

ABSTRAK

Perkembangan teknologi internet khususnya dalam bidang *perfilman* memberikan sarana terbuka bagi masyarakat dalam mengekspresikan pendapat dan emosional. Salah satu pendapat yang seringkali masyarakat keluarkan yaitu berupa penilaian sebuah *film* pada *platform* tertentu seperti IMDB. Ulasan yang dikeluarkan tentunya mengandung emosional yang dibawakan oleh masyarakat itu sendiri, baik emosional positif maupun negatif yang dinamakan sentimen. Sentimen atau opini masyarakat ini perlu dianalisis untuk mengklasifikasikan opini sesuai dengan kelasnya sehingga kecenderungannya terhadap suatu objek dapat diketahui. Adapun metode yang digunakan dalam penelitian ini yaitu metode *data mining* dengan *Knowledge Discovery in Database (KDD)*. Tujuan dari penelitian ini yaitu analisa sentimen IMDB *film review* oleh masyarakat menggunakan algoritma *Support Vector Machine* dan seleksi *Feature importance* untuk memperoleh pola dan hasil yang terbaik. Dengan pengujian validasi akurasi data menggunakan metode split data sederhana dan *k-fold cross validation* yang menghasilkan akurasi sebesar 91.942% dan 87.699%. Lalu Kemudian dilakukan evaluasi menggunakan *confusion matrix* dengan penetapan *max feature* sebesar 10000 untuk memeriksa nilai ketepatan prediksi yang dilakukan oleh model yaitu diperoleh akurasi sebesar 88.033%. Dalam hal ini dapat dibuktikan bahwa kemampuan model dalam melakukan klasifikasi dinilai cukup baik.

Kata kunci : *Data mining, KDD, Feature importance, SVM, Confusion matrix*