

## ABSTRAK

Nama : Maulana Ibrahim  
NIM : 41519010030  
Program Studi : Teknik Informatika  
Judul Laporan Skripsi : Perbandingan Kinerja Algoritma Klasifikasi untuk Deteksi Berita Hoax dalam Bahasa Indonesia menggunakan Naïve Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN).  
Pembimbing : Eliyani Dr. Ir.

Penelitian ini bertujuan untuk membandingkan kinerja ketiga algoritma pengklasifikasian yang berbeda, yaitu *Naïve Bayes*, *SVM (Support Vector Machine)*, dan *KNN (K-Nearest Neighbor)*, dalam mengklasifikasi berita hoax menggunakan bahasa Indonesia. Penyebaran berita hoax atau informasi palsu telah menjadi masalah yang signifikan dalam era digital saat ini. Dengan adanya penyebaran luas melalui platform media sosial dan situs web, berita hoax dengan mudah menyebar dan mempengaruhi opini publik. Penelitian ini menggunakan dataset berita berbahasa Indonesia yang terdiri dari berita hoax dan bukan hoax. Data ini akan digunakan untuk melatih dan menguji kinerja ketiga algoritma pengklasifikasian tersebut. Hasil klasifikasi ditampilkan dalam *precision*, *recall*, dan *f1-score* untuk masing-masing kelas, serta akurasi keseluruhan. Setelah melakukan evaluasi, ditemukan bahwa *SVM* mencapai hasil klasifikasi terbaik dengan akurasi 100%. Karena *SVM* dengan representasi *TF-IDF* yang lebih kompleks tampak menjadi pilihan terbaik. Kemampuan *SVM* dalam menangani data kompleks dan keuntungan dari representasi *TF-IDF* yang akurat mendukung kemungkinan performa yang baik bahkan untuk situasi yang lebih kompleks di masa depan. Namun, perlu dilakukan beberapa peningkatan seperti analisis lebih mendalam terhadap data, teknik resampling untuk mengatasi ketidakseimbangan kelas, dan tuning parameter untuk mengoptimalkan performa model.

**Kata kunci:** pengklasifikasian, berita hoax, bahasa Indonesia, Naïve Bayes, SVM, KNN

## ABSTRACT

Name : Maulana Ibrahim  
NIM : 41519010030  
Study Program : Teknik Informatika  
Title Thesis : Perbandingan Kinerja Algoritma Klasifikasi untuk Deteksi Berita Hoax dalam Bahasa Indonesia menggunakan Naïve Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbors(KNN).  
Counsellor : Eliyani Dr. Ir.

This study aims to compare the performance of three different classification algorithms, namely *Naïve Bayes*, *SVM (Support Vector Machine)*, and *KNN (K-Nearest Neighbor)*, in classifying hoax news using Indonesian. The spread of hoax news or false information has become a significant problem in today's digital era. With widespread dissemination via social media platforms and websites, hoax news easily spreads and influences public opinion. This study uses an Indonesian language news dataset consisting of hoax and non-hoax news. This data will be used to train and test the performance of the three classification algorithms. Classification results are displayed in precision, recall, and f1-score for each class, as well as overall accuracy. After conducting an evaluation, it was found that *SVM* achieves the best classification results with 100% accuracy. Because *SVM* with more complex *TF-IDF* representation seems to be the best choice. *SVM*'s ability to handle complex data and the advantages of an accurate *TF-IDF* representation means the possibility of good performance for even more complex situations in the future. However, some improvements need to be made such as deeper analysis of the data, resampling techniques to overcome class imbalances, and parameter tuning to optimize model performance.

**Keywords:** *classification, hoax news, Indonesian language, Naïve Bayes, SVM, KNN*