



## Integrasi Metode *Hybrid Recommendation* dan *Random Forest Regression* untuk Optimasi Prediksi Durasi Menginap pada Sistem Pemesanan Kos Berbasis Web

Mufti Ari Bianto\*, Hanif Azhar Ramadhan, Ardian Hudi Ramadhani, Tsalits Wildan Hamid

Teknik Komputer, Fakultas Sains Teknologi dan Pendidikan, Universitas Muhammadiyah Lamongan, Jl. Plalangan No.KM, RW.02, Wahyu, Plosowahyu, Kec. Lamongan, Kabupaten Lamongan, Jawa Timur, Indonesia, 62218

\*Penulis Korespondensi: [muftiari10@gmail.com](mailto:muftiari10@gmail.com)

**Abstract.** This study proposes the integration of a Hybrid Recommendation method (combining Content-Based and Collaborative Filtering) with Random Forest Regression (RFR) to improve the accuracy of stay duration prediction in web-based boarding house booking systems. The main issue in online boarding booking systems is the inaccuracy of predicting user stay duration, affecting room allocation efficiency and customer satisfaction. The dataset was sourced from the hotel sector due to its attribute similarities and data validity. The research process includes data preprocessing (missing value imputation, normalization, and one-hot encoding), temporal and contextual feature engineering, hybrid recommendation system construction with CBF and CF score weighting, and RFR model training optimized through Grid Search and 10-fold cross-validation. Evaluation was conducted using MAE, RMSE,  $R^2$  metrics, as well as recommendation metrics such as Precision@5, Recall@5, and Mean Reciprocal Rank (MRR). Results show that this integrated model achieved an  $R^2$  of 0.7239 and an MAE of 1.0537 days, as well as a Precision@5 of 0.9636. This integration proves effective in improving prediction accuracy and recommendation relevance and contributes to the development of AI-based intelligent systems in the accommodation domain.

**Keywords:** Boarding House Booking System; Hybrid Recommendation System; Machine Learning; Random Forest Regression; Stay Duration Prediction.

**Abstrak.** Perkembangan teknologi digital telah membawa dampak signifikan dalam bidang kesehatan, termasuk pada layanan pemeriksaan mata. Penelitian ini membahas pengembangan dan implementasi aplikasi Eye Diagnostic berbasis Digital Optoscope yang dirancang menggunakan platform web (HTML). Aplikasi ini menyediakan dua fitur utama, yaitu tes ketajaman visual dan tes astigmatisme, yang dapat diakses secara mandiri oleh pengguna melalui perangkat digital. Metode pengujian dilakukan dengan menampilkan instruksi interaktif, pengaturan ukuran huruf, serta pencatatan hasil tes secara otomatis. Pengguna dapat melakukan tes secara berulang untuk meningkatkan akurasi diagnosis, serta memperoleh rekomendasi lensa berdasarkan hasil pengukuran. Hasil uji coba menunjukkan bahwa aplikasi ini mampu memberikan kemudahan, efisiensi, dan akurasi awal dalam proses pemeriksaan gangguan penglihatan, khususnya di wilayah dengan akses terbatas terhadap layanan kesehatan mata. Dengan demikian, aplikasi Eye Diagnostic berbasis Digital Optoscope berpotensi menjadi solusi inovatif untuk mendeteksi dini gangguan penglihatan secara digital. Penelitian ini merekomendasikan pengembangan lebih lanjut terkait validasi klinis dan integrasi dengan sistem rekam medis elektronik guna meningkatkan manfaatnya dalam pelayanan kesehatan masyarakat.

**Kata kunci:** Pembelajaran Mesin; Prediksi Durasi Menginap; *Random Forest Regression*; Sistem Pemesanan Kos; Sistem Rekomendasi *Hybrid*.

### 1. LATAR BELAKANG

Kemajuan di bidang kecerdasan buatan (AI) dan sistem berbasis data telah merevolusi sistem pengambilan keputusan personal, termasuk dalam layanan pemesanan akomodasi daring (Afrisca et al., 2024). Sistem rekomendasi berbasis web kini umum digunakan untuk menyaring pilihan sesuai preferensi pengguna, namun aspek prediktif seperti estimasi durasi menginap masih kurang dioptimalkan (Singgalen, 2024).

