

MODEL FORECASTING TREN KUNJUNGAN WISATAWAN DI DIY MENGGUNAKAN REGRESI LOGISTIK BINER

Erma Susanti^{1*}, Windyaning Ustyannie², I Wayan Julianta Pradnyana³, Katherina Irene Dhamayanti⁴

^{1,2,4,3} Universitas AKPRIND Indonesia, *Penulis Koresponden

e-mail: ¹erma@akprind.ac.id, ²wustyannie@akprind.ac.id, ³juliantapradnyana@gmail.com,

⁴katherina3n@gmail.com

ABSTRACT

The Special Region of Yogyakarta (DIY) is a province that has special features in terms of arts and culture, nature, culinary, shopping, and other tourism. Visits by both foreign and domestic tourists to DIY before the COVID pandemic and after the pandemic have fluctuated. The many choices of tourist attractions in DIY attract tourists to come. An increase or decrease in the number of tourists can affect various aspects of people's lives. An increase in tourist visits will affect congestion and traffic on the highway, while a decrease in the number of visits can affect regional income and the income of people who depend on tourism for their livelihood. In addition, the increase in the number of tourists also needs to be anticipated by the government to provide road infrastructure, buildings, city planning, traffic order, waste management, and so on. The existence of a forecasting model that can predict future tourist visits can help stakeholders make decisions to handle problems related to the impact of tourism on the community and to improve the governance of tourist visits. The use of a binary logistic regression algorithm in this case is used to predict the trend of an increasing or decreasing number of tourists for the next two years until 2026. Historical data of tourist visits from 2018 to 2024 from BPS is used for this study. The forecasting results show an increase in tourist visits in June 2025 and 2026. The evaluation results show the accuracy, precision, recall, and f1-score values of 1.0 (for the range 0-1). These results indicate that the forecasting model has a very good accuracy value.

Keywords: binary, forecasting, tourist visits, foreign, logistic regression

INTISARI

Daerah Istimewa Yogyakarta (DIY) merupakan sebuah provinsi yang memiliki keistimewaan baik itu dari segi wisata, seni budaya, alam, kuliner, belanja, dan lainnya. Kunjungan wisatawan khususnya mancanegara ke DIY sebelum pandemi COVID dan setelah pandemi mengalami fluktuasi peningkatan dan penurunan. Adanya tren kenaikan atau penurunan jumlah wisatawan dapat berpengaruh terhadap berbagai segi kehidupan masyarakat. Peningkatan kunjungan wisata akan dapat menimbulkan dampak seperti kemacetan lalu lintas di jalan raya, sedangkan penurunan jumlah kunjungan dapat berpengaruh terhadap pendapatan daerah dan pendapatan masyarakat yang menggantungkan hidup dari pariwisata. Kondisi ini perlu diantisipasi oleh pemerintah untuk menyediakan infrastruktur jalan, bangunan, tata kota, ketertiban lalu lintas, pengelolaan sampah dan sebagainya. Adanya sebuah model *forecasting* (peramalan) yang dapat memprediksi kunjungan wisatawan ke depan dapat membantu pemangku kepentingan untuk dapat mengambil keputusan dalam penanganan permasalahan yang berkaitan dengan dampak wisata untuk masyarakat dan untuk memperbaiki tata kelola kunjungan wisata. Penggunaan algoritma regresi logistik biner dalam kasus ini digunakan untuk meramalkan tren kenaikan atau penurunan jumlah wisatawan selama dua tahun ke depan (2025 sampai 2026). Data historis kunjungan wisata dari tahun 2019 sampai dengan 2024 dari BPS digunakan untuk penelitian ini. Hasil *forecasting* menunjukkan kenaikan kunjungan wisatawan mancanegara pada bulan Juni pada tahun 2025 dan 2026. Hasil evaluasi menunjukkan nilai akurasi, *precision*, *recall*, *f1-score* 1.0 (untuk range 0-1). Hasil tersebut menunjukkan bahwa model *forecasting* memiliki nilai akurasi yang sangat baik.

