

FAKTORISASI MATRIKS MENGGUNAKAN STOCHASTIC GRADIENT DESCENT UNTUK OPTIMASI SISTEM REKOMENDASI HOTEL

Inayatul Maula¹⁾, Abu Tholib²⁾, M. Noer Fadli Hidayat³⁾

^{1, 2, 3)}Teknik Informatika, Universitas Nurul Jadid

e-mail: inytl.maula18@gmail.com¹⁾, ebuenje@gmail.com²⁾, fadli@unuja.ac.id³⁾

ABSTRAK

Di dunia serba digital seperti sekarang, sistem rekomendasi memiliki peranan yang sangat penting untuk membantu pengguna menemukan hotel yang sesuai dengan preferensi mereka. Penelitian ini berfokus pada pengembangan sistem rekomendasi hotel dengan menggabungkan metode faktorisasi matriks dengan algoritma Stochastic Gradient Descent (SGD). Metode faktorisasi matriks digunakan untuk memodelkan data peringkat hotel sebagai hasil kali dari matriks pengguna dan matriks hotel. Sedangkan untuk algoritma Stochastic Gradient Descent (SGD) berperan dalam mengoptimalkan parameter model secara efisien, dimana metode akan diuji pada dataset peringkat hotel atau rating. Evaluasi kinerja model pada penelitian ini, menggunakan metrik seperti Root Mean Squared (RMSE), Mean Squared Error (MSE), dan Mean Absolute Error (MAE). Penelitian ini menunjukkan hasil yang cukup akurat, dengan nilai RMSE sebesar 0.370312, nilai MSE sebesar 0.137131, dan nilai MAE sebesar 0.089932. Hasil ini menunjukkan bahwa penggabungan metode faktorisasi matriks dengan Stochastic Gradient Descent (SGD) dapat menjadi solusi yang efektif untuk membangun sistem rekomendasi hotel sesuai dengan preferensi pengguna.

Kata Kunci: Faktorisasi Matriks, Stochastic Gradient Descent, Sistem Rekomendasi

ABSTRACT

In today's digital world, recommendation systems have a very important role to help users find hotels that match their preferences. This research focuses on developing a hotel recommendation system by combining matrix factorization method with Stochastic Gradient Descent (SGD) algorithm. The matrix factorization method is used to model the hotel ranking data as the product of the user matrix and the hotel matrix. While for the Stochastic Gradient Descent (SGD) algorithm plays a role in optimizing model parameters efficiently, where the method will be tested on hotel rating datasets or ratings. Evaluation of model performance in this study, using metrics such as Root Mean Squared (RMSE), Mean Squared Error (MSE), and Mean Absolute Error (MAE). This study shows fairly accurate results, with an RMSE value of 0.370312, an MSE value of 0.137131, and an MAE value of 0.089932. These results show that combining the matrix factorization method with Stochastic Gradient Descent (SGD) can be an effective solution for building a hotel recommendation system according to user preferences.

Keywords: Matrix Factorization, Stochastic Gradient Descent, Recommendation System.

I. PENDAHULUAN

Tripadvisor merupakan salah satu platform ternama yang menjadi acuan bagi calon wisatawan dalam mencari informasi seputar hotel dan pariwisata[1][2]. Pada Tripadvisor ini, kita dapat menemukan informasi terkait hotel seperti, fasilitas hotel dan pelayanannya, harga per-kamar, rating hotel, bahkan kita dapat melihat ulasan dari pengguna Tripadvisor lain[3]. Namun, meskipun Tripadvisor memberikan informasi yang bermanfaat bagi wisatawan, menemukan sebuah hotel yang sesuai dengan preferensi individu masih menjadi tantangan bagi calon wisatawan karena banyaknya pilihan hotel[4]. Inilah dimana sistem rekomendasi mulai

memasuki panggung, dengan keunggulan utamanya yakni menemukan hotel sesuai dengan preferensi pengguna[5].

