LULUS sebagai Sarjana Ilmu Komputer dari Universitas Mercu Buana. Penulis menyadari bahwa pembuatan Tugas Akhir ini tidak terlepas dari bantuan dan bimbingan berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Ungkapan terima kasih dan penghargaan yang sangat spesial penulis haturkan dengan rendah hati dan rasa hormat kepada kedua orang tua tercinta, Ayahanda Ajat Sudrajat dan Lia Rahmalia yang selalu mendoakan, memberikan semangat, dukungan dan motivasi tiada henti kepada penulis hingga dapat menyelesaikan studi sampai sarjana dan juga dengan segala pengorbanannya tak akan pernah penulis lupakan atas jasa-jasanya. Doa restu, nasihat dan petunjuk yang telah diberikan.
2. Prof. Dr.Ir. Andi Adriansyah, M.Eng selaku Rektor Universitas Mercu Buana.
3. Dr. Bambang Jokonowo, S.Si, MTI selaku Dekan Fakultas Ilmu Komputer.
4. Bagus Priambodo, ST, MTI selaku Ketua Program Studi Teknik Informatika.
5. Wawan Gunawan S.Kom., MT selaku Dosen Pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini.
6. Ibu Saruni Dwiasnati, ST.MM., M.Kom. selaku Dosen Pembimbing Akademik yang telah memantau dan membantu studi saya hingga akhir.
7. Toko Tabasyuk beserta pemilik toko, Reclaudia Dian Arianti yang telah mengizinkan menggunakan datanya untuk diolah dan dijadikan data penelitian.
8. Fadel Muzzaki, Agung Ikbal Apandi, Ferdi, Raihan Dzaky selaku sahabat tercinta penulis yang telah memberikan pengetahuannya beserta waktunya untuk menyelesaikan penulisan tugas akhir ini.

9. Seluruh pihak yang memberikan bantuan kepada penulis namun tidak dapat disebutkan satu persatu. Terima kasih atas bantuan, semangat, dan doa baik yang diberikan kepada penulis selama proses Tugas Akhir ini.

Akhir kata, penulis berharap Tugas Akhir ini dapat menjadi referensi untuk dikembangkan lebih baik lagi kedepannya, karena penulis menyadari hasil Tugas Akhir ini masih jauh dari sempurna. Masih terdapat kekurangan dalam eksperimen, cara penjelasan maupun kekeliruan penulisan. Untuk itu, kritik dan saran pembaca sangat dihargai dan diharapkan. Semoga Tugas Akhir ini dapat memberikan manfaat bagi para pembaca.

Jakarta, 6 Agustus 2023



Muhammad Irfan  
Jatnika Sudrajat

UNIVERSITAS  
MERCU BUANA



## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

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

Nama : Muhammad Irfan Jatnika Sudrajat  
NIM : 41519010128  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma SVM, K-NN,  
dan C4.5 Dalam Mengklasifikasikan Penjualan  
Produk Pada Toko Online Tabasyuk.

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, mengalih media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan namanya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 5 Agustus 2023

Yang menyatakan,



Muhammad Irfan Jatnika Sudrajat

## ABSTRAK

Nama : Muhammad Irfan Jatnika Sudrajat  
NIM : 41519010128  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma SVM, K-NN, dan C4.5  
Dalam Mengklasifikasikan Penjualan Produk Pada Toko  
Online Tabasyuk.  
Pembimbing : Wawan Gunawan S.Kom., MT

