

Implementasi Metode *Collaborative Filtering* pada Algoritma Sistem Rekomendasi Destinasi Wisata di Aceh

Farrah Fahira Murzani¹, Dede Brahma Arianto²

¹ Jurusan Informatika, Universitas Syiah Kuala

Jl. Syech Abdurrauf No.3, Kopelma Darussalam, Kec. Syiah Kuala, Kota Banda Aceh, Aceh 23111

¹farrahfahira2202@gmail.com

²Prodi Magister Informatika, Universitas Islam Indonesia

Jl. Kaliurang Km. 14,5, Krawitan, Umbulmartani, Kec. Ngemplak, Kab. Sleman, Daerah Istimewa Yogyakarta 55584

²dede.brahma2@gmail.com

Abstrak— Provinsi Aceh, kaya akan warisan budaya dan sumber daya alam, menawarkan 833 objek wisata dan 1.003 bangunan bersejarah. Menghadapi tantangan memilih destinasi, sistem rekomendasi, terutama *Collaborative Filtering* (CF), menawarkan solusi. Model-based CF melibatkan langkah awal pengolahan data offline untuk membuat model yang merepresentasikan pola interaksi pengguna-item. Salah satu metodenya adalah pemfaktoran matriks dengan penggunaan teknik embedding. Proses ini memungkinkan model untuk membuat rekomendasi destinasi wisata dengan memahami preferensi pengguna melalui representasi vektor padat. Penelitian ini bertujuan membangun sistem rekomendasi destinasi wisata di Aceh menggunakan metode *collaborative filtering*, memberikan 7 rekomendasi teratas kepada pengguna.

Kata Kunci — Sistem Rekomendasi, Destinasi Wisata Aceh, *Collaborative Filtering*, Matrix Factorization.

I. PENDAHULUAN

Sebagai provinsi yang kaya akan warisan budaya dan sumber daya alam melimpah, Aceh memiliki banyak objek wisata menarik. Tercatat di dalam *Aceh Tourism Magazine* tahun 2022 [1], Aceh memiliki 833 objek wisata dan 1.003 bangunan bersejarah yang bisa dijadikan destinasi wisata bagi wisatawan. Dalam menghadapi beragam pilihan destinasi, wisatawan sering mengalami kesulitan dalam menemukan tempat yang sesuai dengan preferensi dan kebutuhan mereka. Salah satu cara yang dapat membantu wisatawan dalam memilih objek wisata adalah dengan menggunakan sistem rekomendasi.

Sistem rekomendasi dapat membantu mengatasi permasalahan kelebihan informasi seperti ini dengan mencari tahu preferensi penggunaannya dan menyaring item untuk akhirnya memberikan rekomendasi yang sesuai dengan preferensinya [2]. Membangun sistem rekomendasi memerlukan data input beserta algoritma rekomendasi untuk menghasilkan hasil rekomendasi yang sesuai. Algoritma rekomendasi yang paling umum digunakan adalah *Collaborative Filtering* (CF). Selain implementasi dan penjelasannya cukup mudah untuk dipahami, untuk

mengimplementasikan metode ini hanya membutuhkan data interaksi *user-item* [3]. Data interaksi *user-item* mencakup informasi mengenai bagaimana pengguna berinteraksi atau berhubungan dengan berbagai *item* atau objek dalam sistem [4]. Dalam konteks sistem rekomendasi, data ini mencerminkan aktivitas pengguna seperti penilaian, ulasan, atau perilaku interaksi lainnya terhadap objek-objek wisata. Sebagai contoh, data interaksi *user-item* untuk destinasi wisata mungkin mencakup penilaian pengguna terhadap suatu tempat, riwayat kunjungan, atau preferensi yang tergambar dari pencarian dan klik sebelumnya.

Cara kerja metode CF adalah dengan membandingkan preferensi atau perilaku pengguna untuk menghasilkan rekomendasi berdasarkan kesamaan dengan pengguna lain. Dengan menggunakan perbandingan ini, sistem dapat menghasilkan rekomendasi baru untuk pengguna berdasarkan pola perilaku yang serupa [5].