Dari beberapa penelitian sebelumnya terkait sistem rekomendasi, terdapat penelitian yang menggunakan cosine similarity dan algoritma KNN[6], penggabungan Matrix Factorization dengan Neural Networks[7], serta penggunaan metode Restricted Boltzman Machine (RBM) dan Matrix Factorization-Alternating Least Squares (MF-ALS)[8]. Kelebihan dari penelitian menggunakan cosine similarity dan KNN dapat memberikan sistem rekomendasi untuk film dengan hasil yang cukup memuaskan. Namun, jurnal ini juga memiliki kelemahan yakni, masalah sparsity, sehingga sistem sulit memberikan rekomendasi yang akurat.

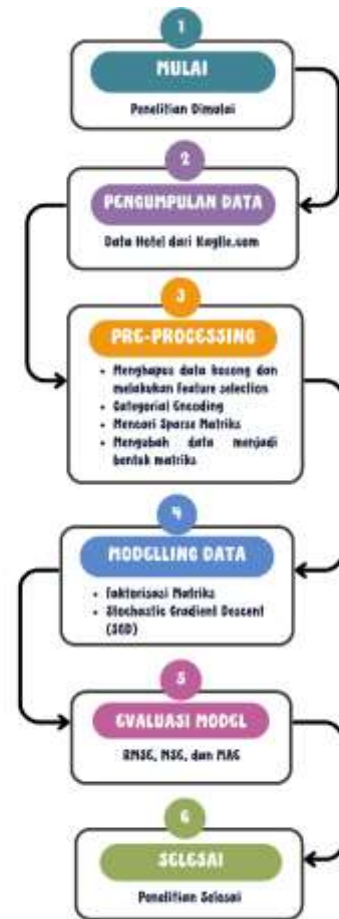
Penggabungan metode *Matrix Factorization* dan *Neural Networks* memiliki kelebihan untuk dapat meningkatkan akurasi dan dapat digunakan dalam di berbagai aplikasi (*e-commerce*, media sosial, dan industri hiburan), akan tetapi kelemahan dari metode ini yaitu, mengalami masalah *sparsity* dan *scalability* yang dapat memengaruhi rekomendasi, terbatas pada cakupan dan akurasi rekomendasi. RBM dan MF-ALS memiliki kelebihan memberikan rekomendasi berdasarkan tipe objek dan rating dari pengguna sebelumnya, sedangkan kelemahannya yaitu, kurangnya fleksibilitas dalam konfigurasi ulang setelah pelatihan selesai dan waktu eksekusi yang lama.

Dilihat dari penelitian sebelumnya, penelitian ini dapat memecahkan masalah yang terjadi seperti, masalah *sparsity* pada tahap pre-processing data, dan waktu eksekusi yang relatif cepat. Metode faktorisasi matriks juga dapat menangkap pola tersembunyi pada data[9], sehingga dapat memberikan rekomendasi yang akurat dengan memanfaatkan implisit dari pengguna dan item[10]. Sedangkan SGD sendiri dapat menangani data berskala besar[11], sehingga cocok digunakan dalam sistem rekomendasi hotel. Tujuan dari penelitian ini untuk mencari sejauh mana nilai akurasi yang dicapai. Dengan memanfaatkan keunggulan ini, diharapkan dapat ditemukan solusi yang tidak hanya meningkatkan akurasi rekomendasi, tetapi juga meningkatkan efisiensi perhitungan, sehingga memberikan pengalaman pengguna yang lebih baik.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam konteks faktorisasi matriks dan optimasi sistem rekomendasi, bukan hanya dalam konteks Tripadvisor, tetapi juga dapat menjadi dasar untuk penelitian lebih lanjut dalam pengembangan metode rekomendasi yang lebih akurat. Dengan demikian, penelitian ini tidak hanya ditujukan untuk meningkatkan kualitas layanan pada suatu platform tertentu, tapi juga memiliki potensi untuk meningkatkan pemahaman kita tentang cara terbaik mengoptimalkan sistem rekomendasi dalam skala yang lebih luas.