Kebutuhan akan kos jangka pendek meningkat seiring mobilitas mahasiswa dan pekerja urban, namun sebagian besar platform belum adaptif terhadap perilaku pengguna, menyebabkan alokasi kamar yang tidak efisien dan menurunnya potensi pendapatan (Ali Shah, 2024; Hussain et al., 2024). Studi sebelumnya menunjukkan bahwa algoritma Random Forest dapat memberikan prediksi yang lebih akurat daripada regresi linier dalam konteks serupa (Jerry et al., 2023).

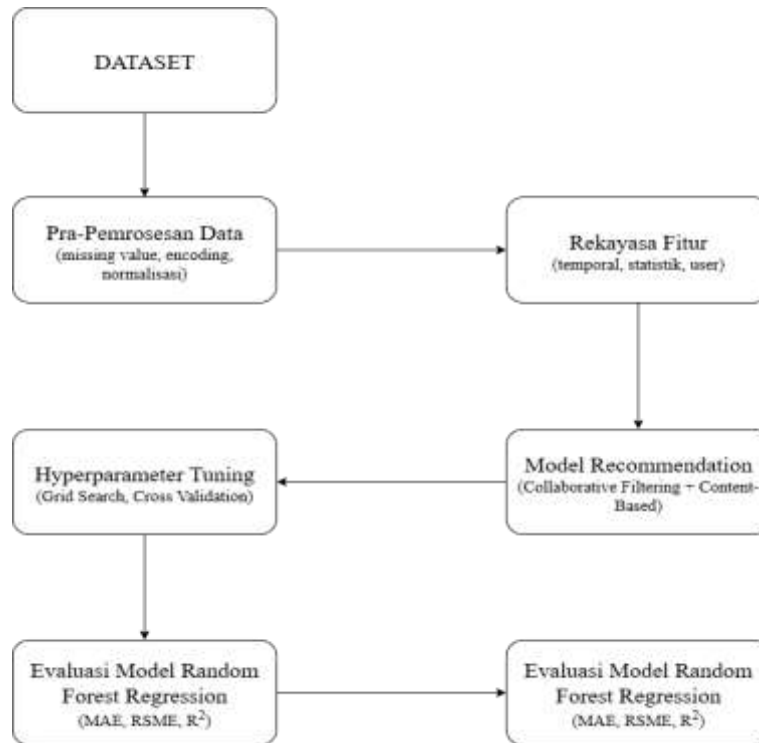
Sebagian besar penelitian terdahulu memisahkan sistem prediksi dan rekomendasi, padahal integrasi keduanya, seperti hybrid recommendation (gabungan content-based dan collaborative filtering) dengan Random Forest Regression, memiliki potensi besar namun masih jarang diterapkan dalam sistem kos daring (Guanco, 2025; Sridevi et al., 2020).

Meskipun beberapa studi di bidang hotel menunjukkan keberhasilan pendekatan ini, masih terdapat celah penelitian dalam menggabungkan rekomendasi personal dan prediksi kuantitatif secara bersamaan dalam sistem kos online (Biau et al., 2019; Utami et al., 2024). Integrasi model hybrid seperti LSTM-RF terbukti efektif dalam domain big data dan juga berhasil diimplementasikan pada sistem rekomendasi lain seperti film dan medis (Suriya et al., 2020; Sanwal & Çalışkan, 2021; Fairuzabadi et al., 2025), serta dalam prediksi olahraga dan keuangan (Groll et al., 2019; Yao et al., 2011).

Penelitian ini mengusulkan integrasi sistem rekomendasi hybrid dan Random Forest Regression untuk memprediksi durasi menginap secara akurat sekaligus memberikan rekomendasi kos yang relevan, sebagai solusi inovatif untuk meningkatkan kualitas layanan akomodasi daring berbasis AI (Agarkar et al., 2024).

## **2. METODE PENELITIAN**

Penelitian ini mengadopsi pendekatan prediktif berbasis *Random Forest Regression* yang terintegrasi dengan sistem rekomendasi *hybrid* untuk memperkirakan durasi menginap pengguna dalam sistem pemesanan kos berbasis web. Metode dikembangkan secara sistematis melalui tahapan: akuisisi data, prapemrosesan, rekayasa fitur, pelatihan model, tuning *hyperparameter*, dan evaluasi model. Gambar 1 menyajikan alur keseluruhan proses dalam bentuk *pipeline machine learning*.