Kata kunci: biner, peramalan, kunjungan wisatawan, mancanegara, regresi logistik

1. PENDAHULUAN

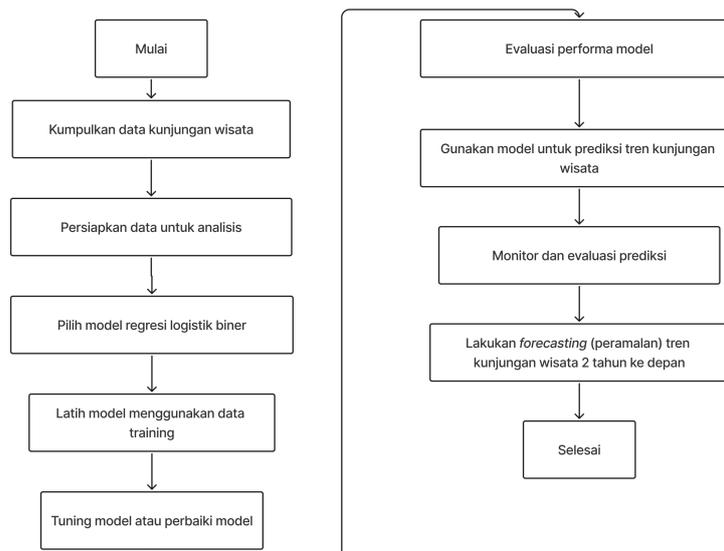
Metode regresi logistik merupakan salah satu metode yang dapat digunakan untuk menyelesaikan masalah baik itu klasifikasi maupun peramalan (*forecasting*). Peramalan data untuk prediksi jumlah wisatawan khususnya mancanegara ke depan akan coba diselesaikan untuk mendapatkan model peramalan yang akurat. Penelitian juga dilakukan untuk mendapatkan model peramalan dengan menggunakan metode regresi logistik biner dan dilakukan perhitungan evaluasi performa akurasi. Hasil penelitian dapat dijadikan sebagai bahan acuan untuk langkah-langkah penanganan wisatawan oleh pemangku kepentingan di masa datang. Peramalan masa depan akan diteliti untuk prediksi kunjungan wisata selama 24 bulan pada tahun 2025 hingga tahun 2026.

Studi kasus peramalan yang akan dilakukan pada penelitian ini yaitu peramalan jumlah wisatawan khususnya mancanegara di Daerah Istimewa Yogyakarta (DIY), dimana *dataset* yang digunakan terbagi menjadi kelas biner (0 dan 1). Nilai 1 jika jumlah kunjungan wisata mengalami kenaikan dan nilai 0 jika mengalami penurunan. Kebaruan penelitian ini ada pada penggunaan studi kasus untuk *forecasting* (peramalan), dimana pada sebagian besar penelitian lebih banyak membahas kasus klasifikasi. Tujuan penelitian ini, selain untuk melakukan peramalan kunjungan wisatawan mancanegara di masa depan (2025-2026), juga untuk mengetahui akurasi performa model regresi logistik biner yang dihasilkan. Selanjutnya dapat diketahui fitur yang perlu dikembangkan untuk penelitian lanjutan.

Penelitian tentang prediksi kunjungan wisatawan domestik pernah dilakukan oleh Nawawi et al. (2021) dengan metode *Holt Winter Additive*. Prediksi kunjungan wisata dengan *Time series forecasting* dan LSTM juga dilakukan oleh (Manullang & Santoso, 2020). Prediksi lainnya menggunakan metode *Decision Tree* dengan algoritma C4.5 oleh Iman dkk (2024). Penelitian prediksi jumlah wisatawan dengan regresi linear dilakukan oleh (Manurung et al., 2024) dan (Sovia et al., 2020). Sedangkan prediksi kunjungan wisatawan mancanegara dengan ARIMA, Glnet, dan Prophet dilakukan dari (Mu'minin et al., 2022), evaluasi kunjungan wisatawan mancanegara dengan metode *forecasting* diteliti oleh (Rais et al., 2020). (Sovia et al., 2020), (Putri et al., 2021), (Sari & Agustini, 2024) dengan *Backpropagation*. (Utami & Maulana, 2020) melakukan prediksi kunjungan wisata dengan model *time series*.

