

PREDIKSI JUMLAH PENGUNJUNG PERPUSTAKAAN MENGUNAKAN ALGORITMA RANDOM FOREST

Haifan Ghani Raihan Santos, Fidi Supriadi, David Setiadi

Informatika, Universitas Sebelas April Sumedang

Jl Angkreng Situ No.19 Sumedang, Indonesia

220660121202@student.unsap.ac.id

ABSTRAK

Perpustakaan memiliki peran penting sebagai pusat literasi dan penyebaran informasi bagi masyarakat. Namun, jumlah pengunjung yang berfluktuasi setiap tahun sering kali menyulitkan pengelola dalam merencanakan kapasitas dan layanan secara optimal. Permasalahan ini dapat diatasi dengan pendekatan berbasis data melalui pemodelan prediktif. Penelitian ini bertujuan untuk memprediksi jumlah pengunjung perpustakaan di Kota Bandung pada tahun 2025 menggunakan algoritma Random Forest Regressor. Data yang digunakan berasal dari portal Open Data Bandung periode 2019–2024. Model dibangun dengan parameter $n_estimators=300$ dan $random_state=42$, serta dievaluasi menggunakan metrik MAE, RMSE, dan MAPE. Hasil evaluasi menunjukkan nilai MAPE sebesar 33,29%, yang tergolong cukup baik dan masih dapat diterima untuk kebutuhan manajemen serta perencanaan layanan perpustakaan. Berdasarkan hasil peramalan, total kunjungan tahun 2025 diperkirakan mencapai 21.525 dengan rata-rata 1.794 pengunjung per bulan. Hasil penelitian ini diharapkan dapat memperkuat strategi pelayanan dan pengelolaan sumber daya perpustakaan secara efisien berbasis data.

Kata kunci : Machine Learning, Random Forest, Prediksi, Perpustakaan, Kota Bandung

1. PENDAHULUAN

Perpustakaan memiliki peran penting sebagai pusat literasi dan penyebaran informasi bagi masyarakat. Sebagai lembaga pendidikan nonformal, perpustakaan tidak hanya berfungsi sebagai tempat penyimpanan bahan pustaka, tetapi juga menjadi ruang belajar dan sumber referensi yang mendukung kegiatan akademik. Jumlah pengunjung perpustakaan dapat menjadi indikator tingkat minat baca dan efektivitas layanan yang disediakan [1], [2]. Di Kota Bandung, tren kunjungan perpustakaan cenderung berfluktuasi dari tahun ke tahun akibat perubahan perilaku masyarakat, kebijakan pemerintah, serta perkembangan teknologi digital. Kondisi ini menuntut strategi pengelolaan berbasis data yang mampu mengidentifikasi pola kunjungan dan memberikan prediksi yang akurat sebagai dasar pengambilan keputusan.

Fluktuasi jumlah pengunjung perpustakaan tidak hanya menunjukkan perubahan perilaku masyarakat terhadap literasi, tetapi juga berdampak pada proses pengelolaan layanan perpustakaan. Ketidakpastian jumlah pengunjung dapat menyebabkan kesulitan dalam perencanaan kapasitas ruang baca, pengelolaan koleksi, serta penjadwalan tenaga layanan [3]. Selain itu, perencanaan program literasi dan kegiatan edukatif juga menjadi kurang optimal apabila tidak didukung oleh estimasi jumlah pengunjung yang akurat [4]. Oleh karena itu, diperlukan suatu metode prediksi yang mampu memperkirakan jumlah pengunjung secara lebih sistematis sehingga dapat membantu pengelola dalam menyusun strategi pelayanan yang lebih efektif dan efisien.

Berbagai penelitian terdahulu telah menerapkan algoritma machine learning untuk mendukung sistem prediksi di berbagai bidang. Salah satu algoritma yang

terbukti efektif adalah Random Forest, yang dikenal karena stabilitas dan kemampuannya menghindari *overfitting* [1], [5]. Algoritma ini telah digunakan dalam optimasi keamanan OTP [1], prediksi risiko obesitas [2], serta peningkatan akurasi klasifikasi berbasis atribut [6]. Selain itu, Random Forest juga diterapkan untuk deteksi anomali jaringan [7], prediksi mahasiswa baru [5], dan peramalan jumlah wisatawan di sektor pariwisata [8]. Dalam konteks lain, penelitian di bidang kesehatan menunjukkan keberhasilan model Random Forest dalam memprediksi distribusi dosis terapi kanker [9], sementara di bidang properti digunakan untuk memprediksi harga rumah dengan hasil akurasi tinggi [10]. Di lingkungan akademik dan literasi, algoritma yang sama juga digunakan untuk memprediksi pola peminjaman buku [3] serta kinerja akademik siswa [4], memperkuat bukti bahwa Random Forest merupakan metode yang fleksibel dan memiliki kinerja prediktif yang konsisten.

