

ABSTRAK

Film merupakan suatu media komunikasi massa yang digunakan sebagai sarana hiburan bagi masyarakat. Film cukup efektif dalam menyampaikan suatu informasi. Dewasa ini, banyak sekali *website* yang menyediakan informasi film-film terbaru. Informasi yang bisa didapatkan salah satunya adalah *genre* film. Satu film bisa memiliki satu atau lebih *genre*. Namun untuk menentukan *genre* sebuah film, orang harus membaca sinopsisnya terlebih dahulu atau menonton film tersebut sehingga membutuhkan waktu yang cukup lama. Tugas akhir ini memberikan solusi pada permasalahan tersebut dengan melakukan klasifikasi *multi-genre* film dari sinopsis film menggunakan *naive bayesian multi-label classifier*. Tahapan yang dilakukan pada penelitian ini yaitu pengambilan data menggunakan teknik *undersampling*, *pre-processing*, *feature selection*, *classifier building*, dan *classifier evaluation*. Pada tahap *pre-processing* yang dilakukan adalah proses *stop words removal*, *word segmentation* dan *stemming* terhadap data yang telah didapatkan di *imdb.com/interfaces*. Pada tahap *feature selection* dan *classifier building* yang dilakukan masing-masing adalah *chi square test* dan penerapan algoritma *naive bayesian multi-label*. Algoritma *naive bayesian multi-label* merupakan pengembangan *multinomial naive bayes* yang setiap kelasnya merupakan kombinasi dari kelas yang terbentuk dari data. Hasil pengujian menunjukkan bahwa untuk klasifikasi *genre* film berdasarkan synopsis yang paling efektif menggunakan *Naive Bayesian Muti-Label (NBML) classifier* dengan *micro average f1-measure* sebesar 87%, dibandingkan dengan beberapa *classifier* yang populer lainnya (SVM classifier sebesar 65,15%, Decision Tree sebesar 50,53%, dan K-NN sebesar 59,52%).

Kata Kunci: klasifikasi, *multi-label*, film, *classifier*, *naive bayesian*