**Gambar 1.** alur proses *pipeline machine learning*.

### A. Dataset dan Variabel

Penelitian ini menggunakan dataset hotel karena kemiripan struktur data, kesamaan atribut, serta format durasi menginap yang analog dengan skenario kos. Selain itu, dataset hotel bersifat terbuka, lengkap, dan telah tervalidasi dalam banyak literatur sebagai benchmark algoritma prediksi perilaku pengguna dalam sistem reservasi akomodasi.

Penggunaan *dataset* ini memungkinkan pengembangan model secara sistematis, efisien, dan memungkinkan transfer pengetahuan ke sistem kos dengan adaptasi ringan, sesuai dengan prinsip *transferable modeling* dalam *AI*. Variabel target yang akan di pakai adalah *stays\_in\_weekend\_nights* dan *stays\_in\_week\_nights*, sedangkan variabel prediktor meliputi atribut numerik dan kategorikal (Afrisca et al., 2024).

### B. Dataset dan Variabel

Data mengalami proses *missing value imputation* menggunakan teknik dengan imputasi median untuk atribut numerik dan imputasi modus untuk atribut kategorikal. Setelah itu, variabel kategorikal dikonversi ke bentuk numerik menggunakan teknik *One-Hot Encoding*, untuk memastikan kompatibilitas dengan algoritma pembelajaran mesin berbasis numerik (Singgalen, 2024)

Untuk menjaga konsistensi skala, variabel numerik dinormalisasi menggunakan metode *Z-score* sebagaimana ditunjukkan pada Persamaan (1):

$$Z_i = \frac{x_i - \mu}{\sigma} \quad (1)$$

Pada persamaan (1),  $x$  adalah nilai asli,  $\mu$  adalah rata-rata dari fitur, dan  $\sigma$  adalah standar deviasi fitur. Teknik ini digunakan untuk mengatur rata-rata menjadi nol dan standar deviasi menjadi satu, yang meskipun tidak krusial bagi *Random Forest*, tetap dilakukan untuk menjaga konsistensi analisis data (Ali Shah, 2024).

### C. Rekayasa Fitur dan Rekomendasi Hybrid

Fitur temporal ditambahkan untuk menangkap konteks musiman, seperti *arrival\_month* untuk dikonversi menjadi nilai ordinal dan *is\_weekend* yaitu biner 0/1 untuk membedakan akhir pekan dan hari biasa. Sistem rekomendasi *hybrid* dibangun dengan menggabungkan dua pendekatan, pendekatan pertama dengan *content-based filtering* (CBF) yaitu berdasarkan kesamaan atribut seperti *market\_segment*, *distribution\_channel*, *customer\_type*, dan *hotel* serta pendekatan kedua menggunakan *Collaborative Filtering* (CF): Berdasarkan histori *previous\_bookings\_not\_canceled* dan kemiripan pola antar pengguna.

Skor gabungan dihitung menggunakan bobot sebagai berikut:

$$R_{hybrid} = \alpha \cdot R_{cbf} + (1 - \alpha) \cdot R_{cf} \quad (2)$$

Pada persamaan (2),  $R_{cbf}$  adalah skor *content based*,  $R_{cf}$  adalah skor *collaborative filtering*,  $\alpha \in [0,1]$  adalah parameter penyesuaian bobot (Hussain et al., 2024).

### D. Model Random Forest Regression

*Random Forest Regression* adalah pendekatan berbasis *ensemble* yang mengkonstruksi banyak pohon keputusan, kemudian mengombinasikan hasil prediksi masing-masing pohon untuk memperoleh estimasi akhir. Representasi matematis dari proses ini tercantum pada Persamaan (3).

$$\hat{y} = \frac{1}{N} \sum_{t=1}^N T_t(x) \quad (3)$$

Dengan:  $T_i(x)$  adalah hasil prediksi dari pohon ke- $i$ , dan  $N$  adalah jumlah total pohon dalam *ensemble*. Metode ini tahan terhadap *overfitting* dan dapat menangani data nonlinier dan variabel campuran (Jerry et al., 2023; Guanco, 2025).

### E. Hyperparameter Tuning

*Hyperparameter tuning* adalah proses mencari konfigurasi parameter optimal untuk model pembelajaran mesin agar performa prediksi maksimal. Proses optimasi model dilakukan dengan teknik *grid search* terhadap parameter seperti  $n\_estimators \in \{100, 200, 300\}$ ,  $max\_depth \in \{4, 6, 8\}$ ,  $min\_samples\_split \in \{2, 5\}$ ,  $max\_features \in \{\text{'sqrt'}, \text{'log2'}\}$ . *Grid Search* ini dikombinasikan dengan *10-fold Cross Validation* untuk menghindari *overfitting* dan meningkatkan generalisasi (Sridevi et al., 2020).