2. METODE PENELITIAN

Proses penelitian ini dimulai dari melakukan studi pustaka terhadap objek yang akan diteliti yaitu peramalan jumlah wisatawan di Yogyakarta. Kelas data dibagi menjadi data *up/down* (kenaikan/penurunan) jumlah wisatawan. Selanjutnya melakukan tahapan-tahapan penelitian mulai dari proses pengumpulan data sampai mendapatkan hasil.



Gambar 1. Diagram Alir Tahapan Penelitian

Gambar 1 merupakan diagram alir tahapan penelitian ini. Tahapan proses peramalan dimulai dari pengumpulan *dataset*. Selanjutnya dilakukan proses untuk memuat *dataset* dan dilakukan *preprocessing* untuk mengkonversi data agar siap digunakan untuk diolah untuk tahapan pemodelan. Berikutnya pelatihan model dan evaluasi performa model. Jika model sudah mendapatkan performa yang baik, maka lakukan peramalan kenaikan atau

penurunan jumlah wisatawan, lalu evaluasi kembali hasil performa, jika sudah baik maka proses selesai dan lakukan dokumentasi. Jika performa model belum baik, maka sesuaikan atau perbaiki kembali model. Jika hasil evaluasi performa peramalan masih belum baik maka ulangi pelatihan model sampai mendapatkan hasil akurasi pelatihan dan peramalan yang baik (*good fit*).

Tahapan prosedur kerja pada penelitian ini antara lain:

1. Melakukan analisis dan pengumpulan data penelitian (*dataset*).
2. Melakukan *preprocessing* (pra-proses) data.
3. Melakukan proses tahapan pembuatan model dengan regresi logistik biner.
4. Melakukan evaluasi terhadap model.
5. Menyusun dokumentasi hasil penelitian.
6. Melakukan publikasi hasil penelitian.

Regresi Logistik Biner

Regresi logistik adalah sebuah pendekatan untuk membuat model prediksi seperti halnya regresi linear atau yang biasa disebut dengan istilah *Ordinary Least Squares (OLS) regression*. Perbedaannya pada regresi logistik, peneliti memprediksi variabel terikat yang berskala dikotomi (skala data nominal dengan dua kategori, misalnya: 'Ya' dan 'Tidak', 'Baik' dan 'Buruk', 'Tinggi' dan 'Rendah'). Regresi logistik merupakan tipe dari analisis regresi yang digunakan untuk memprediksi hasil secara biner (contoh 'yes/no', 'ya/tidak', '0/1').

Model Persamaan Regresi Logistik

Metode regresi logistik adalah suatu metode analisis statistik yang menggambarkan hubungan antara variabel respon berkategori yang mempunyai dua kategori (biner) atau lebih dengan satu atau lebih variabel prediktor. Variabel respon biner adalah variabel respon yang hanya bernilai 1 untuk adanya suatu karakteristik dan 0 untuk tidak adanya karakteristik. Model regresi logistik digunakan untuk melihat probabilitas suatu kejadian dan membandingkan resiko terjadinya suatu kejadian dengan menghitung faktor-faktor yang mempengaruhinya.

Fungsi probabilitas distribusi Bernoulli (Yap et al., 2014) ditunjukkan pada persamaan 1.

$$f(Y_i) = \pi_i^{Y_i} (1 - \pi_i)^{1-Y_i} \quad (1)$$

dimana π_i adalah peluang terjadinya kemunculan ke-i dan variabel acak ke-i. Jika $Y_i = 0$ maka $f(Y_i) = (1 - \pi_i)$ jika $Y_i = 1$ maka $f(Y_i) = \pi_i$.

Fungsi model logit yang terletak diantara rentang 0 dan 1 diperoleh dengan menggunakan fungsi logit pada persamaan 2.

$$\pi(x_i) = \frac{1}{1 + e^{-g(x_i)}} \quad (2)$$

$\pi(x_i)$ adalah fungsi non-linier maka perlu diubah ke dalam bentuk logit untuk memperoleh fungsi yang linier sehingga dapat dilihat hubungan antara variabel bebas dan variabel tidak bebas.