II. METODE PENELITIAN

Pada penelitian kali ini, kami menggunakan metode penggabungan faktorisasi matriks dengan *Stochastic Gradient Descent* untuk mengembangkan sistem rekomendasi hotel. Secara garis besar, tahapan penelitian meliputi identifikasi masalah, pengumpulan data, *pre-processing* data, modelling menggunakan faktorisasi matriks dan *Stochastic Gradient Descent* (SGD), serta evaluasi model. Adapun tahapan pada penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Metode Penelitian

A. Identifikasi Masalah

Dalam tahap ini, kami mengidentifikasi permasalahan utama yang ingin kami selesaikan dalam sistem rekomendasi hotel [12]. Sehingga fokus utama pada penelitian ini yaitu upaya untuk meningkatkan nilai akurasi rekomendasi dan mengurangi kesalahan prediksi[13]. Bukan hanya itu, penelitian ini juga digunakan untuk mengisi kekosongan pengetahuan serta dapat dijadikan sebagai referensi terkait sistem rekomendasi hotel menggunakan metode faktorisasi matriks dengan *Stochastic Gradient Descent*.

B. Pengumpulan Data

Dataset pada penelitian ini diperoleh dari website kaggle.com [14], dimana dataset yang digunakan disini berisi data rating dan review hotel dari pengguna platform Tripadvisor. Dalam penelitian ini, terdapat kurang lebih 2952 baris dan 8 variabel data yang digunakan.

C. Pre-Processing Data

Pre-processing adalah proses pembersihan, transformasi dan penyiapan data sebelum diproses lebih lanjut[15]. Pada tahap ini peneliti melakukan penghapusan data kosong pada dataset dan memilah variabel-variabel yang akan digunakan[16][17], seperti

Reviewer_Name, Location, Title, Rating, Pos_Review, dan Hotel_Name. Kemudian dilakukan *categorical encoding*, sparse matriks, dan mengubah data menjadi bentuk matriks. Hal ini dilakukan untuk meningkatkan kualitas data dan akurasi[18].

D. Modelling Data

Pada tahapan modelling data, setelah data melalui tahapan *pre-processing* data maka akan di proses oleh metode faktorisasi matriks yang dioptimasi oleh *Stochastic Gradient Descent* (SGD)

1) Faktorisasi matriks adalah suatu metode penguraian matriks besar menjadi matriks-matriks yang lebih kecil [19]. Faktorisasi matriks dapat mengubah atribut k menjadi matematika dengan mengambil pengguna u ke dalam bentuk vektor $k \times u$. Begitu juga dengan item i dengan vektor berdimensi $k \times i$. Perkiraan peringkat pengguna u untuk item i dapat dihitung dari perkalian titik dari dua vektor [20].

$$\hat{r}_{ui} = x_u^T \cdot y_i = \sum_k x_{uk} y_{ki} \quad (20)$$

Untuk memeriksa kesalahan menggunakan fungsi [21] berikut :

$$\min_{q^*, p^*} \sum_{(u,i) \in k} (R_{ui} - q_i^T p_u)^2 + \lambda (||q_i||^2 + ||p_u||^2) \quad (21)$$

Keterangan :

- λ = parameter regulasi.
- R = himpunan indeks (u, i) .
- q_i = baris ke- i dari q .
- p_u = baris ke- u dari p .

2) *Stochastic Gradient Descent* (SGD) merupakan algoritma optimasi iteratif yang digunakan dalam pembelajaran mesin[22], termasuk dalam konteks faktorisasi matriks untuk sistem rekomendasi. Pada tahap ini, setiap nilai matriks rating, sistem akan memprediksi r_{ui} dan menghitung kesalahan dengan fungsi berikut:

$$e_{ui} \stackrel{\text{def}}{=} r_{ui} - q_i^T p_u \quad (21)$$

parameter yang dipelajari, diambil turunan parsialnya dari fungsi kerugian pada setiap parameter, lalu diperbarui secara acak menggunakan gradien yang dihitung dari subset data pelatihan. Parameter model nantinya akan bergerak melawan arah gradien, sehingga sedikit demi sedikit nilai objektifnya akan berkurang[23]. Proses ini akan diulang berkali-kali hingga parameter model konvergen ke nilai optimal [24]. Berikut adalah pembaruan turunan parsial:

$$q_i \leftarrow q_i + \gamma \cdot (e_{ui} \cdot p_u - \lambda \cdot q_i)$$

$$p_u \leftarrow p_u + \gamma \cdot (e_{ui} \cdot q_i - \lambda \cdot p_u) \quad (24)$$