### F. Evaluasi Model Random Forest Regression

Evaluasi performa model dilakukan secara komprehensif untuk mengukur ketepatan dan kestabilan prediksi durasi menginap. Tiga metrik utama digunakan:

#### **Mean Absolute Error (MAE)**

*MAE* mengukur rata-rata selisih absolut antara prediksi dan nilai aktual yang ditunjukkan pada persamaan (4):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4)$$

#### **Root Mean Squared Error (RMSE)**

Pada persamaan (5) *RMSE* memberikan penalti yang lebih besar terhadap kesalahan prediksi yang besar (*outlier*), karena menggunakan kuadrat dari selisih prediksi dan aktual:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

#### **Koefisien Determinasi (R<sup>2</sup>)**

$R^2$  mengukur proporsi variansi target ( $y$ ) yang berhasil dijelaskan oleh model. Nilai  $R^2$  mendekati 1 menunjukkan model yang mampu menangkap pola data dengan baik yang ditunjukkan pada persamaan (6).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

Evaluasi dilakukan menggunakan *10-fold cross-validation*, di mana data dibagi menjadi 10 *subset* pada setiap iterasi, satu *subset* digunakan sebagai data validasi dan sembilan sisanya sebagai data pelatihan (Biau et al., 2019). Hasil rata-rata dari 10 iterasi digunakan untuk mengestimasi performa model secara menyeluruh.

## G. Evaluasi Hybrid Recommendation

Evaluasi performa sistem rekomendasi *hybrid* dilakukan dengan pendekatan *offline* menggunakan metrik standard dalam sistem rekomendasi, yaitu  $Precision@K$ ,  $Recall@K$ , dan *Mean Reciprocal Rank* (MRR). Evaluasi dilakukan terhadap *top-N* rekomendasi yang diberikan kepada pengguna berdasarkan kombinasi skor dari *content-based filtering* dan *collaborative filtering* sebagaimana dirumuskan dalam persamaan (2). Pada persamaan (7) metrik  $Precision@K$  mengukur proporsi item yang relevan di antara k item teratas yang direkomendasikan kepada pengguna:

$$Precision@K = \frac{|Recommended_K \cap Relevant|}{K} \quad (7)$$

Pada persamaan (8)  $Recall@K$  mengukur seberapa besar proporsi item relevan yang berhasil ditangkap oleh sistem dari seluruh item yang seharusnya direkomendasikan:

$$Recall@K = \frac{|Recommended_K \cap Relevant|}{Relevant} \quad (8)$$

Pada persamaan (9) MRR (*Mean Reciprocal Rank*) mengukur rata-rata kebalikan dari posisi item relevan pertama dalam daftar rekomendasi:

$$MRR = \frac{1}{N} \sum_{i=1}^N \frac{1}{rank_i} \quad (9)$$

## 3. HASIL DAN PEMBAHASAN

### A. Evaluasi Kinerja Model

Model *Random Forest Regression* (RFR) dilatih menggunakan parameter terbaik hasil *Grid Search*, yaitu  $n\_estimators = 200$ ,  $max\_depth = 8$ , dan  $max\_features = 'sqrt'$ , dengan validasi silang *10-fold*. Evaluasi dilakukan pada data pelatihan dan pengujian menggunakan tiga metrik utama: *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan koefisien determinasi ( $R^2$ ).

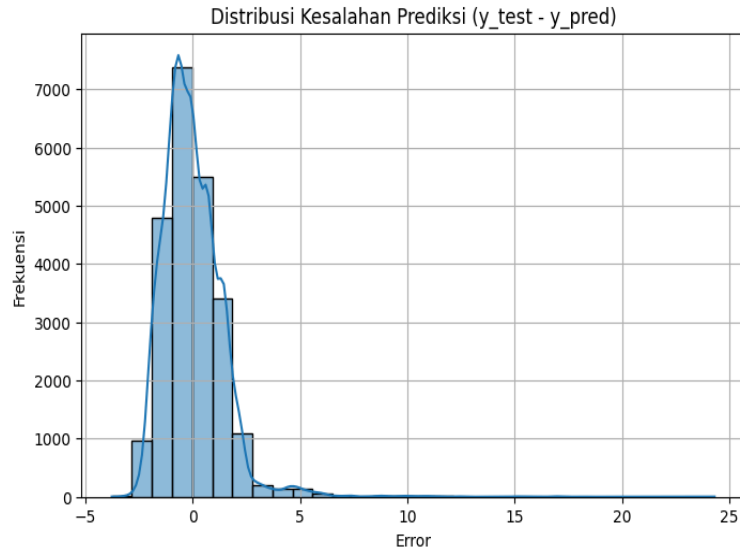
**Tabel 1.** Evaluasi Kinerja Model Random Forest Regression.