Sejalan dengan berbagai studi tersebut, penelitian ini difokuskan pada penerapan algoritma Random Forest untuk memprediksi jumlah pengunjung perpustakaan di Kota Bandung pada tahun 2025. Data yang digunakan berupa data historis kunjungan perpustakaan dari tahun 2019 hingga 2024. Permasalahan utama yang diangkat adalah bagaimana algoritma Random Forest dapat dimanfaatkan untuk menghasilkan prediksi jumlah pengunjung secara akurat guna membantu pengambil kebijakan dalam merencanakan strategi pelayanan, pengadaan fasilitas, serta pengelolaan sumber daya perpustakaan secara efektif dan efisien.

Penelitian ini menerapkan algoritma Random Forest sebagai metode utama dalam proses pemodelan prediksi. Tahapan penelitian meliputi pengumpulan dan pra-pemrosesan data, pembagian data latih dan uji,

pembangunan model, serta evaluasi performa. Proses evaluasi dilakukan dengan tiga metrik utama, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE) sebagai indikator pengukuran hasil model. Ketiga metrik tersebut digunakan untuk menilai tingkat akurasi dan stabilitas prediksi, sehingga hasil penelitian dapat menjadi acuan dalam penerapan teknologi machine learning untuk meningkatkan efektivitas pengelolaan layanan publik, khususnya di bidang perpustakaan daerah.

2. TINJAUAN PUSTAKA

Bagian ini membahas teori-teori dan hasil penelitian terdahulu yang menjadi dasar dalam pengembangan penelitian ini. Tinjauan pustaka mencakup konsep perpustakaan dan perilaku kunjungan, dasar teori machine learning, penjelasan mengenai algoritma Random Forest, serta penelitian relevan yang mendukung penerapan metode tersebut pada berbagai bidang.

2.1. Konsep Perpustakaan dan Kunjungan

Perpustakaan berperan penting dalam mendukung kegiatan literasi, pendidikan, dan penyebaran informasi di masyarakat. Sebagai pusat pengetahuan, perpustakaan kini tidak hanya berfungsi sebagai tempat penyimpanan koleksi fisik, tetapi juga sebagai penyedia layanan informasi digital yang responsif terhadap perkembangan teknologi. Jumlah pengunjung perpustakaan dapat dijadikan indikator tingkat minat baca dan efektivitas layanan yang diberikan kepada masyarakat. Dengan meningkatnya kebutuhan masyarakat terhadap informasi, analisis terhadap pola kunjungan menjadi penting untuk mendukung pengelolaan layanan dan sumber daya perpustakaan secara efisien [3].

2.2. Konsep Machine Learning

Machine learning merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang memungkinkan sistem untuk belajar secara otomatis dari data tanpa pemrograman eksplisit. Proses pembelajaran dilakukan melalui identifikasi pola dan hubungan antar variabel pada data historis untuk menghasilkan prediksi terhadap data baru [1], [2]. Berdasarkan jenisnya, machine learning dibedakan menjadi *supervised learning*, *unsupervised learning*, dan *reinforcement learning*. Dalam konteks penelitian ini, pendekatan *supervised learning* digunakan untuk memprediksi jumlah pengunjung berdasarkan data historis kunjungan perpustakaan. Pendekatan ini efektif dalam kasus peramalan karena memanfaatkan pasangan data masukan dan keluaran sebagai dasar pelatihan model [5], [7].

2.3. Algoritma Random Forest

Random Forest merupakan salah satu algoritma *supervised learning* berbasis ensemble method yang menggabungkan banyak pohon keputusan (*decision*

trees) untuk menghasilkan prediksi yang lebih stabil dan akurat. Algoritma ini bekerja dengan membangun sejumlah pohon keputusan menggunakan metode *bootstrap* sampling, kemudian hasil dari seluruh pohon digabungkan melalui mekanisme voting atau *averaging* untuk menentukan output akhir. Keunggulan utama Random Forest adalah kemampuannya dalam menghindari *overfitting*, menangani data berukuran besar, serta menghasilkan performa yang baik pada data nonlinier [1], [5], [6]. Selain itu, metode ini juga memiliki interpretabilitas yang baik melalui analisis *feature importance* sehingga memudahkan peneliti dalam memahami variabel yang paling berpengaruh terhadap hasil prediksi [8], [10].