Terdapat Banyak *platform* berjualan secara *online (e-Commerce)* yang sudah beredar di kalangan masyarakat, mulai dari berjualan pakaian, kebutuhan pangan, papan, dan lain-lain secara *online*. Dengan perkembangan teknologi ini, banyak sekali data yang dihasilkan oleh teknologi informasi yang canggih, mulai dari bidang industri, ekonomi, ilmu pengetahuan, maupun berbagai aspek di kehidupan lainnya. Dari data yang melimpah, diperlukan pengolahan dan analisa sehingga dapat memberikan nilai lebih dan menghasilkan informasi yang penting bagi para penggunanya. Salah satu yang dapat dimanfaatkan adalah teknologi data mining. Dalam menggali informasi yang bermanfaat dari gudang data perusahaan penjualan, digunakan aplikasi pendukung yang bisa mengimplementasikan data mining, penelitian ini menggunakan google colaboratory untuk melakukan klasifikasi penjualan berdasarkan parameter kurang laris, cukup laris, dan sangat laris. Klasifikasi dapat dilakukan dengan algoritma Support Vector Machine, K-Nearest Neighbor, dan C4.5. Pada penelitian ini, dilakukan perbandingan algoritma dengan pengolahan data yang didapat dari rekap penjualan toko Tabasyuk pada platform *e-commerce* Shopee. Tujuan dari penelitian ini adalah membuat hasil perbandingan algoritma untuk klasifikasi penjualan berdasarkan parameter laris. Selanjutnya hasil klasifikasi data dievaluasi dan divalidasi dengan menggunakan confusion matrix dan *split decision validation* dengan menggunakan bahasa pemrograman Python dan *library* yang dibutuhkan. Dari hasil penilitan ini didapatkan algoritma terbaik yaitu C4.5 dengan tingkat akurasi sebesar 84,37%. Hasil yang didapatkan dengan nilai akurasi klasifikasi bagus (Excellent classification), sehingga dapat disimpulkan bahwa dengan dataset yang didapatkan dari toko Tabasyuk di platform *e-commerce* Shopee algoritma C.45 merupakan algoritma terbaik yang bisa digunakan untuk melakukan klasifikasi. Selain perbandingan algoritma, juga didapatkan hasil produk terlaris menggunakan teknik visualisasi data berdasarkan nama produk dan provinsi dengan penjualan produk terbanyak, sehingga ini bisa dijadikan acuan kepada penjual untuk penjualan kedepannya.

Kata Kunci – Algoritma, Support Vector Machine, K-Nearest Neighbor, C4.5, Klasifikasi, Penjualan, Produk, *e-commerce*.



## ABSTRACT

Name : Muhammad Irfan Jatnika Sudrajat  
NIM : 41519010128  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Performa Algoritma SVM, K-NN, dan C4.5  
Dalam Mengklasifikasikan Penjualan Produk Pada Toko  
Online Tabasyuk.  
Counsellor : Wawan Gunawan S.Kom., MT.

There are numerous online selling platforms (e-Commerce) that have been circulating among the public, ranging from clothing, food necessities, building materials, and various other products sold online. With the advancement of technology, a vast amount of data is generated by sophisticated information technology, spanning various industries, economics, sciences, and other aspects of life. From this abundant data, processing and analysis are required to derive added value and essential information for its users. One of the tools that can be utilized is data mining technology. To extract useful information from the sales data warehouse, supporting applications implementing data mining are used. In this research, Google Colaboratory is employed to perform sales classification based on three categories: low-selling, moderate-selling, and highly-selling. The classification is carried out using the algorithms Support Vector Machine, K-Nearest Neighbor, and C4.5. The study involves a comparison of these algorithms based on the sales data obtained from the Tabasyuk store on the Shopee e-commerce platform. The objective of this research is to present a comparative analysis of algorithms for sales classification based on their performance. Furthermore, the classified results are evaluated and validated using confusion matrix and split decision validation techniques, implemented with the Python programming language and required libraries. The research findings indicate that the best algorithm is C4.5, achieving an accuracy rate of 84.37%. The results demonstrate excellent classification accuracy, leading to the conclusion that with the dataset obtained from the Tabasyuk store on the Shopee e-commerce platform, the C4.5 algorithm is the most suitable for performing sales classification. Apart from the algorithm comparison, the study also yields insights into the best-selling products through data visualization techniques, based on product names and provinces with the highest product sales. This information can serve as a reference for sellers in their future sales strategies.

Keywords – Algorithm, Support Vector Machine, K-Nearest Neighbor, C4.5, Classification, Sales, Product, e-commerce.

