

BERBAHASA INDONESIA

ABSTRAK

Nama : Muhammad Variansjah
NIM : 41519010191
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Sentimen Analisis Mengenai Polusi Udara Menggunakan Algoritma *Support Vector Machine* dan *Random Forest*
Pembimbing : Lukman Hakim, ST, M.Kom

Polusi udara adalah kontaminasi area dalam dan luar ruangan oleh zat kimia, fisik, atau biologis yang mengubah sifat alami atmosfer. Insinerator domestik, mobil, pabrik, dan kebakaran hutan merupakan sumber polusi udara yang umum. Di Indonesia, tidak diragukan lagi kalau polusi udara terjadi karena banyaknya kebakaran hutan di Indonesia. Akibat kasus tersebut, banyak opini masyarakat yang berbeda-beda. Berbagai sentiment terjadi di dunia maya, salah satunya *Twitter*. *Twitter* adalah *social media* yang paling banyak menampung berbagai macam opini positif, negatif maupun netral. Oleh karena itu, peneliti ingin memecahkan masalah dengan implementasi algoritma SVM dan *Random Forest*. *Dataset* didapatkan dari hasil scrapping menggunakan *tweet harvest*. Data yang diperoleh didapatkan sebanyak 8555 *tweet*. Dengan membagi model *dataset* 80% dan 20%, hasil didapat bahwa akurasi algoritma SVM lebih baik dari algoritma *Random Forest*. Akurasi dari algoritma SVM sebesar 83% sedangkan algoritma *Random Forest* sebesar 81%.

Katakunci : *sentiment analisis, SVM, Random Forest, Twitter, Polusi*

ABSTRACT

Nama : Muhammad Variansjah
NIM : 41519010191
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Sentimen Analisis Mengenai Polusi Udara Menggunakan Algoritma *Support Vector Machine* dan *Random Forest*
Pembimbing : Lukman Hakim, ST, M.Kom

The term "air pollution" refers to the contaminating effects of chemical, physical, or biological contaminants on the atmosphere that alter its inherent characteristics in both indoor and outdoor spaces. Common sources of air pollution include automobiles, factories, homes with incinerators, and forest fires. There is no denying that Indonesia's numerous forest fires contribute to air pollution there. This situation has caused divisions in the opinions of a lot of people. Twitter is only one among the numerous diverse emotions that can be felt on the internet. Twitter is the social media site that allows the most variety of unbiased, affirmative, and negative opinions. Therefore, the research's objective is to solve the problem by using the Random Forest and SVM methods. Tweet Harvest was used to scrape results and compile the information. The data collection contained 8555 tweets. by dividing.

Keywords: *sentiment analisis, SVM, Random Forest, Twitter, Polusi*