

## ABSTRAK

Perkembangan media sosial *Twitter* menjadi munculnya informasi tekstual dalam jumlah yang tidak terbatas, sehingga nilai dari informasi ini perlu dikaji. Istilah di *Twitter* yang disebut *Tweet* yang merupakan pesan atau status yang dibuat oleh penggunaanya. Setiap *tweet* yang diekspresikan dalam *twitter* memiliki nilai makna positif dan negatif serta komentar-komentar yang ada dalam *Twitter*. Algoritma yang paling banyak diteliti dalam analisis sentimen yaitu *Support Vector Machine* karena bekerja dengan sangat baik. Meskipun memiliki sedikit kekurangan dalam hal lebih sensitif terhadap fitur yang bisa berpengaruh pada pemrosesan, namun hal ini penulis mencoba mengimplementasi algoritma *chi square* untuk seleksi fitur. Setelah dilakukan pengujian terhadap 2000 sampel data dengan penggunaan 0 sampai 100 % dari jumlah fitur yang sudah diseleksi maka didapatkan hasil akurasi SVM yang paling bagus berada pada penggunaan 80% fitur. Hal ini dibuktikan dengan nilai accuracy score sebesar 0.889 (88.9%), precision score 0.8843873517786561 (88.44%), recall score 0.895 (89.5%) dan f1 score 0.889662027833002 (88.96%).

**Kata Kunci :** Analisis Sentimen, Support Vector Machine, Chi Square

## **ABSTRAC**

The development of social media Twitter has provided unlimited amounts of textual information, so the value of this information needs to be studied. The term on Twitter is called a Tweet which is a message or status created by its users. Every tweet expressed on Twitter has positive and negative meaning values as well as comments on Twitter. The most researched algorithm in sentiment analysis is the Support Vector Machine because it works very well. Even though it has a slight drawback in terms of being more sensitive to features that can influence processing, the author tries to implement the chi square algorithm for feature selection. After testing 2000 sample data using 0 to 100% of the number of features that have been selected, the best SVM accuracy results were obtained when using 80% of the features. This is proven by the accuracy score of 0.889 (88.9%), precision score of 0.8843873517786561 (88.44%), recall score of 0.895 (89.5%) and f1 score of 0.889662027833002 (88.96%).

**Keywords : Sentiment Analysis, Support Vector Machine, Chi Square**