

ABSTRACT

Recommendation system is developed to match consumers with product to meet their variety of special needs and tastes in order to enhance user satisfaction and loyalty. The popularity of personalized recommendation system has been increased in recent years and applied in several areas include movies, songs, books, news, friends recommendations on social media, travel products, and other products in general. Collaborative Filtering methods are widely used in recommendation systems. The Collaborative Filtering method is divided into Neighborhood-based and Model-based. Neighborhood Collaborative Filtering methods have issues such as Scalability (train big amount of data) and Sparsity (many data have not yet rated by users). Another issue on recommender system overfitting, which caused having less prediction accuracy. In this study, we are implementing Matrix Factorization which is part of Model-based that learns latent factor for each user and item and using them to make rating predictions. The method will be trained using Stochastic Gradient Descent and optimization of regularization hyperparameter. However, Matrix Factorization method has issue on processing time. In the end, Neighborhood-based Collaborative Filtering and Matrix Factorization with different values of regularization hyperparameter are compared based on the prediction accuracy and processing time. Our results show that Matrix Factorization method is better than Item-based Collaborative filtering method and even better with tuning the regularization hyperparameter by achieving lowest RMSE score. Our results also show that Matrix Factorization methods take longer processing time than Item-based Collaborative Filtering method and Matrix Factorization with smallest regularization value ($1e-5$) takes longer processing time over the other values ($1e-8$ and $1e-5$)

Keywords : Recommendation System, Collaborative Filtering, Matrix Factorization, Regularization

INTISARI

Sistem rekomendasi dikembangkan untuk mencocokkan konsumen dengan produk yang memenuhi berbagai kebutuhan dan selera khusus mereka sehingga bisa kepuasan dan loyalitas pengguna. Popularitas sistem rekomendasi yang terpersonalisasi telah meningkat dalam beberapa tahun terakhir dan diterapkan di beberapa bidang termasuk film, lagu, buku, berita, rekomendasi teman di media sosial, produk perjalanan, dan produk lainnya secara umum. Metode Collaborative Filtering banyak digunakan dalam sistem rekomendasi. Metode collaborative filtering dibagi menjadi Neighborhood-based dan Model-based. Metode neighborhood collaborative filtering memiliki masalah seperti Scalability (melatih data dalam jumlah besar) dan Sparsity (banyak data belum dinilai oleh pengguna). Masalah lain pada sistem rekomendasi adalah overfitting, yang menyebabkan kurang akuratnya prediksi. Dalam penelitian ini, peneliti menerapkan metode Matrix Factorization yang merupakan bagian dari Model-based yang mempelajari latent factor untuk setiap pengguna dan item dan menggunakannya untuk membuat prediksi peringkat. Metode ini akan dilatih menggunakan Stochastic Gradient Descent dan proses optimalisasi dari regularization hyperparameter. Namun, metode matrix factorization memiliki masalah pada waktu pemrosesan. Pada akhirnya, Neighborhood-based Collaborative Filtering dan Matrix Factorization dengan nilai regularization hyperparameter yang berbeda dibandingkan berdasarkan tingkat akurasi prediksi dan waktu pemrosesan. Hasil kami menunjukkan bahwa metode matrix factorization lebih baik daripada metode item-based collaborative filtering dan bahkan lebih baik dengan perubahan regularization hyperparameter dengan mencapai skor RMSE terendah. Hasil penelitian ini juga menunjukkan bahwa metode matrix factorization membutuhkan waktu pemrosesan yang lebih lama daripada metode item-based collaborative filtering dan matrix factorization dengan nilai regularization terkecil ($1e-5$) membutuhkan waktu pemrosesan yang lebih lama daripada nilai lainnya ($1e-8$ dan $1e-5$)

Kata Kunci : Sistem Rekomendasi, Collaborative Filtering, Matrix Factorization, Regularization