

ABSTRACT

Traffic Jam in Special Capital Region (DKI) Jakarta has become a problem since the beginning until now that bothers people to run their activities. To minimize this problem, people are looking for, and exchange information about congestion on social media. One of the social media that is frequently used is by people of the Capital City is Twitter.

Twitter is one of the social media in Indonesia that has users increased from year to year. Modern society people like Jakarta is also utilizing social media to exchange information, including traffic jam. One of a Twitter account that provides information about traffic is @TMCPoldaMetro Twitter account, official accounts belonging Jakarta's Police City. Every day, an average of 200-300 informs this account tweets about congestion in Jakarta from morning till night. With the huge amount of data, then can be analyzed to determine the pattern Data Mining of a certain data.

In this study, We created a classification system congestion in Jakarta with one of the data mining techniques, namely classification. By using Naive Bayes classification method, the data are taken from congestion Twitter account TMCPoldaMetro will be processed and used as a basis for predicting the possible next traffic jam that will happen. Then, from the results of performance testing for five times with different training data in each test, each obtained accuracy results, the first test 56.25%, the second test 63%; The third test 58.89%; The fourth test 61.54%, the highest accuracy of testing with the test contained in the fifth test with 68.66% accuracy.

keyword: Twitter, traffic jam, Data mining, Naive Bayes

ABSTRAK

Kemacetan lalu lintas di Daerah Khusus Ibukota (DKI) Jakarta sudah menjadi masalah sejak dulu sampai sekarang yang mengganggu masyarakat untuk menjalankan aktivitasnya. Untuk meminimalisir masalah ini, masyarakat mencari , dan saling bertukar informasi mengenai kemacetan di media social. Salah satu media penyebar informasi yang sering digunakan masyarakat ibukota adalah melalui social media, yaitu Twitter.

Twitter adalah salah satu media social yang penggunanya di Indonesia semakin meningkat dari tahun ke tahun.. Masyarakat modern seperti yang tinggal di DKI Jakarta pun memanfaatkan media social ini untuk saling bertukar informasi, termasuk masalah kemacetan.Salah satu akun Twitter yang menyediakan informasi seputar lalu lintas adalah akun Twitter @TMCPoldaMetro, akun resmi milik Kepolisian Polda Metro Jaya. Setiap harinya, rata-rata akun ini menginformasikan 200- 300 *tweet* mengenai kondisi kemacetan di DKI Jakarta dari pagi hingga malam hari. Dengan jumlah data yang sangat besar ini, maka dapat dilakukan analisa *Datamining* untuk mengetahui pola dari suatu data tertentu.

Dalam penelitian ini, dibuat suatu sistem klasifikasi kemacetan di DKI Jakarta dengan salah satu teknik *datamining*,yaitu klasifikasi. Dengan menggunakan metode klasifikasi *Naive Bayes*, data kemacetan yang diambil dari akun Twitter @TMCPoldaMetro akan diolah lalu dijadikan dasar untuk memprediksi kemacetan berikutnya yang mungkin akan terjadi. Lalu , dari hasil pengujian kinerja selama lima kali dengan data *training* yang berbeda di setiap pengujiannya, didapat masing-masing hasil akurasi, pengujian pertama 56,25% ;pengujian kedua 63% ; pengujian ketiga 58,89% ; pengujian keempat 61,54%, pengujian dengan akurasi tertinggi terdapat pada pengujian ke lima dengan akurasi 68,66%.

Kata Kunci : Twitter, kemacetan, *Data mining*, *Naive Bayes*