

ABSTRAK

Meningkatnya kualitas film buatan sineas dalam negeri mendorong masyarakat untuk menonton film karya anak bangsa. Tentu banyak faktor pendukung dibalik kesuksesan sebuah film. Sutradara yang berkualitas serta kemampuan aktor dalam memerankan sebuah karakter dapat menjadi daya tarik lebih dari suatu film agar “laku” dipasaran. Adapun hal lain yang penting untuk diperhatikan adalah seberapa sering sebuah film dibicarakan di media sosial. Semakin banyak diperbincangkan maka akan membuat masyarakat merasa ingin tahu lebih mengenai film tersebut, salah satu media sosial yang dapat dijadikan patokan sebuah film sedang ramai diperbincangkan adalah Twitter. Melalui Twitter dapat diteliti apakah sebuah film akan laku dipasaran atau tidak. Pada penelitian ini data yang digunakan diambil dari Twitter dengan mengedepankan dua sisi yaitu pada sisi *user* dan sisi *tweet*. Pada sisi *user* adapun fitur-fitur yang digunakan yaitu *follower official*, *follower production house*, *follower produser*, *follower sutradara*, dan *follower* dari 3 artis pemeran utama film, untuk sisi *tweet* fitur-fitur yang digunakan adalah data *tweet*, *retweet*, *like*, jumlah *tweet* yang berisi URL, dan jumlah *tweet* yang berisi sentimen positif maupun negatif. Tujuan dari penelitian ini adalah menerapkan teknik *pre-processing* untuk membentuk data masukan yang lebih baik kedalam algoritma *Random Forest*. Hasil pengujian menggunakan algoritma *Random Forest* didapatkan nilai akurasi sebesar 88%.

Kata Kunci : *Twitter, Pre-processing, Random Forest*