2.4. Studi Tekait

Berbagai penelitian telah membuktikan bahwa algoritma Random Forest memiliki performa yang baik dalam berbagai bidang prediksi dan klasifikasi berbasis data historis. Dalam bidang keamanan siber, algoritma ini digunakan untuk meningkatkan efektivitas sistem autentikasi *One-Time Password* (OTP) dengan tingkat akurasi mencapai 98,5% dalam mengklasifikasikan transaksi OTP [1]. Pada bidang kesehatan, Random Forest juga digunakan untuk memprediksi risiko obesitas berdasarkan pola konsumsi makanan dengan akurasi sebesar 81,76%, yang menunjukkan bahwa model mampu mengidentifikasi faktor-faktor penting yang mempengaruhi risiko obesitas seperti aktivitas fisik, jenis transportasi, serta riwayat keluarga [2]. Selain itu, pada penelitian lain algoritma Random Forest mampu menghasilkan tingkat akurasi yang tinggi hingga 95%–99% pada proses klasifikasi data dengan teknik pra-pemrosesan yang tepat serta mampu menangani permasalahan *imbalanced dataset* secara efektif [3].

Dalam penerapan lainnya, Random Forest juga digunakan untuk mendeteksi anomali lalu lintas jaringan dengan performa yang baik, ditunjukkan melalui nilai akurasi sebesar 0,9151 serta F1-score sebesar 0,8510 yang menunjukkan kemampuan model dalam mengenali pola anomali secara stabil [4]. Pada bidang pendidikan, algoritma ini digunakan untuk memprediksi jumlah mahasiswa baru dengan tingkat kesalahan prediksi yang relatif rendah dibandingkan metode regresi linier [5], serta mampu memprediksi prestasi akademik siswa dengan akurasi mencapai 87,40% [10]. Pada sektor pariwisata, Random Forest juga digunakan untuk memperkirakan jumlah kunjungan wisatawan sebagai bagian dari analisis peramalan berbasis data historis [6]. Selain itu, metode ini telah diterapkan untuk memprediksi pola peminjaman buku di perpustakaan dengan akurasi mencapai 89,19% serta menunjukkan keseimbangan performa yang baik antara precision, recall, dan F1-score [7]. Random Forest juga dimanfaatkan dalam bidang kesehatan untuk memprediksi distribusi dosis terapi kanker dengan hasil prediksi yang cukup baik sebagai dasar pengembangan algoritma prediksi dosis

di masa depan [8]. Pada bidang properti, algoritma ini digunakan untuk memprediksi harga rumah dan menghasilkan tingkat akurasi serta nilai F1-score sebesar 88%, yang menunjukkan performa lebih baik dibandingkan beberapa metode klasifikasi lainnya [9].

3. METODE PENELITIAN

Bagian ini menjelaskan metode yang digunakan dalam penelitian untuk memprediksi jumlah pengunjung perpustakaan di Kota Bandung dengan menggunakan algoritma Random Forest. Pendekatan ini dipilih karena memiliki kemampuan yang baik dalam memproses data kompleks, mengenali pola historis, dan menghasilkan hasil prediksi yang stabil.

3.1. Data Penelitian

Data yang digunakan merupakan data sekunder yang diperoleh dari portal Open Data Bandung, mencakup jumlah pengunjung perpustakaan pada periode tahun 2019 hingga 2024. Setiap entri berisi informasi tahun, bulan, dan jumlah pengunjung. Data tersebut digunakan sebagai dasar dalam melakukan pelatihan dan pengujian model untuk menghasilkan prediksi jumlah pengunjung pada tahun 2025.

