

INTISARI

KLASIFIKASI JENIS EMOSI PADA *TWEET* UNTUK DETEKSI BUNUH DIRI MENGUNAKAN *SUPPORT VECTOR MACHINE* DAN *MULTINOMIAL NAIVE BAYES*

Oleh:

Maria Gabriel Okta Yuliana

14/364185/PA/15937

Salah satu bentuk dari berkomunikasi adalah mengungkapkan perasaan atau emosi kepada orang lain. Banyak media yang membantu kita dalam berkomunikasi, salah satunya media sosial Twitter. *Tweet* sebagai representasi pemikiran pengguna dapat dimanfaatkan untuk menganalisis kondisi emosi serta tendensi pengguna terhadap perilaku bunuh diri.

Pada penelitian ini, telah dikumpulkan 7500 *tweet* dari pengguna yang sudah meninggal dengan penyebab bunuh diri dan kecelakaan. Proses klasifikasi emosi dan bunuh diri dilakukan menggunakan algoritma *Support Vector Machine* (SVM) dan *Multinomial Naive Bayes* (MNB). *Term Frequency-Inverted Document Frequency* (TF-IDF) dan *Total Emotions Intensity* (TEI) menjadi fitur pada klasifikasi emosi, sedangkan *output* dari klasifikasi emosi dan bigram menjadi fitur pada klasifikasi bunuh diri.

Penambahan fitur TEI pada klasifikasi emosi dapat meningkatkan akurasi pada klasifikasi emosi dengan menggunakan algoritma SVM metode One-vs-One, dengan hasil 92,32%. Hal ini dikarenakan skor TEI mengindikasikan emosi apa yang paling representatif pada dokumen. Pada klasifikasi bunuh diri dihasilkan akurasi terbaik sebesar 55,05% dengan menggunakan algoritma SVM. Pelaku bunuh diri tidak selalu menunjukkan kecenderungan terhadap emosi tertentu, namun cenderung kondisi emosi yang sering berubah.

Kata Kunci: klasifikasi emosi, klasifikasi bunuh diri, *support vector machine*, *naive bayes*

ABSTRACT

EMOTION DETECTION IN TWEET FOR EARLY IDENTIFICATION OF SUICIDAL BEHAVIOR USING SUPPORT VECTOR MACHINE AND MULTINOMIAL NAÏVE BAYES

By

Maria Gabriel Okta Yuliana

14/364185/PA/15937

Communication is a human need. One of the ways is expressing emotion towards others. There are many media that helps us to communicate and twitter is one of them. A tweet is representing user's idea, emotion or opinion. These tweets can be used to analyze something about the user.

Support Vector Machine (SVM) and Multinomial Naive Bayes (MNB) algorithm will be used in this classification. The data obtained from dead user, caused by suicide and accident. There are 7500 tweets in total. Term frequency-Inverted Document Frequency (TF-IDF) and Total Emotion Intensity (TEI) will be the features for emotion classification, meanwhile bigram will be the feature for suicide classification.

TEI improve the result of emotion classification from 77,86% (Chirawichitchai, 2014) to 92,32% by using SVM One-vs-One method. TEI calculate intensity for each emotion, hence the highest score will indicate which emotion represent the tweet. The best result for suicide classification is 55,05% by using SVM algorithm. The data shows that the suicidal user does not always show a tendency towards a particular emotion, but tends to a changing emotional state quickly.

Keywords: emotion classification, suicide classification, support vector machine, multinomial naive bayes