

Analisis Pengaruh Seleksi Fitur *Information Gain* dan *Mutual Information* pada Klasifikasi Sentimen Ulasan Film Menggunakan *Support Vector Machine*

Hanif Salafi¹, Adiwijaya², Untari Novia Wisesty³

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

¹hanifsalafi@student.telkomuniversity.ac.id, ²adiwijaya@telkomuniversity.ac.id,

³untarinw@telkomuniversity.ac.id

Abstrak

Analisis sentimen adalah teknik identifikasi emosi yang diekspresikan dalam bentuk teks. Tujuan analisis sentimen adalah untuk menentukan pendapat positif atau negatif dalam suatu kalimat atau dokumen. Salah satu cara untuk menentukan positif atau negatif dari suatu pendapat dapat dilakukan dengan klasifikasi teks. Namun, klasifikasi teks mempunyai masalah yaitu banyaknya atribut yang digunakan pada sebuah *dataset*. Seleksi fitur memiliki peran penting untuk mengurangi atribut yang kurang relevan dari suatu *dataset*. Metode pemilihan fitur seperti *Information Gain* (IG) dan *Mutual Information* (MI) adalah metode yang umum digunakan pada klasifikasi teks. Pada penelitian ini domain yang digunakan adalah ulasan film. Masing-masing dari kedua metode pemilihan fitur tersebut dijadikan sebagai seleksi fitur untuk membantu meningkatkan performansi algoritma klasifikasi *Support Vector Machine* (SVM). Hasil dari pengklasifikasian SVM dengan kedua metode seleksi fitur tersebut menunjukkan bahwa penggunaan seleksi fitur dengan SVM dapat meningkatkan nilai akurasi, dimana saat tidak menggunakan seleksi fitur hanya mendapatkan akurasi sebesar 88.75%, sedangkan jika ditambahkan seleksi fitur nilai akurasi meningkat menjadi 89.05%. Kemudian ketika kedua seleksi fitur tersebut dibandingkan, *Information Gain* memberikan hasil yang sama dengan *Mutual Information* dengan nilai akurasi tertinggi sebesar 89.05%. Dengan demikian, penggunaan *Information Gain* sebagai seleksi fitur pada algoritma klasifikasi SVM memberikan hasil yang sama dengan *Mutual Information*.

Kata kunci: Analisis Sentimen, Klasifikasi, Ulasan Film, *Information Gain*, *Mutual Information*, *Support Vector Machine*