Transformasi logit dari model pada persamaan 3.

$$\begin{aligned} g(x_i) &= \ln \left[\frac{\pi(x_i)}{1 - \pi(x_i)} \right] \\ &= \ln(\pi(x_i)) - \ln(1 - \pi(x_i)) \\ &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} \end{aligned} \quad (3)$$

Keterangan:

β_0 : konstanta

β_j : koefisien parameter with $j = 1, 2, \dots, p$.

Regresi Logistik (Hosmer et al., 2013) dapat dibedakan menjadi 2, yaitu:

1. *Binary Logistic Regression* (Regresi Logistik Biner). Regresi Logistik biner digunakan ketika hanya ada 2 kemungkinan variabel respon (Y), misal membeli dan tidak membeli.
2. *Multinomial Logistic Regression* (Regresi Logistik Multinomial). Regresi Logistik Multinomial digunakan ketika pada variabel respon (Y) terdapat lebih dari 2 kategorisasi.

Regresi logistik biner merupakan suatu metode statistika yang di dalamnya menggambarkan hubungan antara variabel respon (y) bersifat biner dengan variabel prediktor (x) yang bersifat kualitatif maupun sebaliknya (Hosmer et al., 2013). Regresi logistik biner digunakan untuk menguji pengaruh dari data dependen (y) yaitu data

ketegorigal yang terdiri dari dua data (contoh kode 1 = ya dan kode 0 = tidak). Sedangkan untuk data independen (x) merupakan data primer dan sekunder yang bertujuan untuk menguji apakah probabilitas terjadinya variabel dependen dapat diprediksi dengan variabel independen (Sujarweni & Utami, 2019).

3. HASIL DAN PEMBAHASAN

Hasil penelitian terdiri dari beberapa bagian seperti hasil pengumpulan *dataset*, hasil *preprocessing*, hasil pemodelan regresi logistik, dan evaluasi model.

3.1. Penyiapan *Dataset*

Data kunjungan wisata di Daerah Istimewa Yogyakarta (DIY) dengan *baseline* dari tahun 2018 sampai dengan tahun 2024 menggunakan data publik dari Badan Pusat Statistik (BPS). *Dataset* yang digunakan menggunakan variabel target yaitu tren kunjungan wisatawan khususnya mancanegara (kenaikan/penurunan). *Dataset* kunjungan wisatawan dari BPS ditampilkan pada **Tabel 1**. Gambar grafik data tahun vs data *baseline* dapat dilihat pada **Gambar 2**.

Tabel 1. Dataset Kunjungan Wisatawan Mancanegara Baseline Tahun 2018

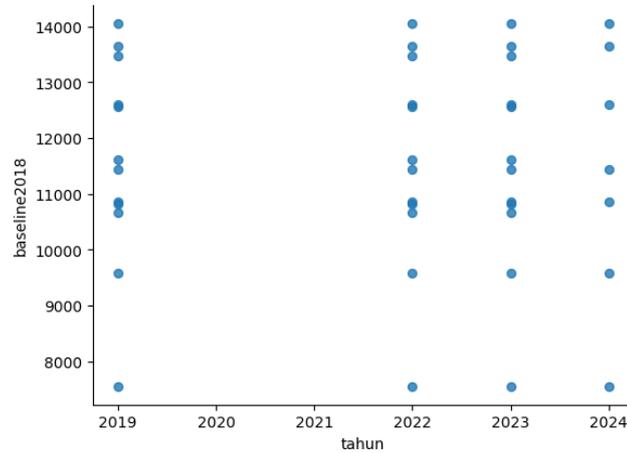
Waktu	Capaian 2018	Target	Class	Waktu	Capaian 2018	Target	Class
Jan-19	10849	8055	0	Jan-23	10849	3883	0
Feb-19	12602	9631	0	Feb-23	12602	4849	0
Mar-19	14055	10623	0	Mar-23	14055	5007	0
Apr-19	11435	9980	0	Apr-23	11435	5592	0
May-19	9584	6688	0	May-23	9584	8160	0
Jun-19	7547	6684	0	Jun-23	7547	11188	1
Jul-19	13642	10532	0	Jul-23	13642	12444	0
Aug-19	13468	12091	0	Aug-23	13468	12974	0
Sep-19	12553	10655	0	Sep-23	12553	11855	0
Oct-19	11613	9577	0	Oct-23	11613	8536	0
Nov-19	10661	9218	0	Nov-23	10661	9675	0
Dec-19	10813	9795	0	Dec-23	10813	11338	1
Jan-22	10849	0	0	Jan-24	10849	6879	0
Feb-22	12602	0	0	Feb-24	12602	9813	0
Mar-22	14055	0	0	Mar-24	14055	6258	0
Apr-22	11435	35	0	Apr-24	11435	8790	0
May-22	9584	260	0	May-24	9584	11323	1
Jun-22	7547	369	0	Jun-24	7547	8134	1
Jul-22	13642	809	0	Jul-24	13642	11120	0
Aug-22	13468	640	0				
Sep-22	12553	962	0				
Oct-22	11613	1933	0				
Nov-22	10661	1947	0				
Dec-22	10813	5169	0				

