

Klasifikasi Komentar *Toxic* di Media Sosial Menggunakan *Support Vector Machine* dan *Chi Square* sebagai *Feature Selection*

Nadhia Salsabila¹, Danang Triantoro Murdiansyah², Kemas M. Lhaksana³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹nadhiasalsabila@student.telkomuniversity.ac.id, ²danangtri@telkomuniversity.ac.id,

³kemasmuslim@telkomuniversity.ac.id

Abstrak

Penggunaan media sosial di masyarakat terus meningkat dari waktu ke waktu. Kemudahan akses dan keakraban media sosial memudahkan pengguna yang tidak bertanggung jawab untuk melakukan hal-hal yang tidak etis seperti menyebarkan kebencian, pencemaran nama baik, radikalisme, pornografi, dll. Meskipun ada peraturan yang mengatur semua aktivitas di media sosial, namun tetap saja tidak bekerja secara efektif karena ketidakmungkinan mengklasifikasikan komentar secara manual. Oleh karena itu, kami melakukan penelitian ini untuk mengklasifikasikan komentar ke dalam kategori toksisitasnya menggunakan metode pembelajaran mesin untuk tujuan kenyamanan penggunaan media sosial. Metode yang kami gunakan dalam penelitian ini adalah SVM dengan TF-IDF sebagai ekstraksi fitur dan Chi Square sebagai seleksi fitur. Kami juga melakukan beberapa skenario eksplorasi, termasuk mengimplementasikan kernel SVM dan tahapan preprocessing untuk mengetahui performa terbaik dari model tersebut. Performa terbaik diperoleh dengan menggunakan model SVM dengan kernel linier, tanpa mengimplementasikan Chi Square, dan menggunakan *stemming* dan *stopwords removal* dengan *F1-Score* sebesar 76,57%.

Kata Kunci: klasifikasi teks, komentar *toxic*, media sosial, *support vector machine*