

INTISARI

Metode Ekstraksi Twitter dan Klasifikasi Berbasis Aturan Untuk Memprediksi Kepadatan Lalu Lintas

Oleh

Arief Wibowo

11/326061/SPA/00415

Kemacetan lalu lintas merupakan masalah yang terjadi di kota-kota besar, termasuk Jakarta sebagai Ibukota Negara Republik Indonesia. Berbagai studi terdahulu telah dilakukan untuk mendapatkan alternatif solusi dalam mengatasi kemacetan lalu lintas. Beberapa studi juga telah memanfaatkan data *Twitter* untuk mengidentifikasi kondisi kepadatan lalu lintas di Jakarta.

Penelitian ini mengusulkan suatu metode baru ekstraksi *Twitter* yang dinamakan *xTRoad*, untuk menghasilkan basis data lalu lintas yang akurat dan spesifik pada wilayah tertentu. Metode ekstraksi *xTRoad* bekerja dengan urutan proses tokenisasi, *stop-word removal* serta *stemming*. Metode usulan ekstraksi ini mampu mendeteksi dan menseleksi data *tweet* lalu lintas tertentu yang memiliki kesamaan nama di beberapa wilayah. Hasil ekstraksi teks dengan metode usulan diuji menggunakan rekaman video CCTV dengan nilai akurasi pada proses pengidentifikasian jalan nasional di wilayah DKI Jakarta mencapai 100% dan akurasi untuk proses mengidentifikasi kondisi kepadatan lalu lintas pada jalan nasional sebesar 62,3%.

Penelitian ini juga mengusulkan metode klasifikasi teks berbasis aturan yang dinamakan *Traffic Road-Prediction (TR-P) Advanced*, untuk memprediksi kelas kepadatan lalu lintas. Metode *Traffic Road-Prediction (TR-P) Advanced* bekerja berdasarkan profil yang dibentuk dari pembobotan seluruh atribut pada data hasil ekstraksi *Twitter*. Hasil studi telah menunjukkan bahwa model dapat memprediksi kelas kepadatan lalu lintas dengan nilai akurasi mencapai 91,3%. Kelas kepadatan lalu lintas juga dapat divisualisasi dalam bentuk warna kepadatan pada peta elektronik.

Kata kunci: Ekstraksi *Twitter*, Metode Klasifikasi, Klasifikasi Berbasis Aturan, Prediksi Kepadatan Lalu Lintas

ABSTRACT

A Twitter Extraction Method and Rule Based Classification For Predicting The Traffic Density

by

Arief Wibowo
11/326061/SPA/00415

Road traffic congestion is an ever-prevalent problem in big cities such as Jakarta as the Capital of the Republic of Indonesia. Past studies have been conducted to find out how to solve traffic congestion problems. Several studies have used Twitter text data to identify road traffic density conditions.

This research has proposed a new method for Twitter text extraction called “*xTRoad*” to generate accurate and specific road traffic databases in specific areas. Proposed extraction methods work with tokenization, stop-word removal and stemming. This method is able to detect and select traffic tweet data about roads that have similar names in some areas. The text extraction with the proposed method were tested using CCTV video recording with accuracy value in the process of identifying national road in Jakarta Capital City reaching 100%, and the accuracy to identify road traffic density reaching 62.3%.

This research has also proposed a rule-based text classification method called “*TR-P Advanced*” to determine the road traffic density class. The “*TR-P Advanced*” method works based on profiles that are formed from weighting all attributes on Twitter extraction data. The results of the study have shown that the model has been able to predict the class of traffic density with an accuracy value of 91.3%. Traffic density classes can also be visualized in the form of density colors on electronic maps.

Keywords: Twitter Extraction, Classification Method, Rule Based Classification, Traffic Density Prediction