Keterangan:

Tahun : Tahun data

Jumlah Wisatawan : Jumlah wisatawan mancanegara

Tren : Kategori target (1 = naik, 0 = turun dibandingkan tahun sebelumnya)



Gambar 2. Data Tahun Vs Baseline

3.2. Pra-proses (*Preprocessing Data*)

Dataset sebelum diproses untuk menghasilkan model dilakukan beberapa proses seperti *balancing* (menyeimbangkan) *dataset*. Selain itu juga dilakukan konversi data ke format yang mendukung untuk regresi logistik. Selanjutnya data dibagi menjadi *data training* dan *data testing* untuk membuat model untuk memprediksi kenaikan atau penurunan wisatawan mancanegara. Contoh data hasil pra-proses dapat dilihat pada Gambar 3.

	waktu	bulan	tahun	baseline2018	target	class
0	Jan-19	1	2019	10849	8055	0
1	Feb-19	2	2019	12602	9631	0
2	Mar-19	3	2019	14055	10623	0
3	Apr-19	4	2019	11435	9980	0
4	May-19	5	2019	9584	6688	0

Gambar 3. Contoh Data Hasil Pra-proses

3.3. Model Regresi Logistik Biner

Model regresi logistik biner dalam penelitian ini digunakan untuk membuat dan melatih model untuk memprediksi kenaikan atau penurunan jumlah wisatawan mancanegara. Regresi logistik biner untuk digunakan untuk memprediksi kenaikan atau penurunan jumlah wisata. Penelitian ini menggunakan bahasa pemrograman *Python* dan *library* (pustaka) pendukung lainnya seperti *Pandas* dan *Scikit-Learn*. Urutan proses untuk menghasilkan model *forecasting* ini antara lain:

1. Proses *import* pustaka yang diperlukan seperti *Pandas* dan *Scikit Learn*

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, classification_report
```

2. Proses memuat *dataset*.

```
data = pd.read_csv("/content/Data Penelitian 2024.csv")
```

3. Membuat variabel untuk target tren dengan kategori kenaikan dan penurunan. Variabel tren ini didapat dari menghitung selisih jumlah wisatawan selama tahun-tahun berturut-turut sebelumnya. Jika terjadi peningkatan maka tren akan diberikan nilai 1 dan jika terjadi penurunan maka tren akan diberi nilai 0.

```
# Membuat variabel target berdasarkan perbandingan dengan data bulan sebelumnya atau baseline 2018
data['class'] = (data['target'] > data['baseline2018']).astype(int)
```

- Menyiapkan fitur dan target tren. Variabel x merupakan variabel penjelas yang berisi tahun dan jumlah kunjungan, sedangkan variabel y adalah tren.

```
# Menyiapkan fitur dan class
X = data[['tahun', 'bulan', 'baseline2018']] # Fitur
y = data['class'] # Target
```

- Membagi *dataset* menjadi *data training* dan *data testing (split)*.