Collaborative Filtering secara umum terbagi dua, yaitu *memory-based* dan *model-based*. *Memory-based* melibatkan penggunaan “memori” atau “rekam jejak” langsung dari data interaksi pengguna dengan item untuk membuat rekomendasi. Tidak ada proses membangun model *machine learning* yang digunakan pada metode ini. *Model-based*, sementara itu melibatkan proses melatih model untuk menghasilkan prediksi terhadap item yang belum di-*rate* oleh user [6].

Pada model-based CF, langkah pertama melibatkan mengolah data secara *offline* untuk pembuatan model yang mewakili pola atau karakteristik dari data interaksi pengguna-item [7]. Proses ini dapat mencakup teknik-teknik *machine learning* seperti regresi, clustering, atau bahkan teknik *deep learning* untuk menggambarkan hubungan kompleks antara pengguna dan item [8].

Salah satu metode yang umum digunakan dalam model-based *Collaborative Filtering* adalah penggunaan teknik pemfaktoran matriks. Pemfaktoran matriks melibatkan representasi vektor yang padat dari entitas, seperti pengguna atau item, dalam ruang fitur yang memiliki dimensi lebih rendah [9]. Dalam konteks rekomendasi, embedding

digunakan untuk merepresentasikan pengguna dan item sehingga model dapat memahami hubungan antara mereka.

Proses menciptakan embedding untuk pengguna dan item melibatkan pembelajaran dari data. Model akan diberikan data interaksi pengguna-item, seperti penilaian atau histori kunjungan, dan melalui proses pelatihan (*training*), model akan mengoptimalkan parameter-parameter untuk menghasilkan *embedding* yang mewakili pola atau karakteristik dari data tersebut.

Setelah model terlatih, *embedding* pengguna dan item dapat digunakan untuk membuat prediksi terhadap interaksi yang belum terjadi. Misalnya, model dapat membuat prediksi sejauh mana suatu pengguna mungkin menyukai destinasi wisata tertentu yang belum dia kunjungi berdasarkan *embedding* pengguna dan item yang telah dipelajari.

Teknik *embedding* ini membantu mengatasi masalah ruang fitur yang besar dan kompleks, di mana setiap pengguna dan item mungkin memiliki banyak atribut atau fitur yang sulit diolah secara langsung [10]. *Embedding* merangkum informasi ini ke dalam representasi vektor padat, memungkinkan model untuk memahami dan menangkap hubungan yang lebih kompleks antara pengguna dan item dalam ruang fitur yang lebih terstruktur.

Pada penelitian ini akan dibangun sistem rekomendasi destinasi wisata di Aceh menggunakan metode *collaborative*

filtering yang melibatkan proses membangun model di dalamnya untuk menghasilkan 7 rekomendasi teratas kepada pengguna.