3.2. Pra-pemrosesan Data

Pada tahap ini, dilakukan pembersihan dan transformasi data agar siap digunakan oleh model. Kolom “bulan” dikonversi dari teks (nama bulan dalam bahasa Indonesia) menjadi nilai numerik (1–12). Seluruh nilai kosong dan data duplikat dihapus, sedangkan kolom target “jumlah pengunjung” dikonversi ke tipe numerik. Selain itu, dibuat beberapa fitur tambahan untuk menangkap pola musiman dan historis, yaitu fitur $\sin_bulan$ dan $\cos_bulan$ untuk representasi siklus bulanan, serta lag_1 dan lag_2 yang merepresentasikan jumlah pengunjung satu dan dua bulan sebelumnya.

3.3. Pemodelan

Algoritma Random Forest Regressor digunakan sebagai model utama dalam penelitian ini. Random Forest merupakan metode *ensemble learning* yang membangun sejumlah pohon keputusan (*decision tree*) dan menggabungkan hasil prediksi dari setiap pohon untuk meningkatkan akurasi serta stabilitas model [3]. Metode ini juga mampu menangani data yang kompleks dan tidak seimbang (*imbalanced dataset*) dengan baik.

Secara matematis, prediksi pada Random Forest Regressor diperoleh dari rata-rata hasil prediksi seluruh pohon keputusan yang terbentuk dalam model. Jika terdapat N pohon keputusan dalam model Random Forest, maka nilai prediksi dapat dihitung sebagai berikut:

$$\hat{y} = 1/N \sum_{i=1}^N f_i(x)$$

di mana $f_i(x)$ merupakan hasil prediksi dari pohon keputusan ke- i dan N adalah jumlah total pohon dalam model Random Forest.

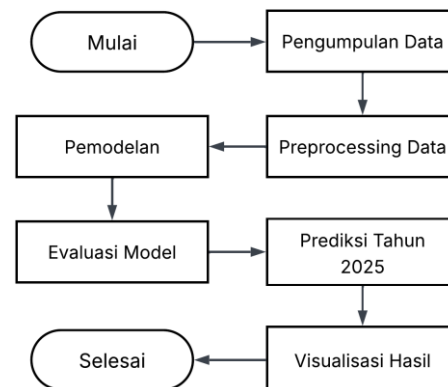
Dataset dalam penelitian ini dibagi menjadi dua bagian, yaitu 80% data pelatihan (*training data*) dan 20% data pengujian (*testing data*) tanpa proses pengacakan (*shuffle=False*) untuk menjaga urutan waktu. Model dibangun dengan parameter utama $n_estimators = 300$ dan $random_state = 42$ untuk menjaga konsistensi hasil pelatihan. Model dilatih menggunakan data tahun 2019–2023 dan diuji menggunakan data tahun 2024.

3.4. Evaluasi Model

Setelah model dilatih, dilakukan pengujian terhadap data uji menggunakan tiga metrik evaluasi, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Metrik MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai aktual dan prediksi, RMSE memperhitungkan besarnya variansi kesalahan, sedangkan MAPE digunakan untuk mengetahui persentase rata-rata kesalahan prediksi terhadap nilai aktual. Nilai MAPE yang diperoleh digunakan untuk mengukur tingkat akurasi model, dengan kategori akurasi sangat baik ($MAPE \leq 10\%$), baik ($\leq 20\%$), cukup ($\leq 50\%$), dan buruk ($> 50\%$).

3.5. Tahapan Penelitian

Tahapan penelitian yang dilakukan dalam studi ini dapat dilihat pada Gambar 1. Diagram tersebut menggambarkan alur proses penelitian dari tahap awal hingga tahap akhir, dimulai dari pengumpulan data hingga evaluasi hasil prediksi.



Gambar 1. Diagram alur penelitian

Diagram alur pada Gambar 1 menunjukkan urutan kegiatan utama dalam penelitian ini. Langkah-langkah penelitian secara ringkas meliputi:

- Pengumpulan data dari portal Open Data Bandung tahun 2019–2024.
- Preprocessing data, meliputi pembersihan data, konversi bulan ke format numerik, serta pembuatan fitur tambahan.
- Pemodelan menggunakan algoritma Random Forest dengan parameter $n_estimators=300$ dan $random_state=42$ untuk menghasilkan prediksi tahun 2025.

- d. Evaluasi hasil prediksi dengan menghitung nilai MAE, RMSE, dan MAPE untuk menilai performa model.
- e. Prediksi jumlah pengunjung tahun 2025 menghasilkan output berupa prediksi per bulan, total kunjungan tahunan, dan rata-rata kunjungan bulanan.
- f. Visualisasi hasil dilakukan dengan menampilkan grafik perbandingan antara data aktual dan hasil prediksi tahun 2025. Grafik ini memperlihatkan tren jumlah pengunjung perpustakaan dari tahun 2019 hingga 2025.