```
# Split data menjadi training dan testing set
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
```

- Melakukan inisialisasi dan melatih model Regresi Logistik Biner. Proses pelatihan model menggunakan model *LogisticRegression* dengan menggunakan *data training*.

```
# Inisialisasi dan training model Regresi Logistik
model = LogisticRegression(max_iter=1000)
model.fit(X_train, y_train)
```

3.4. Evaluasi Model

Melakukan evaluasi model yang dihasilkan. Evaluasi merupakan proses pengujian model untuk mendapatkan nilai akurasi dari model. Evaluasi menggunakan *data testing* sampai mendapatkan *metric* kinerja model. Hasil akurasi model ditunjukkan pada **Gambar 4**. Hasil evaluasi model menunjukkan nilai akurasi, *precision*, *recall* dan *f1-score* bernilai 1.00 (range 0-1.00).

```
# Evaluasi model
y_pred = model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"Akurasi model: {accuracy}")
print("Classification Report:\n", classification_report(y_test, y_pred))
```

```
Akurasi model: 1.00
Classification Report:
              precision    recall  f1-score   support

     0           1.00      1.00     1.00         9
 accuracy          1.00          1.00          1.00         9
 macro avg          1.00          1.00          1.00         9
 weighted avg          1.00          1.00          1.00         9
```

Gambar 4. Hasil Akurasi Model

3.5. Proses Prediksi Tren Kenaikan dan Penurunan Kunjungan Wisata

Model regresi logistik untuk penelitian ini menggunakan fungsi *sigmoid* karena akan digunakan untuk memprediksi probabilitas dari suatu kelas. Proses perhitungan *forecasting* untuk mendapatkan tren kenaikan dan penurunan kunjungan wisatawan selama dua tahun ke depan (2025 dan 2026) dihitung menggunakan regresi logistik biner. Probabilitas prediksi untuk kemungkinan terjadi kenaikan dan penurunan karena biner menggunakan 0 atau 1, dimana probabilitas untuk kelas 1 akan menggunakan formula (4).

$$P(Y = 1|X) = \frac{1}{1+e^{-(\beta_0+\beta_1.X_1+\beta_2.X_2+\dots+\beta_n.X_n)}} \quad (4)$$

Dimana:

Y = variabel target untuk tren kenaikan (1) dan tren penurunan (0)

X₁, X₂, ..., X_n = variabel fitur seperti tahun, bulan, baseline2018

β₁, β₂, ...β_n = koefisien model yang ditentukan setelah proses pelatihan model

Setelah didapatkan probabilitas nilai $P(Y = 1|X)$ maka dilakukan konversi probabilitas ke dalam bentuk biner 0 dan 1. Jika probabilitas $P(Y = 1|X) \geq 0.5$ maka hasil prediksi bernilai 1 (tren kenaikan) Jika probabilitas $P(Y = 1|X) < 0.5$ maka hasil prediksi bernilai 0 (tren penurunan). Proses prediksi untuk peramalan (*forecasting*) selama dua tahun ke depan (2025 dan 2026) dilakukan dengan menggunakan *data testing*.

Perhitungan manual:

Misal dimiliki sebuah titik data pada Januari 2025 dengan detail antara lain:

- Fitur inputan terdiri dari:
tahun = 2025
bulan = 1
baseline2018 = 1000
Koefisien model dihitung setelah pelatihan model.
Koefisien
 $z = \beta_0, \beta_1, \beta_2, \beta_3$ yang akurat dapat diperoleh dengan melatih model dengan menggunakan data histori.
Asumsi nilai antara lain:
Intersep (β_0) = -2.5
Koefisien tahun (β_1) = 0.01
Koefisien bulan (β_2) = -0.05
Koefisien baseline2018 (β_3) = 1×10^{-5} .