E. Evaluasi Model

Selanjutnya dilakukan evaluasi model faktorisasi matriks dan *Stochastic Gradient Descent* (SGD). Evaluasi model adalah proses mengukur kinerja model dalam memprediksi peringkat[25]. Untuk mengevaluasi kinerja model, beberapa metrik evaluasi seperti:

- 1) *Root Mean Squared Error* (RMSE), digunakan untuk mengukur perbedaan antara peringkat sebenarnya dan peringkat yang diprediksi oleh model. Apabila nilai RMSE yang dicapai menunjukkan angka yang rendah, itu artinya kinerja model baik.[26]
- 2) *Mean Squared Error* (MSE), digunakan untuk mengukur rata-rata kuadrat perbedaan antara nilai sebenarnya dan nilai yang diprediksi oleh model[27].
- 3) *Mean Absolute Error* (MAE), digunakan untuk mengukur rata-rata kesalahan absolut antara peringkat sebenarnya dan peringkat yang diprediksi oleh model[28].

Ketiga metrik evaluasi ini dapat dihitung menggunakan nilai peringkat yang diamati dan nilai peringkat yang diprediksi[29].

III.HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dataset yang digunakan pada penelitian ini sebanyak 2952 baris dan 8 variabel dari tahun 2019-2022. Kedelapan variabel ini meliputi Reviewer_Name, Country, Visited_Date, Rating, Title, Pos_Review, Hotel_Name dan Location. Pada penelitian ini, tidak semua variabel yang kami gunakan, tapi hanya 6 variabel saja yang akan kami pakai yaitu, Review_Name, Rating, Title, Pos_Review, Hotel_Name. Untuk kolom Country dan Visited_Date tidak digunakan karena, kedua kolom ini tidak begitu dibutuhkan dalam penelitian ini.

B. Pre-processing Data

Tahap pra-pemrosesan data merupakan langkah penting sebelum menggunakan data untuk membangun model faktorisasi matriks dan *Stochastic Gradient Descent* (SGD). Tujuannya adalah untuk menyempurnakan data sehingga dapat meningkatkan kualitas dataset.

Tahap pertama adalah penghapusan data kosong dan pemilihan variabel pada dataset. Dimana dari delapan variabel yang ada, hanya diambil enam variabel yang akan kami digunakan dalam penelitian ini. Hal ini bertujuan untuk mengoptimalkan dan membuang data

yang tidak relevan agar dapat menghindari terjadinya *overfitting*.

Tabel 1.
Ketiga variabel yang dipilih dan akan digunakan

Review e_Nam e	Location	Title	Rati ng	Pos_Re view	Hotel_ Name
Joshua	Entebbe	Excellent place to transition	10.0	Staff was very helpful and polite	Admas Grand Hotel
Howard	Kampala	Comfortabl e hotel with friendly staff in a quite location	9.0	Comfor table, excelle nt breakfa st and good menu	ONOM O Hotel Kampal a
Cherie	Entebbe	Owner couldn't have done more to accommod ate our needs. Property spotless, simple and wonderful	10.0	Quiet peaceful gardens	BlueMo nkey Guesth ouse
Greygh ost1	Entebbe	Fabulous	8.8	Fabulou s location next to the entranc e to the botanic gardens	Edd Botanic al View
Athele	Kampala	Good	7.0	The actual hotel is great clean etc.	Latitud e 0 degrees

Langkah selanjutnya adalah melakukan *categorical encoding* dari Reviewer_Name dan Hotel_Name. Pada proses ini, user_id memiliki nilai 1 sampai dengan n, dimana n adalah jumlah kategori *unique* user[30]. Setiap nilai dalam kolom mewakili pengguna dalam bentuk angka. Begitupun dengan item_id dalam hal ini Idhotel memiliki nilai dari 1 sampai n, dimana n merupakan jumlah kategori *unique* hotel dan juga diwakilkan dalam bentuk angka. Berikut merupakan hasil dari *categorical encoding* dari Reviewer_Name dan Hotel_Name:

