



INTISARI

KLASIFIKASI GENRE FILM BERDASARKAN SINOPSIS MENGGUNAKAN WORD EMBEDDING

Oleh

Elsa Ramadhani
20/459263/PA/19924

Layanan *streaming* film telah menjadi salah satu layanan yang semakin populer belakangan ini. Melalui layanan tersebut, penggunanya dapat menonton film melalui beragam *device* serta menemukan rekomendasi film lainnya yang tersedia. Terkait dengan pembuatan sistem rekomendasi film, telah dilakukan beberapa penelitian klasifikasi genre film secara otomatis menggunakan teks sinopsis. Teks sinopsis dari suatu film dianggap dapat menyampaikan genre yang dimiliki oleh film tersebut.

Penelitian ini melakukan percobaan klasifikasi genre film berdasarkan sinopsis dengan pendekatan *word embedding*. Terdapat dua metode yang dicobakan, yakni metode kombinasi antara model Word2Vec dengan model LSTM serta metode *fine-tuning* BERT. Pengujian pada metode Word2Vec + LSTM memberikan hasil terbaik dengan nilai *f1-score* 0,56. Sementara itu, metode BERT mendapatkan hasil terbaik dari seluruh percobaan yang dilakukan dengan nilai *f1-score* 0,64. Berdasarkan hasil penelitian ini, penggunaan model BERT dinilai berpotensi untuk memberikan hasil yang lebih baik dalam kasus klasifikasi genre film menggunakan sinopsis.

Kata-kata kunci: Klasifikasi Genre Film, Sinopsis, *Word Embedding*, Word2Vec, LSTM, BERT

ABSTRACT

MOVIE GENRE CLASSIFICATION BASED ON SYNOPSIS USING WORD EMBEDDING

By

Elsa Ramadhani
20/459263/PA/19924

Movie streaming service has become one of the most popular services in recent times. Using these services, users can watch movies through various devices and find recommendations for other available movies. In regard to movie recommendation system, several studies have been conducted on automatic movie genre classification using synopsis text. The synopsis text of a movie is considered to be able to convey the genre of the movie.

This study conducted an experiment to classify movie genres using word embedding approach. Two methods were used in this study, the first one is a combination method between Word2Vec model and LSTM model, and the second one is BERT fine-tuning method. The Word2Vec + LSTM method was able to achieve its best result with f1-score of 0,56. Meanwhile, the BERT method was able to achieve the best result from all the experiments done in this study with f1-score of 0,64. Based on the result of this study, the BERT method is considered to have the potential to provide a better result in terms of movie genre classification using synopsis.

Keywords: Movie Genre Classification, Synopsis, Word Embedding, Word2Vec, LSTM, BERT