4. HASIL DAN PEMBAHASAN

4.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari portal Open Data Kota Bandung, mencakup bulan Januari 2019 hingga Desember 2024. Dataset tersebut berisi jumlah pengunjung perpustakaan per bulan yang mencerminkan aktivitas literasi masyarakat di Kota Bandung. Data ini dijadikan dasar dalam membangun model prediksi jumlah pengunjung tahun 2025 karena memiliki pola musiman dan fluktuasi tahunan.

4.2. Preprocessing Data

Tahapan preprocessing data dilakukan untuk memastikan dataset siap digunakan pada tahap pemodelan. Proses ini meliputi:

- a. Pembersihan data untuk menghapus nilai kosong (*missing value*).
- b. Konversi bulan menjadi format numerik agar dapat diproses oleh model.
- c. Pembuatan fitur musiman menggunakan transformasi *sin_bulan* dan *cos_bulan* guna merepresentasikan pola periodik tahunan.
- d. Pembuatan fitur lag, yaitu *lag_1* dan *lag_2*, yang masing-masing merepresentasikan jumlah pengunjung pada satu dan dua bulan sebelumnya.

4.3. Pemodelan

Pemodelan dilakukan menggunakan algoritma Random Forest Regressor dari pustaka *scikit-learn*. Dataset dibagi menjadi dua bagian, yaitu 80% data pelatihan (57 data) dan 20% data pengujian (15 data) menggunakan fungsi *train_test_split* dengan parameter *shuffle=False* untuk menjaga urutan waktu.

Model dibangun dengan parameter utama *n_estimators=300*, yaitu jumlah decision tree yang digunakan, serta *random_state=42* untuk memastikan hasil pelatihan yang konsisten. Data tahun 2019–2023 digunakan untuk melatih model, sedangkan data tahun 2024 digunakan untuk pengujian. Pemilihan algoritma ini didasarkan pada kemampuannya dalam menangani data dengan hubungan non linier dan menghindari risiko *overfitting* melalui kombinasi pohon keputusan. Implementasi kode program untuk pembagian data dan latihan model ditunjukkan pada Gambar 2.

```
1 # Siapkan fitur & target
2 x = df[['tahun', 'sin_bulan', 'cos_bulan', 'lag_1', 'lag_2']]
3 y = df['jumlah_pengunjung']
4
5 # Split data menjadi train dan test (80% train, 20% test)
6 if len(df) < 15:
7     x_train, x_test, y_train, y_test = x, x, y, y
8 else:
9     x_train, x_test, y_train, y_test = train_test_split(
10         x, y, test_size=0.2, random_state=42, shuffle=False
11     )
12
13 # Inisialisasi dan latih model
14 model = RandomForestRegressor(n_estimators=300, random_state=42)
15 model.fit(x_train, y_train)
```

Gambar 2. Split data & pemodelan

4.4. Evaluasi Model

Evaluasi model dilakukan menggunakan data tahun 2024 yang tidak dilibatkan dalam proses pelatihan, dengan tujuan untuk mengukur sejauh mana model mampu melakukan prediksi terhadap data baru secara akurat sebelum digunakan pada peramalan tahun 2025. Hasil evaluasi model Random Forest terhadap data pengujian ditunjukkan pada Tabel 1.

Tabel 1. Evaluasi Model

MAE	963.04
RMSE	1249.70
MAPE	33.29%
Akurasi	Cukup

Berdasarkan hasil evaluasi pada Tabel 1, model Random Forest menghasilkan nilai MAE sebesar 963,04 yang menunjukkan bahwa rata-rata selisih antara nilai prediksi dan data aktual sekitar 963 pengunjung. Nilai ini menggambarkan tingkat kesalahan absolut model dalam memprediksi jumlah pengunjung perpustakaan pada setiap periode pengujian. Sementara itu, nilai RMSE sebesar 1249,70 menunjukkan bahwa model memberikan penalti yang lebih besar terhadap kesalahan prediksi yang tinggi, sehingga metrik ini sensitif terhadap adanya deviasi yang cukup besar pada beberapa periode tertentu. Perbedaan nilai RMSE yang lebih tinggi dibandingkan MAE mengindikasikan bahwa terdapat beberapa bulan dengan kesalahan prediksi yang relatif lebih besar dibandingkan bulan lainnya.

