

## Deteksi *Cyberbullying* pada Twitter Menggunakan Metode Klasifikasi *Support Vector Machine*

Ni Luh Putu Mawar S.P.W.<sup>1</sup>, Dade Nurjanah, S.T., M.T., Ph.D.<sup>2</sup>, Hani Nurrahmi, S.Kom., M.Kom<sup>3</sup>

<sup>1,2,3</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>1</sup>mawarsilveria@students.telkomuniversity.ac.id, <sup>2</sup>dadenurjanah@telkomuniversity.ac.id,

<sup>3</sup>haninurrahmi@telkomuniversity.ac.id

---

### Abstrak

Tindakan *bullying* yang dilakukan untuk menyerang seseorang atau sekelompok orang dengan agresif secara berulang-ulang kerap terjadi. Dengan adanya perkembangan teknologi internet masyarakat sangat dipermudah untuk melakukan tindakan negatif *bullying* untuk menyerang seseorang atau kelompok orang yang dapat menyebabkan korban tersakiti, tindakan ini disebut dengan *cyberbullying*. Twitter merupakan media sosial yang dapat digunakan masyarakat untuk berbagi informasi berupa tulisan dan dapat disalahgunakan untuk melakukan tindakan *cyberbullying* dengan mengirimkan pesan (*tweet*) terhadap korban yang dituju. Pada tugas akhir ini telah dibangun sebuah sistem untuk mendeteksi tindakan *cyberbullying* pada Twitter. Sistem yang dibangun menggunakan metode *Support Vector Machine* untuk mengklasifikasi apakah *tweet* yang dibagikan termasuk *cyberbullying* atau bukan. Selain itu, tugas akhir ini juga menggunakan ekstraksi fitur *Term Frequency-Inverse Document Frequency* (TF-IDF) dan N-gram untuk data yang sudah melalui tahap *pre-processing*. Dalam pengambilan data, penulis melakukan *crawling tweet* berdasarkan *keyword* 'jelek', 'bodoh', 'goblok', 'brengeks', 'bangsat', 'memalukan', 'laknat', 'bacot' dan 'pelacur'. Hasil penelitian terbaik menghasilkan akurasi 76.2%, *precision* 73.2%, *Recall* 78.2% dan *F1-Score* 75.6% yang dihasilkan oleh kernel RBF dengan jumlah  $n=1$ .

**Kata Kunci:** *cyberbullying*, twitter, *support vector machine*, n-gram, TF-IDF.

---

### Abstract

Bullying is when someone or a group of individuals is continuously attacked. Because of the advancement of the internet, it has become very easy for society to engage in harmful acts of bullying by attacking a person or group of people who can hurt the victim, this is known as cyberbullying. Twitter is a social media platform that may be used by the society to share information and can also be used to perpetrate cyberbullying actions by sending messages (tweets) that addressed to the victims. This final project was developing a system to detect cyberbullying on Twitter. The system uses the Support Vector Machine method to classify whether the tweets that are shared include cyberbullying or not. In addition, this research also uses Term Frequency-Inverse Document Frequency (TF-IDF) and N-gram feature extraction for data that has gone through the pre-processing stage. In collecting data, the author crawled tweets based on the keywords 'jelek', 'bodoh', 'goblok', 'brengeks', 'bangsat', 'memalukan', 'laknat', 'bacot' and 'pelacur'. The best performance results of the research is 76.2% accuracy, 73.2% precision, 78.2% recall and 75.6% F1-Score generated by the RBF kernel with a total of  $n=1$ .

**Keywords:** Cyberbullying, Twitter, Support Vector Machine, N-gram, TF-IDF