Abstract

The application of recommendation systems (RS) has been applied in various types of platforms, especially applications for watching movies such as Netflix and Disney+. The designed RS purposed to make users easy to choosing a movie, because currently the number of movie productions is increasing every day, so this movie RS makes it easy for users. This research purposed to produce a Content Based Filtering (CBF) movie RS by comparing the performance of several different semantic methods to evaluate the accuracy, and efficiency of each approach, and find the most suitable approach to be able to improve the quality of performance and get the best rating prediction results in order to avoid problems from suboptimal system performance results. The semantic approaches compared are Term Frequency-Inverse Intentional Frequency (TF-IDF), Bidirectioal Encoder Representations from Transformers (BERT), Generative Pre-Trained Transformer 2 (GPT-2), Robustly optimized BERT approach (RoBERTa), and implementing Recurrent Neural Network (RNN) model for rating prediction result classification. The data used to generate a RS by applying the use of 854 movie data, and 39 accounts with a total of 34,056 movie reviews on Twitter. In the classification process, RNN model and Stochastic Gradient Descent optimization, the measurement results with confusion matrix by classifying RoBERTa rating predictions obtained an evaluation value, accuracy 95.59%, precision 95.76%, recall 95.41% and F-1 value 95.58%, the evaluation shows the RoBERTa model produces excellent performance in classifying accurate and consistent ratings. So this research has successfully obtained a method that produces rating predictions.