

INTISARI

Ekstraksi Informasi Kondisi Lalu Lintas dari Media Sosial Twitter

Oleh

Deny Prasetya Dewantara

13/348755/PA/15478

Twitter merupakan media sosial yang memungkinkan penggunanya mengirim berbagai informasi termasuk informasi mengenai kondisi lalu lintas di sekitarnya. Untuk memanfaatkan informasi tersebut perlu dilakukan ekstraksi informasi yang memiliki tugas dasar berupa *Named Entity Recognition* (NER) dan ekstraksi relasi. Tidak semua tweet relevan untuk diekstraksi informasi lalu lintasnya. Oleh karena itu diperlukan pemfilteran tweet yang akan diekstraksi informasi.

Pada penelitian ini dilakukan ekstraksi informasi kondisi lalu lintas dengan melakukan pemfilteran tweet untuk mendapat tweet yang relevan, melakukan NER untuk mengenali entitas lokasi, waktu, dan kondisi lalu lintas, serta melakukan ekstraksi relasi untuk menemukan relasi kondisi, detail, dan arah pada entitas yang telah dikenali. Ketiga fungsi tersebut diimplementasikan menggunakan algoritma *machine learning* yaitu *Support Vector Machine* (SVM).

Hasil dari penelitian ini menunjukkan bahwa sistem ekstraksi informasi yang dibuat menggunakan algoritma SVM memiliki nilai *f-measure* pada komponen filter tweet sebesar 0.93, pada komponen NER sebesar 0.87, dan pada komponen ekstraksi relasi sebesar 0,83

Kata Kunci : Ekstraksi Informasi, Twitter, Lalu Lintas, NER, Relation Extraction, Support Vector Machine

ABSTRACT

Traffic Condition Information Extraction from Twitter

by

Deny Prasetya Dewantara

13/348755/PA/15478

Twitter is a social media that allows users to send various information including information about traffic conditions. To utilize the information, extraction information which has the basic task of Named Entity Recognition (NER) and relation extraction is needed. Not all tweets are relevant for extracting traffic information. Therefore it is necessary to filter tweet before the information extraction process.

In this study traffic condition information extraction is done by filtering tweets to get relevant tweets, performing NER to recognize location, times, and traffic conditions entities also extracting relations to find condition, detail, and direction relation between recognized entities. Those three tasks are implemented using Support Vector Machine (SVM) algorithm.

The result of this research shows that the information extraction system made using SVM algorithm has f-measure value of 0.93 on filtering tweet component, 0.87 on NER component, and 0.83 on relation extraction component.

Keywords: Information Extraction, Twitter, Traffic, NER, Relation Extraction, Support Vector Machine