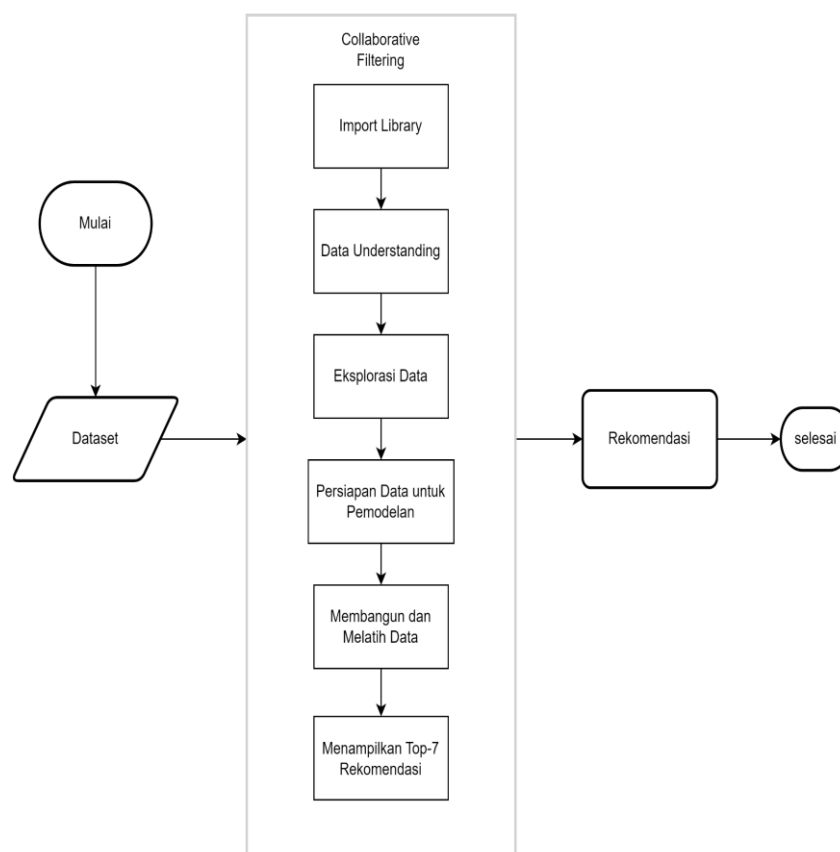
II. DATASET

Dataset yang digunakan pada penelitian ini didapatkan dengan cara melakukan *web scraping* terhadap data ulasan pada situs *Google Maps*. Destinasi wisata yang digunakan adalah objek-objek wisata yang ada di Provinsi Aceh

Terdapat dua macam data yang digunakan, yaitu data *rating* yang berupa data interaksi *user-item*, dan data *destination* yang berupa data informasi mengenai destinasi wisata. Data *rating* digunakan untuk membangun model *collaborative filtering* sedangkan data informasi item hanya digunakan untuk keperluan eksplorasi data seperti menampilkan visualisasi data.

III. METODE PENELITIAN

Adapun alur dari penelitian ini secara garis besar terbagi menjadi enam tahapan, yaitu *import library*, data understanding, eksplorasi data, persiapan data untuk pemodelan, membangun dan melatih data, dan menampilkan top-7 rekomendasi.



Gambar 1. Alur pengerjaan

A. Import Library

Mengimport *library* yang dibutuhkan adalah tahapan yang pertama kali harus dilakukan. Dalam membangun sistem rekomendasi, perlu mengimport beberapa *library* untuk proses pengolahan data, visualisasi, dan pemodelan *machine learning*, yaitu:

- 1) Library untuk pengolahan data: Pandas dan NumPy.
- 2) Library untuk visualisasi data: Seaborn dan Matplotlib.
- 3) Library untuk pemodelan machine learning: keras dan tensorflow.

B. Data Understanding

Pada tahapan ini dilakukan pemahaman awal terhadap data yang akan digunakan. Dua file data dibaca kemudian ditampilkan tiga baris pertama dari masing-masing dataset untuk mendapatkan gambaran awal mengenai struktur dan isinya.

Selanjutnya, kita mencetak informasi mengenai dataset rating menggunakan metode *info()*. Hal ini berguna untuk mengetahui tipe data, jumlah nilai non-null, dan penggunaan memori.

Dengan langkah-langkah ini, kita dapat memahami struktur dan ukuran dari kedua dataset, serta mendapatkan informasi awal yang berguna untuk analisis lebih lanjut.

C. Eksplorasi Data

Pada tahapan ini akan dilakukan eksplorasi data. Tujuannya adalah mendapatkan wawasan yang lebih mendalam. Pada proses ini akan dibuat visualisasi data.

D. Persiapan Data untuk Pemodelan

Pada tahapan ini akan dilakukan persiapan data untuk memodelkan sistem rekomendasi. Langkah pertama melibatkan *encoding* pada kolom 'user_id' dan 'place_name' agar dapat direpresentasikan dalam bentuk nilai numerik. Fungsi *dict_encoder* digunakan untuk mengubah setiap nilai unik dalam kolom menjadi angka, dan hasilnya disimpan dalam dua set kamus untuk masing-masing arah konversi.

Selanjutnya, proses mapping dilakukan untuk mengaitkan kolom 'user_id' dan 'place_name' dengan representasi numerik yang telah dihasilkan. Hal ini memungkinkan pengguna dan destinasi direpresentasikan secara numerik dalam model. Jumlah pengguna (*num_users*) dan destinasi (*num_places*) dihitung berdasarkan kamus encoding yang telah dibuat.

Data rating kemudian diubah menjadi tipe data float32 untuk memastikan kesesuaian dengan model. Selanjutnya, nilai minimum dan maksimum *rating* diidentifikasi untuk memahami rentang nilai rating dalam dataset.

