

Abstrak

Pada saat ini, industri perfilman dunia berkembang cukup pesat dan digemari masyarakat dunia. Untuk mengetahui menarik atau tidaknya suatu film kita perlu melihat ulasan dari penonton sebelumnya pada film tersebut. Beberapa ulasan mungkin terlihat jelas untuk dikategorikan kedalam ulasan positif atau negatif, tetapi masih terdapat ulasan yang belum jelas untuk dikategorikan. Salah satu teknik otomatisasi untuk membantu pengklasifikasian ulasan yaitu dengan analisis sentimen. *K-Nearest Neighbor* (KNN) adalah metode pengklasifikasian yang cukup sederhana. Namun, *K-Nearest Neighbor* memiliki masalah terkait dimensi fitur yang tinggi. Untuk mengatasi masalah ini, penelitian kali ini melakukan analisis sentimen terhadap ulasan film dengan menggunakan metode *K-Nearest Neighbor* dan seleksi fitur *Gini Index* terhadap dataset *Internet Movie Database* (IMDB) yang terbagi menjadi sentimen positif dan sentimen negatif. Seleksi fitur menggunakan *Gini Index* dapat mengurangi dimensi fitur yang tinggi untuk meningkatkan performansi *K-Nearest Neighbor*. Dengan memperhatikan nilai k dan *threshold Gini Index* diperoleh rata-rata akurasi penggunaan metode *K-Nearest Neighbor* tanpa seleksi fitur sebesar 79,6% dan rata-rata akurasi terbaik penggunaan metode *K-Nearest Neighbor* dengan seleksi fitur *Gini Index* sebesar 82,1%. Ini membuktikan seleksi fitur *Gini Index* dapat meningkatkan performansi dari metode *K-Nearest Neighbor*.

Kata kunci : ulasan film, analisis sentimen, *K-Nearest Neighbor*, *Gini Index*, seleksi fitur