Selain itu, nilai MAPE sebesar 33,29% menunjukkan bahwa rata-rata kesalahan prediksi model berada pada kisaran sekitar 33% dari nilai aktual. Berdasarkan kategori evaluasi MAPE, nilai tersebut termasuk dalam kategori cukup, sehingga model masih mampu memberikan gambaran pola kunjungan pengunjung meskipun tingkat kesalahan prediksi masih perlu diperbaiki. Secara keseluruhan, hasil evaluasi ini menunjukkan bahwa model Random Forest telah mampu menangkap pola umum dari data kunjungan perpustakaan, namun masih terdapat variasi data yang menyebabkan perbedaan antara nilai prediksi dan nilai aktual pada beberapa periode pengamatan.

4.5. Prediksi Tahun 2025

Model Random Forest yang telah melalui tahap evaluasi kemudian digunakan untuk memprediksi jumlah pengunjung perpustakaan di Kota Bandung pada tahun 2025. Proses prediksi dilakukan menggunakan fitur-fitur yang telah dibentuk pada tahap pra-pemrosesan, yaitu tahun, sin/cos bulan, serta nilai lag bulan sebelumnya.

Hasil prediksi mencakup estimasi jumlah pengunjung untuk setiap bulan pada tahun 2025, total keseluruhan pengunjung selama satu tahun, serta rata-rata kunjungan bulanan. Rincian hasil prediksi dapat dilihat pada Tabel 2.

Tabel 2. Hasil prediksi jumlah pengunjung 2025

Bulan	Jumlah Pengunjung
Januari	1,881
Februari	1,785
Maret	1,766
April	1,596
Mei	1,763
Juni	1,753
Juli	1,796
Agustus	1,788
September	1,877
Oktober	1,879
November	1,877
Desember	1,864

Berdasarkan hasil prediksi pada Tabel 2, jumlah pengunjung perpustakaan di Kota Bandung pada tahun 2025 diperkirakan mencapai sekitar 21.525 pengunjung dengan rata-rata sekitar 1.794 kunjungan per bulan. Secara umum, nilai prediksi bulanan menunjukkan pola yang relatif stabil dengan kisaran jumlah pengunjung antara sekitar 1.596 hingga 1.881 pengunjung per bulan. Prediksi tertinggi diperkirakan terjadi pada bulan Januari dengan jumlah sekitar 1.881 pengunjung, sedangkan jumlah kunjungan terendah diperkirakan terjadi pada bulan April dengan sekitar 1.596 pengunjung.

Jika dilihat dari pola bulanan, sebagian besar bulan menunjukkan jumlah pengunjung yang tidak mengalami fluktuasi yang terlalu besar, yang mengindikasikan bahwa aktivitas kunjungan perpustakaan cenderung stabil sepanjang tahun. Setelah mengalami penurunan pada bulan April, jumlah kunjungan diperkirakan kembali meningkat secara bertahap pada bulan-bulan berikutnya hingga mencapai nilai yang relatif tinggi pada periode September hingga November. Pola ini menunjukkan adanya kecenderungan peningkatan aktivitas kunjungan pada paruh kedua tahun. Dengan demikian, hasil prediksi ini dapat memberikan gambaran awal bagi pihak pengelola perpustakaan dalam merencanakan strategi pengelolaan layanan, pengadaan koleksi, maupun penyediaan fasilitas untuk mengantisipasi jumlah kunjungan pada tahun 2025.

4.6. Visualisasi Model



Gambar 3. Visualisasi perbandingan data aktual dan hasil prediksi

Visualisasi hasil prediksi dilakukan untuk memperlihatkan perbandingan antara data aktual jumlah pengunjung perpustakaan (2019–2024) dengan hasil prediksi tahun 2025 menggunakan algoritma Random Forest. Grafik ini menampilkan tren perubahan jumlah pengunjung dari waktu ke waktu, sekaligus menunjukkan kemampuan model dalam mengikuti pola data historis.