Kolom user pada tabel 2,3,dan 4 didapatkan dari data kolom Review_Name berdasarkan dataframe `df\_user\_item`, yang kemudian diubah menjadi tipe data *category*. Begitupun dengan kolom hotel yang diisi berdasarkan data kolom Hotel_Name dari dataframe `df\_user\_item`. Dari proses ini, kami telah

mendapatkan kolom baru yakni userId dan Idhotel, dimana untuk setiap kolom user dan hotel melalui proses pengkodean dari 1 sampai n menggunakan metode `cat.codes` dari *library* pandas. Kemudian, kami melakukan tiga pemanggilan `display` untuk menampilkan empat baris pertama, empat baris terakhir dan beberapa baris sampel dari dataframe. Langkah-langkah ini dilakukan agar data dapat diidentifikasi dan diproses oleh model.

Tabel 2.
Empat sampel baris pertama dari dataframe dari pengkodean Review_Name dan Hotel_Name

user	hotel	user Id	Idh otel	ratings	location
Angus	Ataco Country Resort	77	12	7.0	Fort Portal
Pavel	Ataco Country Resort	909	12	9.0	Fort Portal
Cidri	Kampala Forest Resort	209	50	10.0	Kampal a
Josiane	Kampala Forest Resort	537	50	10.0	Kampal a

Tabel 3.
Empat sampel baris terakhir dari dataframe dari pengkodean Review_Name dan Hotel_Name

user	hotel	user Id	Idh otel	ratings	location
Agaba	Luwafu Guest House	21	63	10.0	Kampal a
Joseph	Parons Nest 1bedroom apartment	534	86	10.0	Kampal a
Alex	Parons Nest 1bedroom apartment	40	86	10.0	Kampal a
Dedan	Parons Nest 1bedroom apartment	247	86	10.0	Kampal a

Tabel 4.
Empat sampel baris random dari dataframe dari pengkodean Review_Name dan Hotel_Name

user	hotel	user Id	Idh otel	ratings	location
C	Speke Apartements	170	103	7.0	Kampal a
Thorsten	Kitante Savannah Suites	113 6	96	10.0	Kampal a
Kornel	Lake Nyamirima cottages	623	59	10.0	Fort Portal
Gershim	Ntinda View Apartments	378	79	7.0	Kampal a

Berikutnya merupakan tahapan yang sangat penting yakni membuat sparse matriks. Tahapan ini digunakan untuk dimasukkan ke dalam pembelajaran mesin, dimana hasil yang diperoleh semua matriksnya 0.

Berikut adalah hasil dari sparse matriks:

```
(113, 1232)
(1232, 113)
[[0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 ...
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]]
```

Gambar 2. Sparse Matriks

Sparse matriks ini bertujuan untuk memprediksi nilai nol agar dapat melihat tingkat kekosongan pada matriks. Dari data yang digunakan, diperoleh tingkat kekosongan matriks sebesar 98.67%. artinya dalam data ini aman digunakan[30].

Mengubah data menjadi bentuk matriks, menjadi tahapan terakhir pada *pre-processing* data, tujuannya untuk mempersiapkan data bagi model rekomendasi. Data dimasukkan kedalam mode `pivot`, dimana kolom *userId* digunakan sebagai parameter *index*, dan *Idhotel* digunakan sebagai parameter *columns*. Sedangkan kolom *ratings* akan digunakan sebagai parameter *value*.

Tabel 5.
Mengubah data menjadi bentuk matriks

locati on	Ente bbe	Fort Port al	Fort Prot al	Jinja	Kab ale	Kaja nsi	Kala ngal a	Kam pala
userId								
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.0
3	0.0	0.0	0.0	10.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

C. Modelling Data

Model yang digunakan pada tahap ini, adalah mengimplementasi faktorisasi matriks menggunakan *Stochastic Gradient descent* (SGD) pada sistem rekomendasi hotel. Untuk membangun model faktorisasi matriks, dibutuhkan beberapa masukan seperti, matriks peringkat pengguna-item `R`, jumlah faktor laten `K`, `learning\_rate`, parameter regulasi `reg\_param`, dan jumlah iterasi `epochs`. Kemudian, dilakukan proses pelatihan model menggunakan SGD melalui sejumlah iterasi sebanyak 300 putaran. Nilai faktor laten P dan Q akan terus diperbarui secara bertahap selama proses iterasi-iterasi tersebut.