Dengan langkah-langkah ini, data telah siap untuk digunakan dalam pemodelan sistem rekomendasi, dan informasi mengenai jumlah pengguna, destinasi, serta rentang nilai *rating* telah diperoleh untuk memandu konfigurasi model.

E. Membangun dan Melatih Model

Pada tahapan ini akan dilakukan langkah-langkah untuk membangun dan melatih model rekomendasi telah dilakukan. Pertama, data dibagi menjadi 80% data pelatihan dan 20% data validasi.

Selanjutnya model rekomendasi, yang diberi nama *RecommenderNet*, dibangun menggunakan lapisan-lapisan embedding untuk merepresentasikan pengguna, destinasi, dan bias khususnya. Penggabungan representasi ini dilakukan dalam metode *call*, dan outputnya diaktivasi menggunakan fungsi sigmoid.

Setelah model dibangun, langkah selanjutnya adalah kompilasi menggunakan fungsi kerugian dan optimizer tertentu. Model kemudian dilatih menggunakan data pelatihan dan validasi selama 100 epoch. Sebuah *callback* (*myCallback()*) digunakan untuk memantau proses pelatihan dan memberhentikannya jika kondisi tertentu terpenuhi.

Dengan demikian, model rekomendasi telah berhasil dibangun dan dilatih, siap untuk memberikan rekomendasi berdasarkan input pengguna dan destinasi.

F. Menampilkan Top-7 Rekomendasi

Pada tahap ini, model rekomendasi digunakan untuk memprediksi rating pengguna untuk destinasi yang belum dikunjungi (*user_place_array*). *Rating* hasil prediksi diurutkan, dan indeks destinasi teratas (top 7) diambil sebagai rekomendasi.

Selanjutnya, hasil rekomendasi tersebut diinterpretasikan dan ditampilkan. Pertama, ditampilkan tempat-tempat dengan *rating* wisata paling tinggi dari pengguna berdasarkan data yang telah dikunjungi sebelumnya. Kemudian, ditampilkan juga top 7 destinasi yang direkomendasikan oleh model. Informasi seperti nama tempat, kategori, dan peringkat diberikan untuk memberikan pandangan yang lebih lengkap.

IV. HASIL DAN PEMBAHASAN.

A. Top 10 Destinasi Wisata dengan Jumlah Rate Terbanyak

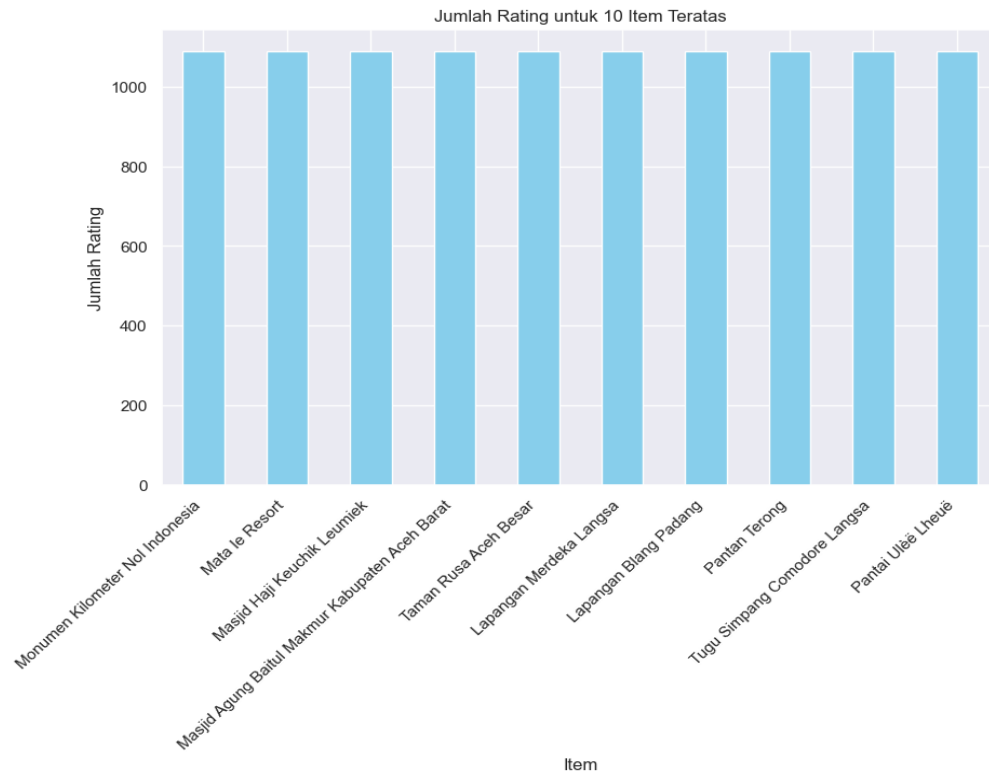
Didapatkan informasi top 10 destinasi wisata dengan jumlah rate terbanyak adalah Monumen Kilometer Nol Indonesia, Mata Ie Resort, Masjid Haji Keuchik Leumiek, Masjid Agung Baitul Makmur Kabupaten Aceh Barat, Taman Rusa Aceh Besar, Lapangan Merdeka Langsa, Lapangan Blang Padang, Pantan Terong, Tugu Simpang Comodore Langsa, dan Pantai Ulee Lheue.