## 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.....	vi
ABSTRAK .....	vii
ABSTRACT .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL .....	xi
DAFTAR GAMBAR.....	xii
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1    LATAR BELAKANG.....	1
1.2    PERUMUSAN MASALAH.....	3
1.3    TUJUAN PENELITIAN .....	4
1.4    MANFAAT PENELITIAN.....	4
1.5    BATASAN PENELITIAN.....	4
<b>BAB II TINJAUAN PUSTAKA.....</b>	<b>5</b>
2.1    PENELITIAN TERDAHULU .....	5
2.2    TEORI PENDUKUNG .....	13
<b>BAB III METODE PENELITIAN .....</b>	<b>17</b>
3.1    JENIS PENELITIAN .....	17
3.2    TAHAP PENELITIAN .....	17
3.3    IDENTIFIKASI MASALAH .....	18
3.4    PENGUMPULAN DATA .....	18
3.5    PENGOLAHAN DATA (PREPROCESSING) .....	18
3.6    SELEKSI FITUR.....	19
3.7    PEMBENTUKAN MODEL.....	19

3.8	IMPLEMENTASI ALGORITMA.....	19
3.9	SUPPORT VECTOR MACHINE (SVM).....	20
3.10	K-NEAREST NEIGHBOR (K-NN).....	21
3.11	ALGORITMA C4.5.....	21
3.12	PENGUJIAN MODEL.....	22
3.13	EVALUASI DAN VALIDASI MODEL.....	23
<b>BAB IV HASIL DAN PEMBAHASAN .....</b>		<b>24</b>
4.1	DATASET .....	24
4.2	IMPORT LIBRARY DAN DATASET .....	25
4.3	PRE-PROCESSING .....	28
4.4	SELEKSI FITUR.....	37
4.5	PEMBENTUKAN MODEL.....	39
4.6	PEMODELAN SPLIT DATA.....	39
4.7	IMPLEMENTASI ALGORITMA SVM.....	40
4.8	IMPLEMENTASI ALGORITMA K-NN.....	42
4.9	IMPLEMENTASI ALGORITMA C4.5.....	45
4.10	ANALISIS HASIL DAN VISUAL ALGORITMA .....	47
<b>BAB V KESIMPULAN DAN SARAN .....</b>		<b>56</b>
5.1	KESIMPULAN.....	56
5.2	SARAN .....	56
<b>DAFTAR PUSTAKA.....</b>		<b>57</b>
<b>LAMPIRAN .....</b>		<b>61</b>
LAMPIRAN 1.FORM ABSENSI BIMBINGAN .....		61
LAMPIRAN 2.LAMPIRAN HALAMAN PERSETUJUAN .....		62
LAMPIRAN 3.LAMPIRAN LUARAN TUGAS AKHIR .....		63
LAMPIRAN 4.BUKTI SUBMIT / PUBLISHED ARTIKEL ILMIAH / HKI.....		64
LAMPIRAN 5.NASKAH ARTIKEL JURNAL.....		65
LAMPIRAN 6.CURRICULUM VITAE .....		75
LAMPIRAN 7.CEK PLAGIARISME .....		76
LAMPIRAN 8.SURAT PERNYATAAN HKI .....		77
LAMPIRAN 9.SERTIFIKAT BNSP.....		78



## DAFTAR TABEL

TABEL 2.1 REFERENSI JURNAL PENELITIAN TERDAHULU .....	5
TABEL 3.1 BENTUK KE-5 EKSPERIMEN .....	23
TABEL 4.1 DATASET SEBELUM <i>PRE-PROCESSING</i> .....	24
TABEL 4.2 DATASET SETELAH <i>DATA CLEANING</i> DAN <i>DATA REDUCTION (HEAD)</i> .....	29
TABEL 4.3 DATASET SETELAH <i>DATA CLEANING</i> DAN <i>DATA REDUCTION (TAIL)</i> .....	30
TABEL 4.4 DATASET SEBELUM <i>DATA TRANSFORMATION</i> .....	34
TABEL 4.5 DATASET SETELAH <i>DATA TRANSFORMATION</i> .....	36
TABEL 4.6 HASIL KE-5 EKSPERIMEN DENGAN SVM .....	49
TABEL 4.7 HASIL KE-5 EKSPERIMEN DENGAN KNN .....	51
TABEL 4.8 HASIL KE-5 EKSPERIMEN DENGAN C4.5 .....	54
TABEL 4.9 PERBANDINGAN KESELURUHAN ALGORITMA .....	55