- Perhitungan nilai *Linear Combination* (Z) pada formula (5).

Ini merupakan kombinasi linear dari fitur dan koefisien:

$$z = \beta_0 + \beta_1 \cdot \text{tahun} + \beta_2 \cdot \text{bulan} + \beta_3 \cdot \text{baseline2018} \quad (5)$$

Nilai substitusi:

- $0.01 \cdot 2025 = 20.25$
- $-0.05 \cdot 1 = -0.05$
- $1 \times 10^{-5} \cdot 100000 = 1$

Gabungkan hasil sehingga:

$$z = -2.5 + 20.25 - 0.05 + 1 = 18.7$$

- Hitung probabilitas dengan menggunakan fungsi *sigmoid* untuk mengkonversi z menjadi probabilitas berikut

$$P(Y = 1|X) = \frac{1}{1 + e^{-z}}$$

Substitusi nilai $z = 18.7$ menjadi

$$P(Y = 1|X) = \frac{1}{1 + e^{-18.7}}$$

Hitung eksponen:

$e^{-18.7} \approx 7.57 \times 10^{-9}$ nilai ini hampir mendekati nol karena nilainya sangat kecil.

- Prediksi Kelas

Probabilitas $P(Y = 1|X)$ mendekati 1, maka prediksi untuk data ini adalah 1 yaitu tren naik. Hasil perhitungan untuk bulan Januari 2025 didapatkan hasil prediksi dengan tren naik.

3.6. Model *Forecasting* Dua Tahun Ke Depan

Selanjutnya setelah mendapatkan model prediksi, dilakukan *forecasting* (peramalan) untuk memprediksi kategori kenaikan atau penurunan jumlah wisatawan dua tahun ke depan (dalam hal ini sampai dengan tahun 2026). Hasil keluaran (*output*) akan menghasilkan prediksi jika mengalami kenaikan dengan angka 1 dan jika penurunan dengan angka 0. Prediksi *forecasting* masa depan dibuat untuk 24 bulan ke depan dari 2025 hingga 2026 berdasarkan *baseline* data kunjungan pada data tahun 2018. *Output* hasil prediksi dapat dilihat pada **Gambar 5**.

	tahun	bulan	predicted_trend
0	2025	1	0
1	2025	2	0
2	2025	3	0
3	2025	4	0
4	2025	5	0
5	2025	6	1
6	2025	7	0
7	2025	8	0
8	2025	9	0
9	2025	10	0
10	2025	11	0
11	2025	12	0
12	2026	1	0
13	2026	2	0
14	2026	3	0
15	2026	4	0
16	2026	5	0
17	2026	6	1
18	2026	7	0
19	2026	8	0
20	2026	9	0
21	2026	10	0
22	2026	11	0
23	2026	12	0

Gambar 5. Hasil *Forecasting* (Peramalan) Kunjungan Wisatawan untuk Tahun 2025 dan 2026

4. KESIMPULAN

Hasil penelitian *forecasting* dengan *Regresi Logistik Biner* menghasilkan prediksi kenaikan pada bulan juni untuk tahun 2025 dan 2026. Nilai akurasi, *precision*, *recall*, dan *f1-score* dari model menghasilkan nilai 1.0 untuk *range* (0-1.0), sehingga akurasi model menunjukkan hasil yang sangat baik. Regresi Logistik Biner pada penelitian ini memiliki keterbatasan karena hanya melakukan klasifikasi menjadi kelas naik atau turun, sehingga hasil prediksi yang lebih spesifik tidak dapat diperoleh dalam penelitian ini. Penggunaan model prediksi *time series* seperti ARIMA untuk meningkatkan akurasi model dapat dipertimbangkan. Selain itu penggunaan model *machine learning* seperti *Random Forest*, *Gradient Boosting*, dan SVM dapat juga digunakan untuk memprediksi kunjungan wisata secara lebih spesifik. Model *forecasting* yang dihasilkan pada penelitian ini menggunakan model dan data yang sederhana. Ke depan perlu dikembangkan model yang melibatkan banyak variabel eksternal. Karena variabel-variabel lainnya merupakan faktor-faktor eksternal yang dapat menimbulkan dampak kenaikan atau penurunan. Penambahan variabel lain seperti indeks kebijakan wisata, faktor ekonomi, *event*, promosi wisata, pasar global perlu untuk penelitian lanjutan.

UCAPAN TERIMA KASIH