B. Jumlah Destinasi Wisata Berdasarkan Kabupaten/Kota

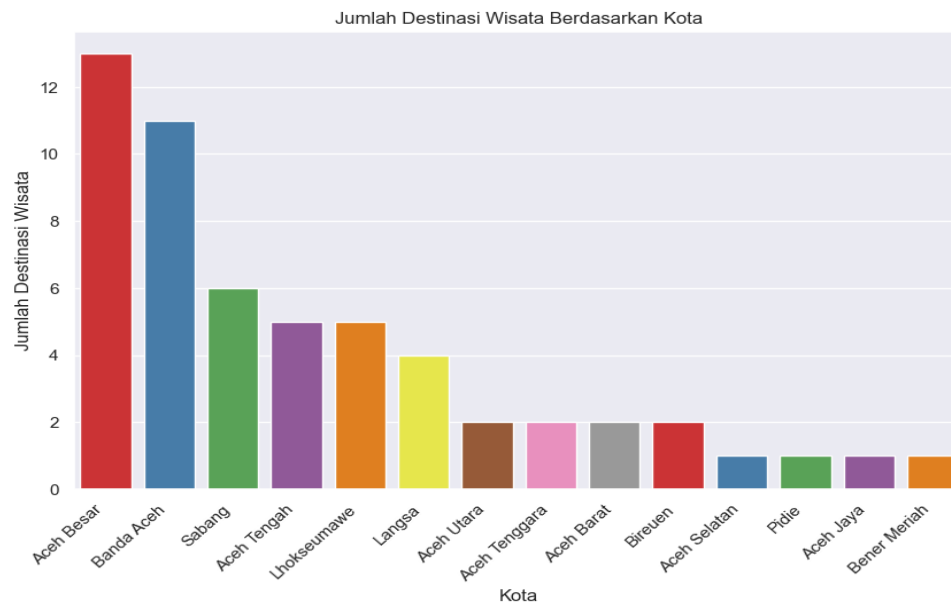
Didapatkan informasi berupa jumlah destinasi wisata yang ada pada tiap kota.

- Pada Kabupaten Aceh Besar, terdapat 13 destinasi wisata.
- Pada Kota Banda Aceh, terdapat 11 destinasi wisata.
- Pada Kota Sabang, terdapat 6 destinasi wisata.
- Pada Kabupaten Aceh Tengah, terdapat 5 destinasi wisata.
- Pada Kota Lhokseumawe, terdapat 5 destinasi wisata.
- Pada Kota Langsa, terdapat 4 destinasi wisata.

- Pada Kabupaten Aceh Utara, terdapat 2 destinasi wisata.
- Pada Kabupaten Aceh Tenggara, terdapat 2 destinasi wisata.
- Pada Kabupaten Aceh Barat, terdapat 2 destinasi wisata.
- Pada Kabupaten Bireuen, terdapat 2 destinasi wisata.
- Pada Kabupaten Aceh Selatan, terdapat 1 destinasi wisata.
- Pada Kabupaten Pidie, terdapat 1 destinasi wisata.
- Pada Kabupaten Aceh Jaya, terdapat 1 destinasi wisata.
- Pada Kabupaten Bener Meriah, terdapat 1 destinasi wisata.



