

## ABSTRAK

Peningkatan jumlah dokumen membuat mahasiswa semakin sulit memperoleh informasi sesuai dengan apa yang diinginkan khususnya referensi tugas akhir. Masalah ini memerlukan teknik pengolahan teks yang mengorganisasikan dokumen sesuai dengan kategorinya. Salah satunya klasifikasi teks. Klasifikasi teks dapat mengordinasikan dokumen sesuai dengan yang telah ditentukan sebelumnya secara otomatis. Salah satu klasifikasi vector teks yang terkenal adalah *support vector machine* (SVM) yang berusaha mencari bidang pemisah terbaik pada *input space*. Menurut Chenometh et al (209) *support vector machine* (SVM) merupakan algoritma klasifikasi terbaik dibanding dengan metode klasifikasi metode vektor yang lain yaitu Rocchio, k-Nearest Neighbor dan decision tree. Melalui serangkaian pengujian terhadap sejumlah dokumen yang telah diolah menjadi suatu representasi data berupa matriks vektor, SVM berhasil mengklasifikasikan dokumen dengan akurasi tertinggi sebesar 82,14% dalam waktu 23 detik.

Kata kunci : klasifikasi dokumen teks, *text mining*, *support vector mechine*, *chi-square*

## **ABSTRACT**

The increase in the number of documents to make students more difficult to obtain information specifically thesis. This problem requires a text processing techniques to organize documents according to category. Text classification can be coordinated the documents according to predefined categories automatically. One well known text classification is a support vector machine (SVM) is trying to find the best interface on the input space. According Chenomoth et al (2009) support vector machine (SVM) is the best classification algorithm compared with other vector classification that Rocchio, K-Nearest Neighbor and Decision Tree. Through a series of test on a number of documents that have been processed into a data representation in the form of a matrix vector, SVM managed to classify document with the highest accuracy of 82,14% with 23 seconds.

**Keyword :** text document classification, text mining, support vector machine, chi-square