Terima kasih kepada DP2M dan Universitas AKPRIND Indonesia atas pendanaan untuk penelitian internal ini. Terima kasih juga untuk Fakultas Sains dan Teknologi Informasi (FSTI) atas kesempatan dan dukungan untuk dapat menyelesaikan penelitian ini.

DAFTAR PUSTAKA

- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression: Third Edition. In *Applied Logistic Regression: Third Edition*. <https://doi.org/10.1002/9781118548387>
- Manullang, J., & Santoso, A. (2020). Prediksi Kunjungan Wisatawan Taman Nasional Gunung Merbabu dengan Time Series Forecasting dan LSTM. *Jurnal Teknik Sipil*, 11(2), 132–140. <http://ejournal.uajy.ac.id/id/eprint/26642>
- Manurung, B. A., Gea, A., & Silalahi, A. P. (2024). Penerapan Algoritma Regresi Linear Untuk Memprediksi Jumlah Wisatawan. *METHOSISFO: Jurnal Ilmiah Sistem Informasi*, 4(1), 1–9. <https://ejournal.methodist.ac.id/index.php/methosisfo/article/view/2850>
- Mu'minin, F., Fauziah, F., & Gunaryati, A. (2022). Prediksi Kunjungan Wisatawan Mancanegara Melalui Pintu Udara Menggunakan ARIMA, Glnmet, dan Prophet. *Techno. Com*, 21(1), 149–157. <https://publikasi2.dinus.ac.id/index.php/technoc/article/view/5695>
- Nawawi, A., Herawati, S., & Prastiti, N. (2021). Implementasi metode holt winter additive untuk prediksi kunjungan wisatawan nusantara kabupaten sumenep. *Jurnal SimanteC*, 10(1), 25–30. <https://journal.trunojoyo.ac.id/simantec/article/view/12466>
- Putri, O., Poningsih, P., & Tambunan, H. S. (2021). Prediksi kunjungan wisatawan mancanegara ke indonesia menggunakan jaringan saraf tiruan dengan algoritma backpropagation. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer Dan Manajemen)*, 2(1), 1–7. <http://tunasbangsa.ac.id/pkm/index.php/kesatria/article/view/51>
- Rais, A., Rousyati, R., Thira, I., Kholifah, D., Purwanti, N., & Kristania, Y. (2020). Evaluasi Metode Forecasting pada Data Kunjungan Wisatawan Mancanegara ke Indonesia. *EVOLUSI: Jurnal Sains Dan Manajemen*, 8(2). <https://ejournal.bsi.ac.id/ejournal/index.php/evolusi/article/view/8971>
- Sari, M., & Agustini, D. (2024). Model Prediksi Kunjungan Wisata: Mengoptimalkan Arsitektur Algoritma Backpropagation untuk Prediksi Kunjungan Wisata Mancanegara (ASIA). *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer Dan Manajemen)*, 5(1), 240–244. <http://www.pkm.tunasbangsa.ac.id/index.php/kesatria/article/view/332>
- Sovia, R., Yanto, M., & Melati, P. (2020). prediksi jumlah kunjungan wisata mancanegara dengan algoritma Backpropagation. *Jurnal Media Informatika Budidarma*, 4(2), 355–362. <http://ejournal.stmik-budidarma.ac.id/index.php/mib/article/view/2048>
- Sujarweni, W. V., & Utami, L. R. (2019). The Master Book of SPSS. In *Start Up*.
- Utami, R., & Maulana, M. W. I. (2020). Visualisasi Prediksi Kunjungan Wisatawan Mancanegara Menggunakan Model Time Series. *Joutica*, 5(2), 356–362. <https://jurnalteknik.unisla.ac.id/index.php/informatika/article/view/436>
- Yap, B. W., Rani, K. A., Rahman, H. A. A., Fong, S., Khairudin, Z., & Abdullah, N. N. (2014). An Application of Oversampling, Undersampling, Bagging and Boosting in Handling Imbalanced Datasets BT - Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013). *Lecture Notes in Electrical Engineering*.