

ABSTRAK

Judul : Analisis Sentimen Pelecehan Seksual Di Media Sosial X (Twitter) Menggunakan Metode Klasifikasi Naïve Bayes Classifier, Support Vector Machine, Random Forest
Nama : Keffinsyah Kahffi
Program Studi : Teknik Informatika

Pengguna aktif media sosial terus meningkat, dengan X (Twitter) menjadi salah satu platform yang populer. Media sosial, termasuk X (Twitter), menjadi tempat bagi pengguna untuk mengekspresikan pendapat, termasuk mengenai isu pelecehan seksual. Penelitian ini bertujuan untuk menganalisis sentimen tweet pengguna X (Twitter) terkait pelecehan seksual menggunakan metode klasifikasi seperti Naïve Bayes Classifier, Support Vector Machine (SVM), dan Random Forest. Pengumpulan data tweet menggunakan metode crawling menggunakan API yang disediakan oleh X (Twitter) dengan kata kunci yang berhubungan dengan pelecehan seksual. Data tweet yang digunakan sebanyak 11.284 tweet, yang selanjutnya dibersihkan pada tahap preprocessing dan diberi label dengan hasil Vader 78,7% Negatif, 18,2% Positif, dan 3,1% Netral, InSet menghasilkan 65,1% Negatif, 30,5% Positif, dan 4,4% Netral, Pelabelan yang dilakukan oleh pakar 73,4% Negatif, 22,3% Positif, dan 4,3% Netral terkait pelecehan seksual data yang sudah diberi label sebagai data untuk pengujian pada algoritma yang digunakan. Dari hasil pengujian performa pada algoritma yang diuji dapat disimpulkan bahwa algoritma Random Forest memiliki hasil performa yang lebih tinggi dibandingkan dengan Naive Bayes Classifier dan Support Vector Machine dengan Hasil akurasi: 82.41%, presisi: 81.31%, recall: 82.41%, dan f1-scores: 81.07%. Berdasarkan hasil dari Identifikasi dari kata " 'korban,' 'laku,' dan 'seksual'" menunjukkan indikasi terhadap pelecehan seksual. Penelitian ini Juga memberikan Word cloud untuk memberikan gambaran visual tentang kata-kata yang paling sering muncul dalam tweet. Kata-kata yang dominan dapat memberikan indikasi tentang topik atau isu yang paling mendominasi terkait pelecehan seksual.

Kata kunci : *Analisis Sentimen, Twitter, Naive Bayes Classifier, Random Forest, Pelecehan Seksual. Wordcloud*

ABSTRACT

Title : Analysis of Sexual Harassment Sentiment on Social Media X (Twitter) Using the Naïve Bayes Classifier, Support Vector Machine, Random Forest Classification Method
Name : Keffinsyah Kahffi
Study Program : Informatics Engineering

Active social media users continue to increase, with X (Twitter) becoming one of the popular platforms. Social media, including X (Twitter), is a place for users to express opinions, including regarding the issue of sexual harassment. This research aims to analyze the sentiment of user X (Twitter) tweets regarding sexual harassment using classification methods such as Naïve Bayes Classifier, Support Vector Machine (SVM), and Random Forest. Tweet data collection uses a crawling method using the API provided by X (Twitter) with keywords related to sexual harassment. The tweet data used was 11,284 tweets, which were then cleaned at the preprocessing stage and labeled with Vader results of 78.7% Negative, 18.2% Positive, and 3.1% Neutral, InSet produced 65.1% Negative, 30.5% % Positive, and 4.4% Neutral, Labeling carried out by experts 73.4% Negative, 22.3% Positive, and 4.3% Neutral regarding sexual harassment data that has been labeled as data for testing the algorithm used. From the results of performance testing on the tested algorithm, it can be concluded that the Random Forest algorithm has higher performance results compared to the Naive Bayes Classifier and Support Vector Machine with accuracy results: 82.41%, precision: 81.31%, recall: 82.41%, and f1-scores: 81.07%. Based on the results of the Identification of the words "'victim,' 'perpetrator,' and 'sexual'" show indications of sexual harassment. This research also provides a word cloud to provide a visual representation of the words that appear most frequently in tweets. Dominant words can provide an indication of the most dominant topics or issues related to sexual harassment

Keywords : *Sentiment Analysis, Twitter, Naive Bayes Classifier, Random Forest, Sexual Harrasment*