

Comparison of Naive Bayes and SVM Methods for Identifying Anxiety Based on Social Media

Endri Rizki Nugraha¹, Warih Maharani²

Fakultas Informatika, Universitas Telkom, Bandung

exdxix@student.telkomuniversity.ac.id,

wmaharani@telkomuniversity.ac.id

Abstrak

Penelitian ini bertujuan untuk mendeteksi pola kecemasan dari unggahan di media sosial menggunakan algoritma Naive Bayes (NB) dan Support Vector Machine (SVM). Cuitan dikumpulkan menggunakan teknik Data Crawling, kemudian diberi label berdasarkan kuesioner Depression Anxiety Stress Scale (DASS-42) serta diseimbangkan menggunakan Random Oversampler untuk menangani ketidakseimbangan dataset. NB dan SVM dipilih karena efektivitasnya dalam klasifikasi sentimen teks. Studi ini mengintegrasikan fitur tekstual yang diperoleh dari metode Term Frequency-Inverse Document Frequency (TF-IDF) dan Bag of Words (BoW). Perbandingan dilakukan untuk mengidentifikasi keunggulan dan keterbatasan masing-masing metode dalam menangani data sentimen teks. Penelitian ini menganalisis data sentimen dengan menghitung akurasi, recall, dan F1-score guna menentukan hasil kinerja paling optimal. Hasil penelitian menunjukkan bahwa SVM dengan ekstraksi fitur TF-IDF mencapai akurasi tertinggi sebesar 72% dan rata-rata F1-Score sebesar 61%, sedangkan NB dengan BoW mencapai akurasi sebesar 56% dan rata-rata F1-Score sebesar 49%. Temuan ini menyoroti efektivitas kombinasi SVM dan fitur TF-IDF, yang meningkatkan kinerja model, dengan SVM menghasilkan hasil terbaik dalam mengidentifikasi kecemasan dari data media sosial.

Kata Kunci : Anxiety, Naive Bayes, SVM, DASS-42, Sentiment Analysis, Social Media..