Gambar 2 Destinasi Wisata dengan Rating Terbanyak

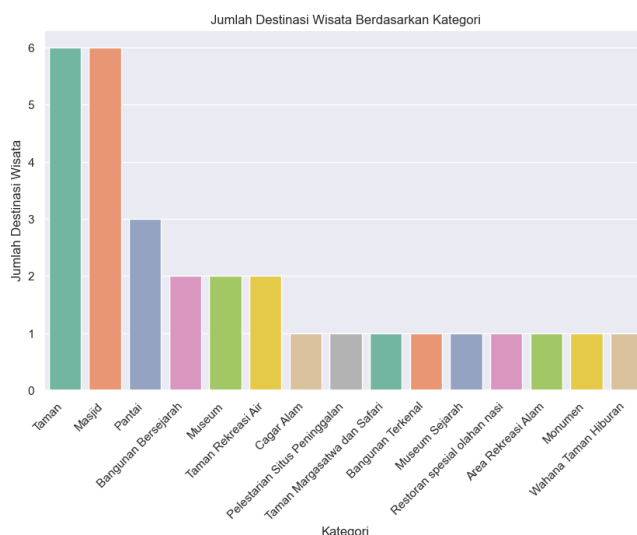


Gambar 3 Jumlah Destinasi Wisata Berdasarkan Kota

C. Jumlah Destinasi Wisata untuk Tiap Kategori

Jumlah Destinasi Wisata untuk tiap kategorikan disajikan dalam diagram batang.

- Kategori Taman dan Masjid memiliki 6 destinasi wisata.
- Kategori Pantai memiliki 3 destinasi wisata.
- Kategori Bangunan Bersejarah, Museum, dan Taman Rekreasi Air memiliki 2 destinasi wisata.
- Kategori Cagar Alam, Pelestarian Situs Peninggalan, Taman Margasatwa dan Safari, Bangunan Terkenal, Museum Sejarah, Restoran spesial olahan nasi, Area Rekreasi Alam, Monumen, dan Wahana Taman Hiburan memiliki 1 destinasi wisata.



Gambar 4 Jumlah Destinasi Wisata Berdasarkan Kategori

D. Hasil Evaluasi Model Menggunakan RMSE

Root Mean Square Error (RMSE) adalah metrik evaluasi yang umum digunakan dalam menilai kualitas prediksi. RMSE digunakan untuk mengukur sejauh mana perbedaan antara hasil prediksi dan nilai aktual yang ingin diprediksi. Semakin kecil nilai RMSE yang dihasilkan, semakin baik kualitas prediksi [11].

Semakin kecil nilai RMSE, maka akan semakin bagus. Hal ini berarti semakin akurat model dalam memprediksi nilai sebenarnya. Nilai RMSE yang rendah menunjukkan bahwa perbedaan antara prediksi model dan nilai sebenarnya cenderung minim [12].

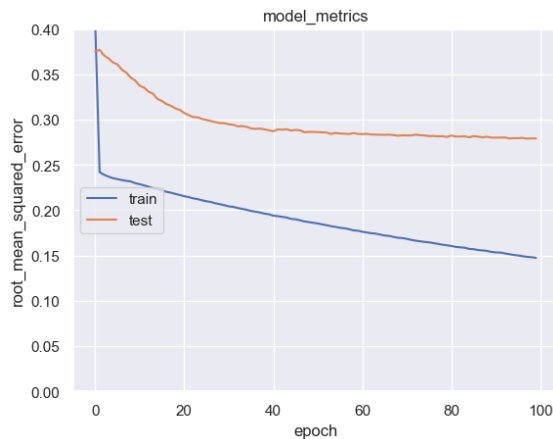
Hasil RMSE pada data latih dan data uji yang dihasilkan dalam rentang epoch 1-100 dapat dilihat pada gambar 5..

E. Hasil Top-7 Rekomendasi Destinasi Wisata di Aceh untuk Active User

Diberikan hasil top 7 rekomendasi destinasi wisata di Aceh untuk user dengan user_id 113207903499711820739.