UNIVERSITAS  
MERCU BUANA

## DAFTAR GAMBAR

GAMBAR 2.1 ALGORITMA <i>SUPPORT VECTOR MACHINE</i> .....	14
GAMBAR 2.2 ALGORITMA <i>K-NEAREST NEIGHBOR</i> .....	15
GAMBAR 3.1 ALUR PENELITIAN.....	17
GAMBAR 3.2 ALUR PROGRAM.....	20
GAMBAR 3.3 FLOWCHART <i>SUPPORT VECTOR MACHINE</i> .....	20
GAMBAR 3.4 FLOWCHART <i>K-NEAREST NEIGHBOR</i> .....	21
GAMBAR 3.5 FLOWCHART <i>C4.5</i> .....	22
GAMBAR 4.1 UPLOAD DATASET .....	26
GAMBAR 4.2 INFORMASI DATASET .....	27
GAMBAR 4.3 TAMPILAN DATASET <i>HEAD</i> .....	27
GAMBAR 4.4 TAMPILAN DATASET <i>TAIL</i> .....	28
GAMBAR 4.5 VISUALISASI PESANAN PER-TRANSAKSI.....	31
GAMBAR 4.6 VISUALISASI PESANAN TERBANYAK.....	32
GAMBAR 4.7 VISUALISASI PESANAN TERBANYAK BERDASARKAN PROVINSI.....	33
GAMBAR 4.8 DATASET SEBELUM <i>DATA TRANSFORMATION</i> .....	34
GAMBAR 4.9 DATASET SETELAH <i>DATA TRANSFORMATION</i> .....	36
GAMBAR 4.10 HASIL PENCARIAN 10 FITUR TERBAIK .....	37
GAMBAR 4.11 VISUALISASI 10 FITUR TERBAIK .....	38
GAMBAR 4.12 KOLOM ATRIBUT SETELAH <i>FEATURE SELECTION</i> .....	39
GAMBAR 4.13 <i>CONFUSION MATRIX SVM KE-1</i> .....	48
GAMBAR 4.14 <i>CONFUSION MATRIX SVM KE-2</i> .....	48
GAMBAR 4.15 <i>CONFUSION MATRIX SVM KE-3</i> .....	48
GAMBAR 4.16 <i>CONFUSION MATRIX SVM KE-4</i> .....	48
GAMBAR 4.17 <i>CONFUSION MATRIX SVM KE-5</i> .....	49
GAMBAR 4.18 <i>CONFUSION MATRIX KNN KE-1</i> .....	50
GAMBAR 4.19 <i>CONFUSION MATRIX KNN KE-2</i> .....	50
GAMBAR 4.20 <i>CONFUSION MATRIX KNN KE-3</i> .....	50
GAMBAR 4.21 <i>CONFUSION MATRIX KNN KE-4</i> .....	50
GAMBAR 4.22 <i>CONFUSION MATRIX KNN KE-5</i> .....	50
GAMBAR 4.23 <i>CONFUSION MATRIX C4.5 KE-1</i> .....	52
GAMBAR 4.24 <i>CONFUSION MATRIX C4.5 KE-2</i> .....	52

GAMBAR 4. 25 <i>CONFUSION MATRIX</i> C4.5 KE-3 .....	52
GAMBAR 4.26 <i>CONFUSION MATRIX</i> C4.5 KE-4.....	52
GAMBAR 4.27 <i>CONFUSION MATRIX</i> C4.5 KE-5.....	52
GAMBAR 4.28 <i>DECISION TREE</i> C4.5 KE-1 .....	53
GAMBAR 4.29 <i>DECISION TREE</i> C4.5 KE-2 .....	53
GAMBAR 4.30 <i>DECISION TREE</i> C4.5 KE-3 .....	53
GAMBAR 4.31 <i>DECISION TREE</i> C4.5 KE-4 .....	54
GAMBAR 4.32 <i>DECISION TREE</i> C4.5 KE-5 .....	54



UNIVERSITAS  
MERCU BUANA