

## 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 colab 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 penelitian 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.