Pada setiap iterasi, sistem akan menghitung error dengan membandingkan nilai peringkat dari data dengan nilai peringkat yang diprediksi menggunakan faktor laten P dan Q saat itu. Selanjutnya, nilai P dan Q pada matriks akan diperbarui dengan

mempertimbangkan error tersebut, dimana learning rate yang menentukan kecepatan pembaruan sebesar 0.001, serta parameter regulasi yang berfungsi untuk mencegah terjadinya *overfitting*, yang dalam penelitian ini parameter regulasinya sebesar 0.001.

Setelah proses pelatihan model selesai, model dapat digunakan untuk membuat prediksi peringkat baru. Dimana prediksi peringkat dapat dihasilkan dari perkalian faktor laten P dan Q yang sudah ditransposisi.

D. Evaluasi Model

Untuk mengevaluasi kinerja model yang diusulkan, kami menggunakan beberapa matriks evaluasi yang populer, yaitu *Root Mean Squared Error* (RMSE), *Mean Squared Error* (MSE), dan *Mean Absolute Error* (MAE). Dimana, hasil penelitian ini menunjukkan bahwa metode faktorisasi matriks dengan SGD dapat mencapai nilai RMSE sebesar 0.370312, nilai MSE sebesar 0.137131, dan nilai MAE sebesar 0.089932. Berdasarkan nilai-nilai ini, dapat diartikan bahwa metode yang diusulkan mampu memberikan akurasi prediksi yang sangat baik dalam memprediksi peringkat hotel.

RMSE: 0.3703129583136226
MSE: 0.13713168709498677
MAE: 0.08993280885258523

Gambar 3. Hasil evaluasi model.

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat kami simpulkan bahwa metode faktorisasi matriks menggunakan *Stochastic Gradient Descent* (SGD) bisa digunakan untuk optimasi sistem rekomendasi hotel. Hal ini dapat dilihat dari hasil evaluasi yang menunjukkan nilai kesalahan prediksi seperti, nilai RMSE sebesar 0.370312, nilai MSE sebesar 0.137131, dan nilai MAE sebesar 0.089932. Dimana, dari ketiga nilai kesalahan prediksi ini menunjukkan hasil yang cukup rendah, artinya kinerja model sangat baik dalam memberikan rekomendasi hotel sesuai dengan preferensi pengguna. metode ini berpotensi meningkatkan kepuasan pengguna dengan menyediakan rekomendasi hotel yang personalia dan sesuai dengan keinginan mereka. Selain itu, penelitian ini juga dapat membuka peluang untuk pengembangan lebih lanjut guna meningkatkan akurasi dan efisiensi sistem rekomendasi hotel di masa mendatang.

REFERENCES

- [1] S. Seimahuira, "Implementasi datamining dalam menentukan destinasi unggulan berdasarkan online reviews tripadvisor menggunakan algoritma K-Means," *Technologia: Jurnal Ilmiah*, vol. 12, no. 1, pp. 53–58, 2021.