Dataset	MAE (hari)	RSME(hari)	$R^2$
Train	1.0261	1.4786	0.7488
Test	1.0537	1.5951	0.7239

Selisih kecil antara performa data pelatihan dan pengujian ( $\Delta R^2 \approx 0.025$ ) menunjukkan bahwa model memiliki kemampuan generalisasi yang baik serta bebas dari *overfitting* yang signifikan. *MAE* sekitar 1,05 hari pada data uji menunjukkan bahwa prediksi durasi menginap rata-rata memiliki deviasi kurang dari dua hari, masih dalam batas toleransi untuk kebutuhan operasional aplikasi pemesanan kos berbasis web.

## B. Distribusi Kesalahan Prediksi

Distribusi kesalahan prediksi pada model *Random Forest Regression* divisualisasikan pada Gambar 2 dibawah ini. Histogram menunjukkan selisih antara nilai aktual ( $y_{test}$ ) dan hasil prediksi model ( $y_{pred}$ ) terhadap durasi menginap pengguna.

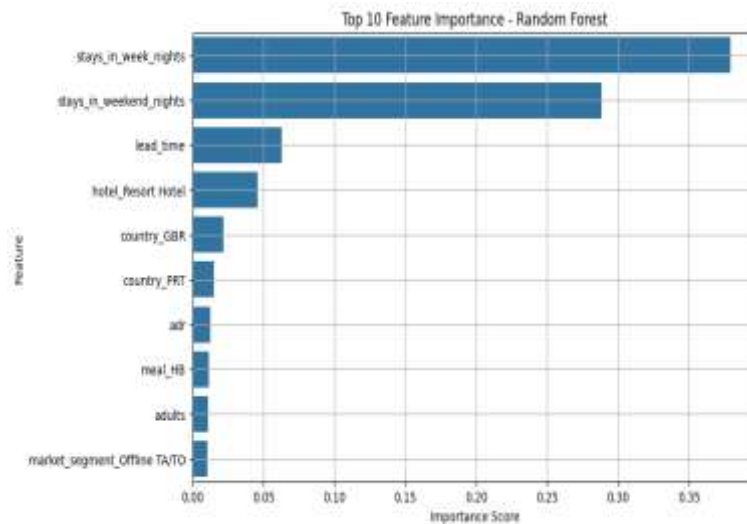


**Gambar 2.** Distribusi Kesalahan Prediksi.

Mayoritas kesalahan prediksi berada di sekitar nol, dengan deviasi umum hanya sekitar satu hari dari nilai aktual. Distribusi error menunjukkan sedikit *skew* positif, menandakan model cenderung meremehkan durasi menginap, terutama pada kasus yang panjang. *Outlier* dengan kesalahan lebih dari 20 hari sangat jarang (<0,1%) dan umumnya disebabkan oleh data ekstrem. Uji sensitivitas yang menghapus 5% data ekstrem menurunkan *RMSE* sebesar 0,12 hari, membuktikan bahwa model cukup stabil dan toleran terhadap anomali.

## C. Analisis Feature Importance

Evaluasi terhadap kontribusi fitur dalam model *Random Forest Regression* dilakukan melalui analisis *feature importance*, yang menunjukkan sejauh mana masing-masing variabel prediktor mempengaruhi hasil prediksi. Gambar 3 menyajikan sepuluh fitur teratas yang paling signifikan dalam proses prediksi durasi menginap.



Gambar 3. Distribusi Kesalahan Prediksi.