Garis berwarna biru merepresentasikan data aktual, sedangkan garis merah putus-putus menunjukkan hasil prediksi untuk tahun 2025. Secara visual, model mampu mempertahankan pola tren historis dengan fluktuasi yang relatif stabil pada tahun prediksi.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil membangun model prediksi jumlah pengunjung perpustakaan di Kota Bandung menggunakan algoritma Random Forest dengan data periode 2019–2024. Model yang dihasilkan menunjukkan performa yang cukup baik dengan nilai MAPE sebesar 33,29%, yang menunjukkan bahwa model mampu menangkap pola umum kunjungan pengunjung meskipun masih terdapat tingkat kesalahan prediksi pada beberapa periode. Berdasarkan hasil prediksi, jumlah pengunjung perpustakaan pada tahun 2025 diperkirakan mencapai sekitar 21.525 kunjungan dengan rata-rata 1.794 pengunjung per bulan, yang menunjukkan tren kunjungan yang relatif stabil. Hasil prediksi ini dapat menjadi referensi awal bagi pengelola perpustakaan dalam merencanakan kapasitas layanan, pengelolaan fasilitas, serta pengembangan strategi layanan berbasis data.

Untuk penelitian selanjutnya, disarankan untuk menggunakan data historis yang lebih panjang serta menambahkan variabel lain seperti hari kerja, kegiatan akademik, hari libur, maupun faktor eksternal lainnya yang dapat mempengaruhi jumlah kunjungan perpustakaan. Selain itu, penelitian selanjutnya juga dapat melakukan perbandingan dengan metode machine learning lainnya untuk memperoleh model prediksi dengan tingkat akurasi yang lebih baik.

DAFTAR PUSTAKA

- [1] Derman Janner Lubis dan A. N. R. Andri, "Implementasi Algoritma Random Forest Untuk Optimasi Keamanan Autentikasi One-Time Password (OTP)," *Informatech J. Ilm. Inform. Dan Komput.*, vol. 1, no. 1, hlm. 23–29, Mei 2024, doi: 10.69533/eyp7ag46.
- [2] A. Muliawan, D. A. Fauziah, dan E. Afrianto, "Obesity Risk Prediction Using Random Forest Based on Eating Habit Parameters".
- [3] Ega Ranaldi Pebriansyah, Susanti, Rahmiati, dan Triyani Arita Fitri, "Prediction of Library Book Borrowing Patterns Using The Random Forest Algorithm," *J. Ris. Inform.*, vol. 7, no. 4, hlm. 327–335, Sep 2025, doi: 10.34288/jri.v7i4.409.
- [4] R. A. Saputri dan L. Rosnita, "A Random Forest-Based Predictive Model for Student Academic Performance: A Case Study in Indonesian Public High Schools," vol. 9, no. 3.
- [5] M. Nurdin dan F. Fauziah, "Analytical Study Forecasting Students Using Random Forest and Linear Regression Algorithms," *sinkron*, vol. 8, no. 4, hlm. 2369–2378, Okt 2024, doi: 10.33395/sinkron.v8i4.13886.
- [6] A. Nugroho dan A. Husin, "Performance Analysis of Random Forest Using Attribute Normalization," *SISTEMASI*, vol. 11, no. 1, hlm. 186, Jan 2022, doi: 10.32520/stmsi.v11i1.1681.
- [7] T. Agustina, M. Masrizal, dan I. Irmayanti, "Performance Analysis of Random Forest Algorithm for Network Anomaly Detection using Feature Selection," *sinkron*, vol. 8, no. 2, Apr 2024, doi: 10.33395/sinkron.v8i2.13625.
- [8] R. H. Hirzi, U. Hidayaturrohman, K. Kertanah, M. H. Amaly, dan R. Satriawan, "Prediksi Jumlah Wisatawan Menggunakan Metode Random Forest, Single Exponential Smoothing dan Double Exponential Smoothing di Provinsi NTB," *Jambura J. Probab. Stat.*, vol. 4, no. 1, hlm. 47–55, Mei 2023, doi: 10.34312/jjps.v4i1.17088.
- [9] R. Ramlah, M. Fadli, J. Valerian, P. Prajitno, dan D. S. Sihono, "RANDOM-FOREST MODEL PREDICTION OF DOSE DISTRIBUTION IN IMRT PLANNING FOR LUNG CANCER," *Iran. J. Med. Phys.*, no. Online First, Agu 2022, doi: 10.22038/ijmp.2022.64319.2101.
- [10] R. Tanamal, N. Minoque, T. Wiradinata, Y. Sockamto, dan T. Ratih, "House Price Prediction Model Using Random Forest in Surabaya City," *TEM J.*, hlm. 126–132, Feb 2023, doi: 10.18421/TEM121-17.