

## **ANALISIS SENTIMEN MENGGUNAKAN ALGORITMA NAIVE BAYES PADA TWITTER TERHADAP PENGGUNAAN TRANSPORTASI UMUM DALAM KOTA SEMARANG**

**M ANANG AMIRUDIN**

(Pembimbing : ARIPIN, M.Kom)

*Teknik Informatika - S1, FIK, Universitas Dian Nuswantoro*

*www.dinus.ac.id*

*Email : 111201207051@mhs.dinus.ac.id*

### **ABSTRAK**

Analisis sentimen pada twitter terhadap penggunaan transportasi umum merupakan analisis yang dilakukan dengan menggunakan opini masyarakat pengguna transportasi umum di kota Semarang khususnya trans Semarang dengan cara crawling data yang ada di twitter. Crawling data twitter ini menggunakan Twitter Search API yang telah disediakan oleh pihak twitter. Data tweet atau opini yang diambil merupakan tweet yang mengandung kata trans Semarang. Metode yang digunakan adalah Naive Bayes. Penelitian dimulai dari pengumpulan data yang didapat dari twitter. Dari keseluruhan data yang terkumpul akan dilakukan proses seleksi kemudian dilakukan proses preprocessing (case folding, normalisasi fitur, tokenizing dan stemming) dan selanjutnya diklasifikasi. Naive Bayes merupakan salah satu classifier yaitu metode yang dapat digunakan untuk mengklasifikasi data. Data set yang terkumpul sebanyak 305 data set, yang dibagi menjadi data training dan data testing. Untuk data training sebanyak 210 data yang dibagi menjadi kelas positif dan kelas negatif, masing-masing kelas mendapat jumlah data yang sama yaitu 105. Data testing yang digunakan sebanyak 95. Penelitian ini akan menghasilkan klasifikasi dimana data akan dikategorikan menjadi kelas positif dan kelas negatif menggunakan algoritma Naive Bayes. Dataset diuji sebanyak sepuluh kali menggunakan K-Fold Cross Validation dan dihasilkan akurasi rata-rata sebanyak 95,81%. Dari hasil yang diperoleh tersebut maka Naive Bayes layak digunakan sebagai metode dalam analisis sentimen pada twitter terhadap penggunaan transportasi umum di Semarang.

Kata Kunci : Sentimen Analisis; Naive Bayes; twitter; Transportasi Umum; Cross Validation

## **SENTIMENT ANALYSIS USING NAIVE BAYES ALGORITHM ON TWITTER ABOUT THE USE OF PUBLIC TRANSPORTATION AT SEMARANG**

**M ANANG AMIRUDIN**

(Lecturer : ARIPIN, M.Kom)

*Bachelor of Informatics Engineering - S1, Faculty of Computer  
Science, DINUS University*

*www.dinus.ac.id*

*Email : 111201207051@mhs.dinus.ac.id*

### **ABSTRACT**

Twitter sentiment analysis about the use of public transportation is an analysis performed using the public opinion of public transportation users in Semarang especially Trans Semarang with the data crawling on twitter. Data Crawling twitter using the Twitter Search API has been provided by the twitter. Data tweet or opinions taken from the tweet that containing word Trans Semarang. The method used is Naive Bayes. The Research started by collecting data that obtained from twitter. All of the data that collected will be select and process, then the researcher do the preprocessing (case folding, normalization features, tokenizing and stemming) and further classified. Naive Bayes classifier is one method that can be used to classify data. The data set are collected as many as 305 data sets, which are divided into training data and data testing. For 210 of training data are divided into classes of positive and negative class, each class still got the same amount of data that is 105. The testing data are used as much as 95. The research will include the classification where data will be categorized into the positive and negative class using naive bayes algorithm. Dataset is tested ten times using K-Fold Cross Validation and the resulting accuracy average of 95.81%. Finally the results that obtained from the Naive Bayes fits for use as the method in sentiment analysis on twitter about the use of public transportation at Semarang.

**Keyword** : Sentiment Analysis; Naive Bayes; Twitter; Public Transport; Cross Validation