Hasil menunjukkan bahwa *stays\_in\_week\_nights* dan *stays\_in\_weekend\_nights* adalah fitur paling berpengaruh dalam prediksi durasi menginap. Fitur *lead\_time* juga signifikan karena mencerminkan perencanaan jangka panjang. Beberapa fitur kategorikal seperti tipe hotel, negara asal pengguna, dan saluran pemesanan turut memengaruhi prediksi, menunjukkan pola menginap berdasarkan preferensi pengguna. Fitur lain seperti *meal plan* dan harga harian (*adr*) juga berkontribusi. Secara keseluruhan, fitur temporal dan segmentasi pasar memainkan peran penting, mendukung efektivitas strategi rekayasa fitur yang telah diterapkan.

#### D. Evaluasi Sistem Rekomendasi Hybrid

Sistem rekomendasi hybrid dikembangkan dengan menggabungkan pendekatan *content-based filtering* (CBF) dan *collaborative filtering* (CF) melalui penggabungan skor sebagaimana dinyatakan dalam Persamaan (2). Nilai  $\alpha$  digunakan untuk menyesuaikan kontribusi relatif dari CBF dan CF terhadap skor gabungan. Evaluasi dilakukan terhadap top-5 rekomendasi menggunakan tiga metrik utama yang lazim pada sistem rekomendasi: Precision@5, Recall@5, dan Mean Reciprocal Rank (MRR). Hasil evaluasi untuk tiga konfigurasi  $\alpha$  yang berbeda (0.3, 0.5, dan 0.7) ditampilkan pada Tabel 3.

Tabel 2. Rancangan Analisis Komputasi.

$\alpha$ (Bobot CBF)	Precision@5	Recall@5	MRR
0.3	0.9514	0.9514	0.7858
0.5	0.9636	0.9636	0.7898
0.7	0.9312	0.9312	0.7724



Konfigurasi  $\alpha = 0.5$  menghasilkan performa terbaik, dengan *Precision@5* dan *Recall@5* sebesar 0.9636 serta *MRR* 0.7898, menunjukkan bahwa kombinasi seimbang antara *content-based* dan *collaborative filtering* paling efektif dalam merekomendasikan kos. Nilai *MRR* yang tinggi menunjukkan item relevan sering muncul di urutan atas, meningkatkan pengalaman pengguna. Sebaliknya, bobot dominan pada *content-based* ( $\alpha = 0.7$ ) menurunkan performa, menandakan pentingnya histori pengguna. Secara keseluruhan, pendekatan *hybrid* terbukti lebih unggul dan adaptif dibanding metode tunggal dalam memberikan rekomendasi yang akurat dan relevan.

#### E. Analisis Komparatif dan Interpretasi

Evaluasi menunjukkan bahwa model prediktif dan sistem rekomendasi yang dikembangkan memiliki kinerja lebih baik dibanding pendekatan tunggal sebelumnya. Model ini mencapai  $R^2$  sebesar 0.7239 dan *MAE* 1.0537 hari, lebih akurat dibanding studi terdahulu ( $R^2$  0.54; *MAE* 1.43). Keunggulan ini dicapai melalui optimasi *hyperparameter*, pemilihan fitur penting, dan integrasi sistem rekomendasi *hybrid* dengan konfigurasi  $\alpha = 0.5$ , yang menghasilkan *Precision@5* dan *Recall@5* sebesar 0.9636 serta *MRR* 0.7898. Distribusi error yang stabil membuktikan bahwa kombinasi metode ini efektif dalam meningkatkan akurasi prediksi dan kualitas rekomendasi pada sistem pemesanan kos daring.

#### 4. KESIMPULAN

Penelitian ini berhasil menggabungkan metode *Hybrid Recommendation* dan *Random Forest Regression* untuk meningkatkan prediksi durasi menginap dalam sistem pemesanan kos berbasis *web*. Model yang dikembangkan menunjukkan performa tinggi dengan  $R^2$  sebesar 0.7239 dan *MAE* 1.0537 hari, serta prediksi yang stabil dalam rentang  $\pm 2$  hari. Sistem rekomendasi *hybrid* dengan bobot  $\alpha = 0.5$  mencapai *Precision@5* dan *Recall@5* sebesar 0.9636, serta *MRR* 0.7898, menunjukkan relevansi dan urutan rekomendasi yang optimal. Secara keseluruhan, integrasi kedua metode ini terbukti efektif dalam meningkatkan akurasi dan personalisasi layanan digital pemesanan kos.