- [2] R. Sarudin, dan Achmad Ismail, D. Tanggal Bulan Tahun, and D. Tanggal Bulan Tahun, "ANALISIS ONLINE REVIEW TRIPADVISOR.COM TERHADAP MINAT PEMBELIAN PRODUK JASA AKOMODASI DI HOTEL MANHATTAN TRIPADVISOR.COM REVIEW ONLINE ANALYSIS ON THE INTEREST OF BUYING ACCOMMODATION SERVICES IN HOTEL MANHATTAN," vol. 7, no. 4, pp. 33–43, 2021, doi: 10.30813/v7i1.2634.
- [3] V. Jap, J. Hartanto, and S. Wijaya, "ANALISIS KEPUASAN DAN KETIDAKPUASAN TAMU HOTEL NOVOTEL: PENDEKATAN TEXT MINING ATAS ULASAN DARING PADA SITUS WEB TRIPADVISOR," *Jurnal Manajemen Perhotelan*, vol. 9, no. 2, pp. 71–80, Sep. 2023, doi: 10.9744/jmp.9.2.71-80.
- [4] P. D. A. Prastika and B. Suharto, "Analisis Efisiensi Reservasi Kamar Hotel Secara Online dari Prespektif Sistem Digital," *Jurnal Pendidikan Tambusai*, vol. 7, no. 3, pp. 13807–13817, 2023.
- [5] A. Simorangkir, P. I. Sihite, C. L. Kiareni, R. Priskila, and V. H. Pranatawijaya, "PEMODELAN CHATBOT REKOMENDASI HOTEL DENGAN MENGGUNAKAN NATURAL LANGUAGE PROCESSING," *Simtek: jurnal sistem informasi dan teknik komputer*, vol. 9, no. 1, pp. 46–50, 2024.
- [6] R. H. Singh, S. Maurya, T. Tripathi, T. Narula, and G. Srivastav, "Movie Recommendation System using Cosine Similarity and KNN," *Int J Eng Adv Technol*, 2020, doi: 10.35940/ijeat.E9666.069520.
- [7] S. C. Aishwavya, "OPTIMASI MODEL SISTEM REKOMENDASI FILM DENGAN NEURAL NETWORK (STUDI KASUS: PLATFORM LETTERBOXD)," 2022.
- [8] Y. Pratama, R. A. Sianturi, D. Chandra, K. Lumbantoruan, and I. T. Tampubolon, "Restricted Boltzmann Machine and Matrix Factorization-Alternating Square Algorithm for Development Tourist Recommendation System," in *Journal of Physics: Conference Series*, Institute of Physics, 2022, doi: 10.1088/1742-6596/2394/1/012004.
- [9] X. Luo, Y. Zhou, Z. Liu, and M. Zhou, "Fast and accurate non-negative latent factor analysis of high-dimensional and sparse matrices in recommender systems," *IEEE Trans Knowl Data Eng*, vol. 35, no. 4, pp. 3897–3911, 2021.
- [10] Z. Munawar, N. I. Putri, and D. Z. Musadad, "Meningkatkan Rekomendasi Menggunakan Algoritma Perbedaan Topik," *J-SIKA| Jurnal Sistem Informasi Karya Anak Bangsa*, vol. 2, no. 02, pp. 17–26, 2020.
- [11] A. Z. Nisa, O. N. Pratiwi, and R. Y. Fa'rifah, "Klasifikasi Soal Sejarah Tingkat SMA Berdasarkan Level Kognitif Revised Bloom's Taxonomy Menggunakan Metode Stochastic Gradient Descent," *eProceedings of Engineering*, vol. 10, no. 2, 2023.
- [12] N. R. Madjid, Y. Vitriani, E. Haerani, and F. Kurnia, "Sistem Rekomendasi Hotel Di Provinsi Riau Dengan Metode AHP dan SAW," *Media Online*, vol. 3, no. 6, pp. 945–956, 2023, doi: 10.30865/klik.v3i6.931.
- [13] M. Noer and F. Hidayat, "PENENTUAN GIZI ANAK MENGGUNAKAN KOMPARASI METODE C4.5 DAN K-NEAREST NEIGHBOR (KNN)," 2020.
- [14] B. Joseph, M. Kasozi, M. Bogere, and C. Nabuyaka, "Hotel Reviews on Tripadvisor." Accessed: Mar. 26, 2024. [Online]. Available: <https://www.kaggle.com/datasets/josephbrianmkasozi/review-of-ugandan-hotels-as-of-2022>
- [15] J. E. Prayogo, A. Suharso, and A. Rizal, "Analisis Perbandingan Model Matrix Factorization dan K-Nearest Neighbor dalam Mesin Rekomendasi Collaborative Berbasis Prediksi Rating," *Jurnal Informatika Universitas Pamulang*, vol. 5, no. 4, p. 506, Dec. 2021, doi: 10.32493/informatika.v5i4.7379.
- [16] L. Uswatun Hasanah *et al.*, "ANALISIS PREDIKSI HARGA RUMAH DI JABODETABEK MENGGUNAKAN MULTIPLE LINEAR REGRESSION," *JIK*, vol. 7, no. 2, 2023.
- [17] A. Tholib and M. Kom, "Implementasi Algoritma Machine Learning Berbasis Web dengan Framework Streamlit," *Pustaka Nurja*, 2023.
- [18] V. Virdausih Putri *et al.*, "DETEKSI KAGGLE BOT ACCOUNT MENGGUNAKAN DEEP NEURAL NETWORKS," 2023.
- [19] K. L. Du, M. N. S. Swamy, Z. Q. Wang, and W. H. Mow, "Matrix Factorization Techniques in Machine Learning, Signal Processing, and Statistics," *Mathematics*, vol. 11, no. 12, MDPI, Jun. 01, 2023, doi: 10.3390/math11122674.
- [20] Chen Denise, "Recommender System - Matrix Factorization." Accessed: Mar. 28, 2024. [Online]. Available: <https://towardsdatascience.com/recommendation-system-matrix-factorization-d61978660b4b>
- [21] kim Haneul, "Matrix Factorization Part 1: Understanding all the processes and how Stochastic Gradient Descent step is derived." Accessed: Mar. 28, 2024. [Online]. Available: <https://haneulkim.medium.com/matrix-factorization-part-i-from-derivation-of-stochastic-gradient-descent-step-to-learning-17a7d8975965>
- [22] R. Umar, I. Riadi, and P. Purwono, "Klasifikasi Kinerja Programmer pada Aktivitas Media Sosial dengan Metode Stochastic Gradient Descent," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 5, no. 2, p. 55, 2020, doi: 10.31328/jointecs.v5i2.1324, 2020.
- [23] C. Leo, "Stochastic Gradient Descent: Math and Python Code." Accessed: Mar. 28, 2024. [Online]. Available: <https://towardsdatascience.com/stochastic-gradient-descent-math-and-python-code-35b5e66d6f79>
- [24] P. S. Rao, T. V. M. Rao, S. Kurumalla, and B. Prakash, "Matrix Factorization Based Recommendation System using Hybrid Optimization Technique," *EAI Endorsed Transactions on Energy Web*, vol. 8, no. 35, pp. 1–7, 2021, doi: 10.4108/eai.19-2-2021.168725.
- [25] B. Walek and V. Fojtik, "A hybrid recommender system for recommending relevant movies using an expert system," *Expert Syst Appl*, vol. 158, p. 113452, 2020.