V. KESIMPULAN

Dalam penelitian ini, kami telah berhasil mengimplementasikan model Collaborative Filtering menggunakan pendekatan pengembangan model keras untuk meningkatkan kualitas sistem rekomendasi destinasi wisata di Aceh. Hasil evaluasi kinerja model menunjukkan bahwa Root Mean Square Error (RMSE) yang dihasilkan sebesar 0.1475. Angka RMSE yang rendah ini menandakan bahwa model yang dikembangkan mampu memberikan rekomendasi destinasi wisata dengan tingkat akurasi yang tinggi, menciptakan pengalaman pengguna yang lebih memuaskan.



Gambar 5 Nilai RMSE terhadap Data Latih dan Data Uji

Daftar rekomendasi untuk: User 113207903499711820739
=====

Tempat dengan rating wisata paling tinggi dari user

Arena Pacuan Kuda Belang Bebangka : Tujuan Wisata
Bur Telege - Takengon : Tujuan Wisata
Geurutee Highland : Tujuan Wisata
Wahana Impian Malaka 69 : Wahana Taman Hiburan
Wisata Pulau Pinang : Tujuan Wisata

Top 7 place recommendation

- 1 . Air Terjun Silangit-langit
Tujuan Wisata
- 2 . Batu Gajah Namobuaya
Tujuan Wisata
- 3 . Ketambe Adventure | Trekking | Rafting | Hiking | Orang Utan
Tujuan Wisata
- 4 . Kolam Wisata Mata Air Sarindan
Lahan Piknik
- 5 . Masjid Haji Keuchik Leumiek
Masjid
- 6 . Masjid Raya Baiturrahman
Masjid
- 7 . Panjang Beach Resort
Tujuan Wisata

=====

Gambar 6 Hasil dari Rekomendasi yang Diberikan pada Active User

REFERENSI

- [1] *Aceh Tourism Magazine*, 2022.
- [2] S. Z. Putri, M. Hariadi, dan R. F. Rachmadi, "Rekomendasi Barang pada E-commerce menggunakan Content Based Recommender System berbasis Ethereum Blockchain," *Jurnal Teknik ITS*, 2023.
- [3] M. N. S. Bahri, I. P. Y. D. Jaya, B. Dirgantoro, I. Istikmal, U. A. Ahmad, dan R. R. Septiawan, "Implementasi Sistem Rekomendasi Makanan pada Aplikasi EatAja Menggunakan Algoritma Collaborative Filtering," *Jurnal Multinetics*, vol. 7, no. 2, November 2021.
- [4] G. Behera dan N. Nain, "Trade-off between memory and model-based collaborative filtering recommender system," dalam *Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences*, hal. 137–146, Springer, 2022.
- [5] M. Nilashi, K. Bagherifard, O. Ibrahim, H. Alizadeh, L. A. Nojeem, dan N. Roozgar, "Collaborative filtering recommender systems," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 5, no. 16, hal. 4168–4182, 2013. <https://doi.org/10.19026/rjaset.5.4644>
- [6] C. C. Aggarwal, "Recommender systems," Springer International Publishing, Yorktown Heights, NY, USA, 2016.
- [7] P. Resnick dan H. R. Varian, "Recommender Systems," *Communications of the ACM*, vol. 40, no. 3, hal. 56–58, 1997.
- [8] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, dan R. Kashef, "Recommendation Systems: Algorithms, Challenges, Metrics, and Business Opportunities," *Appl. Sci.*, vol. 10, no. 21, 2020, doi: 10.3390/app10217748.
- [9] M. Nilashi, K. Bagherifard, O. Ibrahim, H. Alizadeh, L. A. Nojeem, dan N. Roozgar, "Collaborative filtering recommender systems," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 5, no. 16, hal. 4168–4182, 2013. <https://doi.org/10.19026/rjaset.5.4644>
- [10] D. Das, L. Sahoo, and S. Datta, "A Survey on Recommendation System," *International Journal of Computer Applications*, vol. 160, no. 7, pp. 6–10, 2017, doi: 10.5120/ijca2017913081.
- [11] P. Nagarnaik and A. Thomas, "Survey on recommendation system methods," *2nd International Conference on Electronics and Communication Systems, ICECS 2015*, vol. 137, no. 7, pp. 1603–1608, 2015, doi:10.1109/ICECS.2015.7124857
- [12] Manjula, R., & Chilambuchelvan, A. 2016. Content Based Filtering Techniques in Recommendation System using user preferences. *International Journal of Innovations in Engineering and Technology*, 7(4), 151