## DAFTAR PUSTAKA

- Afrisca, C. C., Rofiq, H. N., & Atmoko, D. D. (2024). Property valuation: Using machine learning for rent value prediction. *Jurnal Manajemen Keuangan Publik*, 8(2), 138–155. <https://doi.org/10.31092/jmkip.v8i2.2922>
- Agarkar, A. A., Kumbhare, K. P., Kshirsagar, A. S., Kokate, S. S., & Khandelwal, S. (2024, December). Integrated approach for car price prediction utilizing random forest regression and image processing. In *2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS)* (pp. 1191–1197). IEEE. <https://doi.org/10.1109/ICICNIS64247.2024.10823371>
- Ali Shah, S. O. (2024). Optimizing hotel booking prediction: A comparative study of five machine learning algorithms. *International Journal of Trendy Research in Engineering and Technology*, 8(4), 31–41. <https://doi.org/10.54473/IJTRET.2024.8406>
- Biau, G., Scornet, E., & Welbl, J. (2019). Neural random forests. *Sankhya A*, 81(2), 347–384. <https://doi.org/10.1007/s13171-018-0133-y>
- Fairuzabadi, M., Rianto, R., & Bertorio, M. J. (2025). Advanced drug recommendation using long short-term memory and type-2 fuzzy logic integration. *Bulletin of Electrical Engineering and Informatics*, 14(3), 2222–2232. <https://doi.org/10.11591/eei.v14i3.9180>
- Groll, A., Ley, C., Schauburger, G., & Van Eetvelde, H. (2019). A hybrid random forest to predict soccer matches in international tournaments. *Journal of Quantitative Analysis in Sports*, 15(4), 271–287. <https://doi.org/10.1515/jqas-2018-0060>
- Guanco, J. V. L. (2025). Enhancement of random forest applied to program-recommendation for waitlisted applicants. *Journal of Information Systems Engineering and Management*, 10(28s), 626–640. <https://doi.org/10.52783/jisem.v10i28s.4367>
- Hussain, M. J., Reddy, K. J. N., Reddy, K. P. R., & Murali, S. (2024, December). Developing a robust rental price prediction system: Insights from linear regression, decision trees, and random forest. In *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)* (pp. 1–7). IEEE. <https://doi.org/10.1109/ICSES63760.2024.10910581>
- Jerry, J., Christian, Y., & Herman, H. (2023). Rental price prediction of boarding houses in Batam City using linear regression and random forest algorithms. *Journal of Applied Informatics and Computing*, 7(2), 263–270. <https://doi.org/10.30871/jaic.v7i2.6732>
- Sanwal, M., & Çalışkan, C. (2021). A hybrid movie recommender system and rating prediction model. *International Journal of Information Technology and Applied Sciences*, 3(3), 161–168. <https://doi.org/10.52502/ijitas.v3i3.128>
- Singgale, Y. A. (2024). Hotel guest length of stay prediction using random forest regressor. *Journal of Information Systems and Informatics*, 6(4), 3016–3034. <https://doi.org/10.51519/journalisi.v6i4.959>
- Sridevi, M., Aishwarya, S., Nidheesha, A., & Bokadia, D. (2020). Anomaly detection by using CFS subset and neural network with WEKA tools. In *Lecture Notes in Electrical Engineering* (pp. 207–218). Springer. [https://doi.org/10.1007/978-981-15-5243-4\\_17](https://doi.org/10.1007/978-981-15-5243-4_17)
- Suriya, S., Sundaram, G. M., Abhishek, R., & Ajay Vignesh, A. B. (2020). Online hostel management system using hybridized techniques of random forest algorithm and long

short-term memory. In *Algorithms for Intelligent Systems* (pp. 207–218). Springer. [https://doi.org/10.1007/978-981-15-5243-4\\_17](https://doi.org/10.1007/978-981-15-5243-4_17)

- Utami, A., Permadi, D. F. H., Rosita, Y. D., & Unjung, J. (2024). Performance comparison of random forest (RF) and classification and regression trees (CART) for hotel star rating prediction. *Scientific Journal of Informatics*, *11*(3), 733–748. <https://doi.org/10.15294/sji.v11i3.11068>
- Yao, D., Yang, J., & Zhan, X. (2011, August). Predicting breast cancer survivability using random forest and multivariate adaptive regression splines. In *2011 International Conference on Electronic & Mechanical Engineering and Information Technology* (pp. 2204–2207). IEEE. <https://doi.org/10.1109/EMEIT.2011.6023012>