

**IDENTIFICATION OF TRAFFIC FLOW IN YOGYAKARTA USING SUPPORT
VECTOR MACHINE ALGORITHM WITH TWITTER DATASET**

Proposed by

Iqbal Yudiar Syahputra

15/386879/PA/17056

Congestion is a common thing in big cities. Yogyakarta is one of the big cities that is often congested. Because there is a traffic jam, people usually inform the condition via Twitter. But Twitter has the limitation of not being able to identify the location and the time of the traffic situation.

In this research, identification and classification of traffic flow conditions in Yogyakarta will be carried out using the Support Vector Machine algorithm and data obtained from Twitter. The data obtained are 1244 data. Then the irrelevant data is deleted, leaving 1022 data. Then the data is processed to obtain results from the classification and identification process.

The results of the classification process are an accuracy value of 94%, a precision value of 92%, a recall value of 91%, and an f1-score value of 91%. And the results obtained from the identification process in the form of days, time and street names traffic conditions. The identification can produce street names, day. and times when the flow of traffic is happening. And the most frequently jammed roads are Malioboro, Bener, and Magelang roads.

MENGIDENTIFIKASI ARUS LALU LINTAS DI YOGYAKARTA MENGGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE* DENGAN KUMPULAN DATA TWITTER

Oleh:

Iqbal Yudiar Syahputra

15/386879/PA/17056

Kemacetan adalah hal yang sering terjadi di kota - kota besar. Yogyakarta adalah salah satu kota besar yang sering terjadi kemacetan. Karena terjadi macet biasanya orang menginformasikannya melalui Twitter. Namun Twitter memiliki keterbatasan tidak dapat mengidentifikasi lokasi dan waktu dari keadaan lalu lintas.

Pada penelitian ini akan dilakukan pengidentifikasian dan pengklasifikasian kondisi arus lalu lintas di Yogyakarta dengan menggunakan algoritma *Support Vector Machine* dan data didapatkan dari Twitter. Data yang diperoleh berjumlah 1244 data. Kemudian data yang tidak relevan dihapus sehingga menyisakan 1022 data. Kemudian data tersebut diolah untuk didapatkan hasil dari proses klasifikasi dan identifikasi.

Hasil dari proses klasifikasi berupa nilai akurasi 94%, nilai presisi 92%, nilai *recall* 91%, nilai *f1-score* 91%. Dan didapatkan juga hasil dari proses identifikasi berupa hari, waktu dan nama jalan pada kondisi lalu lintas. Dan jalan yang paling sering macet adalah jalan Malioboro, jalan Bener, dan jalan Magelang.