- [26] S.-I. Kuan, J. Kim, O.-H. Kwon, and H.-J. Song, "Canopy K-means Combined Collaborative Filtering Using RMSE-minimization," in *2022 IEEE International Conference on Big Data and Smart Computing (BigComp)*, IEEE, 2022, pp. 31–34.
- [27] Z. Arif, N. A. S. Adhi, and A. Muttaqin, "Penerapan Jaringan Syaraf Tiruan dan Certainty Factor untuk Peramalan Penjualan Sepeda Motor," *REMIK: Riset dan E-Jurnal Manajemen Informatika Komputer*, vol. 7, no. 3, pp. 1523–1533, 2023.
- [28] Z. Munawar, N. I. Putri, and D. Z. Musadad, "Meningkatkan Rekomendasi Menggunakan Algoritma Perbedaan Topik," *Jurnal Sistem Informasi*, vol. 01, 2020.
- [29] D. Yulion, B. N. Yahya, and V. J. L. Engel, "Building a robust transition matrix using causal matrix for route recommendation," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 768–775. doi: 10.1016/j.procs.2021.12.199.
- [30] H. A. Afif, "Membuat Recommendation Engine, menggunakan faktorisasi matriks, untuk data implicit (Part 1)," medium.com. Accessed: Mar. 28, 2024. [Online]. Available: <https://medium.com/@hafizhan.aliady/membuat-recommendation-engine-menggunakan-faktorisasi-matriks-optimasi-als-alternating-least-b2823d88949f>