

Abstrak

Industri perfilman merupakan salah satu dampak dari perkembangan teknologi yang sangat pesat. Hal itu menyebabkan industri perfilman mengalami peningkatan setiap tahunnya. Selain itu, perkembangan teknologi juga berpengaruh pada masyarakat untuk mempermudah mengakses berbagai film dari berbagai website. Dengan banyak pilihan film, masyarakat perlu mengetahui kualitas dari berbagai film dengan cara mengetahui ulasan film tersebut dari orang lain. Namun, banyaknya ulasan dari penonton terhadap suatu film membuat masyarakat kesulitan dalam mengategorikan film yang bagus dan film yang tidak bagus. Solusi dari permasalahan tersebut adalah dengan melakukan analisis sentimen terhadap ulasan film. Pada penelitian ini, metode klasifikasi yang digunakan adalah Modified Balanced Random Forest. Metode ini dipilih karena dapat mengatasi data yang imbalanced dan dapat meningkatkan akurasi serta mengurangi kompleksitas waktu. Pada penelitian ini juga, Word2Vec digunakan sebagai ekstraksi fitur. Ekstraksi fitur ini dipilih karena pada penelitian sebelumnya menjelaskan bahwa Word2Vec memiliki kelebihan yaitu dapat menunjukkan kesamaan kontekstual dua kata pada vektor yang dihasilkan. Model terbaik yang dihasilkan dari penelitian ini yaitu, model yang dibangun tanpa menggunakan stemming pada tahap preprocessing, menggunakan 300 dimensi pada Word2Vec, dan menggunakan metode klasifikasi Modified Balanced Random Forest yang menghasilkan f1-score sebesar 84.15%.

Kata kunci : Analisis Sentimen, Ulasan Film, Modified Balanced Random Forest, Word2Vec