ABSTRAK

Joshua Yordana:

Skripsi

Prediksi Genre dan Rating Penonton terhadap Sinopsis Film Menggunakan Bidirectional Encoder Representations from Transformers (BERT)

Proses penulisan dan pengembangan sinopsis film membutuhkan evaluasi yang mendalam agar dapat sesuai dengan minat pasar. Penelitian ini bertujuan untuk membandingkan akurasi dua model kecerdasan buatan, yaitu IndoBERT dan Indonesian RoBERTa, dalam memprediksi genre dan rating penonton terhadap sinopsis film. Model ini diimplementasikan dalam sebuah aplikasi web yang dikembangkan menggunakan bahasa pemrograman Python dan framework Flask.

Dataset yang digunakan dalam penelitian ini terdiri dari 1.824 sinopsis film berbahasa Indonesia yang diambil dari API The Movie Database (TMDb) dan Kaggle. Setiap film memiliki atribut seperti judul, rating, genre, dan sinopsis. Data diolah menggunakan Auto Tokenizer dari *huggingface*.

Hasil pengujian menunjukkan bahwa kedua model mampu memprediksi genre dan rating dengan akurasi yang tinggi. Untuk prediksi rating, model IndoBERT menunjukkan sedikit keunggulan dengan rata-rata perbedaan sebesar 0.4 dari rating asli. Untuk prediksi genre, kedua model menunjukkan hasil yang hampir identik dengan genre sebenarnya.

Aplikasi ini diuji oleh produser, sutradara, dan penulis naskah film indie melalui survei yang menunjukkan bahwa mereka menemukan aplikasi ini bermanfaat dan akurat dalam membantu evaluasi sinopsis film. Dengan nilai rata-rata 4.333 untuk akurasi prediksi genre dan 4.5 untuk prediksi rating menggunakan model IndoBERT, aplikasi ini terbukti dapat memberikan rekomendasi yang relevan dan berguna bagi para profesional di industri film.

Kata kunci:
IndoBERT, Indonesian RoBERTa, prediksi genre, prediksi rating, sinopsis film, aplikasi web

ABSTRACT

Joshua Yordana:

Undergraduate Thesis

Prediction of Genre and Audience Rating for Movie Synopsis Using Bidirectional Encoder Representations from Transformers (BERT)

The process of writing and developing movie synopsis requires thorough evaluation to align with market interests. This study aims to compare the accuracy of two artificial intelligence models, namely IndoBERT and Indonesian RoBERTa, in predicting the genre and audience rating of movie synopsis. These models are implemented in a web application developed using the Python programming language and the Flask framework.

The dataset used in this study consists of 1,824 Indonesian movie synopsis sourced from The Movie Database (TMDb) API and Kaggle. Each movie has attributes such as title, rating, genre, and synopsis. The data is processed using the Auto Tokenizer from Huggingface.

The testing results show that both models are capable of predicting genre and rating with high accuracy. For rating prediction, the IndoBERT model shows a slight advantage with an average difference of 0.4 from the actual rating. For genre prediction, both models demonstrate results that are almost identical to the actual genres.

The application was tested by producers, directors, and indie film script writers through a survey, which showed that they found the application useful and accurate in helping to evaluate movie synopsis. With an average score of 4.333 for genre prediction accuracy and 4.5 for rating prediction using the IndoBERT model, the application proves to be able to provide relevant and useful recommendations for professionals in the film industry.

Keywords:

IndoBERT, Indonesian RoBERTa, genre prediction, rating prediction, movie synopsis, web application

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR SEGMEN PROGRAM

DAFTAR TABEL