

## BAB V PENUTUP

### 5.1 Kesimpulan

Berikut merupakan kesimpulan dari hasil penelitian berdasarkan rumusan masalah :

1. Pengujian *Mean Squared Difference* dan *Pearson Correlation* dengan algoritma *K-Nearest Neighbor* berbasis *item* melakukan langkah-langkah berupa pengumpulan data, pembersihan data, pembagian data, pencarian nilai *similarity*, perhitungan KNN dan pengukuran evaluasi menggunakan RMSE.
2. *Mean Squared Difference* merupakan perhitungan *similarity* yang sedikit lebih baik dibandingkan dengan *Pearson Correlation* pada penelitian ini yang menggunakan dataset *rating software*. Terlihat pada hasil pengujian RMSE dengan nilai 1.65350053 serta kecepatan waktu latih dan waktu uji yang lebih cepat dibandingkan metode *Pearson*.
3. Berdasarkan hasil yang telah dilakukan pada tahap evaluasi, akurasi dari sistem rekomendasi dipengaruhi oleh banyaknya  $k$  yang diberikan baik yang menggunakan metode MSD maupun *Pearson*. Pada metode MSD,  $k$  terbaik adalah  $k = 8$  dengan nilai RMSE 1.647825302. Untuk kecepatan metode MSD, *train time* terbaik pada  $k = 8$  dengan waktu 4.79854 detik dan *test time* terbaik pada  $k = 1$  dengan waktu 0.03077 detik. Sedangkan pada metode *Pearson*,  $k$  yang terbaik adalah  $k = 3$  dengan nilai RMSE 1.657728093. Untuk kecepatan metode *Pearson*, *train time* terbaik pada  $k = 7$  dengan waktu 9.08306 detik dan *test time* terbaik pada  $k = 1$  dengan waktu 0.03318 detik.
4. Metode *Mean Squared Difference* dan metode *Pearson Correlation* tampak memiliki perbedaan nilai pengukuran akurasi RMSE saat menggunakan nilai  $k$  yang berbeda-beda. Setiap tambahan nilai  $k$  pada metode MSD, menghasilkan nilai RMSE yang rendah nilainya sehingga dapat menghasilkan prediksi yang semakin baik. Sedangkan nilai RMSE pada *Pearson* cenderung stabil meski  $k$  meningkat.

5. Terdapat beberapa faktor yang menyebabkan besar kecilnya suatu nilai akurasi RMSE berdasarkan penelitian ini yaitu banyaknya data dan juga pemilihan nilai  $k$  yang dapat mempengaruhi hasil dari evaluasi meski tidak memberikan hasil yang signifikan.

## 5.2 Saran

Berikut merupakan saran dari penulis agar penelitian ini dapat menghasilkan pengujian sistem rekomendasi yang lebih maksimal :

1. Melakukan perbandingan dengan menggunakan library selain yang digunakan pada penelitian ini.
2. Melakukan pengujian tidak hanya dengan 25000 baris data saja, akan tetapi menggunakan seluruh data yang ada pada dataset.
3. Mengatasi fenomena *memory leaks* dengan cara menggunakan perangkat yang lebih baik, menyewa *cloud computing* di *Google Collab* atau dengan menggunakan antrian proses komputasi teknologi seperti *redis/rabbitMQ*.
4. Melakukan tuning hyperparameter untuk mencari  $k$  yang paling optimal, sehingga dapat menghasilkan sistem rekomendasi yang lebih baik.