

Analisis Kinerja Model Transformers untuk Klasifikasi Genre Musik = Transformers Performance Analysis for Music Genre Classification

Evan Eka Wijaya, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920519767&lokasi=lokal>

Abstrak

Klasifikasi genre musik merupakan salah satu bidang dari Music Information Retrieval (MIR) yang menggunakan pola-pola spektral dalam rekaman audio digital sebagai fitur untuk membentuk sebuah sistem yang dapat menentukan genre dari sebuah musik secara otomatis. Beberapa model deep learning telah dikembangkan untuk memperoleh performa terbaik dalam melakukan klasifikasi genre musik. Tiga di antaranya adalah Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), dan model hybrid CNN-LSTM. Walaupun model-model tersebut mampu memberikan hasil yang cukup memuaskan, model-model tersebut memiliki kekurangan masing-masing. Model CNN kurang dapat memperhitungkan urutan-urutan fitur pada data berurutan dan model LSTM tidak dapat melakukan komputasi secara paralel. Ketiga model tersebut juga membutuhkan pengulangan dan konvolusi yang kompleks, serta waktu yang cukup panjang untuk perhitungan berurutan. Transformers merupakan arsitektur model yang tidak lagi mengandalkan recurrence/pengulangan, melainkan mekanisme attention yang dapat memperhitungkan urutan-urutan data pada data berurutan dan melakukan perhitungan paralel sehingga jangka waktu yang dibutuhkan dalam perhitungan lebih singkat. Melihat keberhasilan dan kepopuleran dari Transformer pada berbagai bidang seperti Bidirectional Encoder Representations from Transformers (BERT) pada bidang Natural Language Processing dan Vision Transformers pada bidang Computer Vision, pada skripsi ini dilakukan analisis mengenai kinerja model Transformers dalam permasalahan klasifikasi genre musik dibandingkan dengan model CNN, LSTM, dan CNN-LSTM.

.....Music genre classification is one of the fields of Music Information Retrieval (MIR) that uses spectral patterns in digital audio recording as features to build a system that can automatically classify a music's genre. Several deep learning models have been developed to get the best performance in classifying music genres. Three of them are Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and hybrid CNN-LSTM model. Although those models can give satisfactory results, each model has their own weakness. CNN is less able to consider the sequences in sequential data and LSTM is not able to do parallel computation. All these models also require complex recurrences and convolutions, as well as quite a long time for sequential calculations. Transformers is a model architecture that no longer relies on recurrences, but rather on an attention mechanism that can consider the sequences in data and perform parallel calculations so that the time required for calculation is shorter. Looking into the success and popularity of Transformers in various fields such as BERT in the field of NLP and Vision Transformers in the field of Computer Vision, this thesis analyzes the performance of Transformers on music genre classification compared to CNN, LSTM, and CNN-LSTM.