

DETEKSI RADIKALISME DALAM TWEET BERBAHASA INDONESIA DENGAN MENGGUNAKAN METODE SVM

Muhammad Rizky Nirwan¹, Dade Nurjanah², Hani Nurahmi³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung
¹mrinran@students.telkomuniversity.ac.id, ²dadenurjanah@telkomuniversity.ac.id,
³haninurahmi@telkomuniversity.ac.id,

Abstrak

Dengan mudahnya mengakses Internet, maka semakin mudah juga untuk mengakses media sosial. Dengan kemudahan tersebut kebebasan untuk membagikan hal apapun didalam Internet, memungkinkan penggunanya untuk menyebarkan konten radikalisme di sosial media. Penelitian ini Berfokus terhadap bagaimana mendeteksi konten radikalisme tersebut. Studi kasus yang digunakan dalam penelitian ini adalah media sosial Twitter, Twitter merupakan sosial media dengan pengguna yang banyak. Metode yang digunakan dalam penelitian ini adalah Support Vector Machine linier. Dilakukan juga pembobotan kata dengan menggunakan metode TF/IDF. Dataset diperoleh dari proses *crawling tweet*. Pra-pemrosesan dilakukan dengan menggunakan metode cleansing/filtering stop word untuk membuang kata kata yang kurang memiliki makna. Berdasarkan hasil pengujian yang dilakukan dengan menggunakan K - *Cross Validation* dengan nilai K = 10 untuk pembagian datatest sebanyak 10% adalah 70%, untuk K = 5 untuk pembagian datatest sebanyak 20% adalah 69%, untuk K = 4 untuk pembagian dataset sebanyak 25% adalah 67% dan untuk K = 2 untuk dataset sebanyak 50% adalah 71%. Hasilnya pembagian data 50% dengan K= 2 memiliki nilai yang paling tinggi.

Kata kunci : Radikalisme, *support vector machine*, Twitter

Abstract

With easy access to the Internet, it is also easier to access social media. With this convenience the freedom to share anything on the Internet, allows users to spread Radicalism content on social media. This research focuses on how to detect the content of radicalism. The case study used in this research is social media Twitter, Twitter is a social media with many users. The method used in this study is linear Support Vector Machine. Word weighting was also performed using the TF / IDF method. The dataset is obtained from the tweet crawling process. Pre-processing is done by using the stop word cleansing / filtering method to get rid of words that lack meaning. Based on the results of tests conducted using K-Cross Validation with a value of K = 10 for the distribution of datatest of 10% is 70%, for K = 5 for the distribution of datatest of 20% is 69%, for K = 4 for the distribution of datasets by 25% is 67% and for K = 2 for the dataset 50% is 71%. The result is that 50% data sharing with K = 2 has the highest value.

Keywords: Radicalism, *support vector machine*, Twitter

1. Pendahuluan

1.1 Latar Belakang

Di era digitalisasi ini, terdapat banyak manfaat dalam teknologi informasi, tetapi dalam kemudahan mendapatkan informasi tersebut juga membawa sisi negatif, contohnya tindakan *Cyber terrorism*. *Cyber terrorism* adalah bentuk terorisme yang menggunakan teknologi informasi, jaringan komputer untuk mengancam keselamatan orang lain atau suatu negara[1]. Karena pentingnya jaringan internet, beberapa ahli berpendapat bahwa *Cyber terrorism* lebih berbahaya daripada teroris tradisional[1]. Definisi dari konten radikal sendiri sangat berbeda di berbagai negara. Di Indonesia, radikal lebih identik dan sering dikaitkan dengan Suku Ras Agama dan Antar golongan (SARA)[1]. Menurut Badan Nasional Penanggulangan Terorisme (BNPT), konten radikal mendorong, memprovokasi orang melakukan kekerasan atas nama agama, dan melakukan pemboman bunuh diri dari kelompok yang tidak termasuk kelompok mereka[1]. Dalam beberapa tahun terakhir, radikalisme telah bermigrasi ke lingkungan virtual internet dimana banyak organisasi teroris yang menggunakan media sosial untuk mempromosikan ideology dan propaganda mereka, serta merekrut individu untuk tujuan mereka[2]. Dengan adanya media sosial dan komunikasi terenkripsi, tidak hanya radikalisasi tetapi juga perencanaan operasional dapat dengan mudah terjadi di internet [2].