

Pendekatan Hybrid Statistik dan Machine Learning untuk Peramalan Jumlah Kunjungan Turis

Henny Leidiyana^{1*}, Siti Nurajizah²

^{1,2}Program Studi Informatika/Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika

^{1,2}Program Studi Sistem Informasi Akuntansi/Fakultas Teknik dan Informatika, Universitas Bina Sarana Informatika

*email: henny.hnl@bsi.ac.id

DOI: <https://doi.org/10.31603/komtika.v9i2.14909>

Received: 26-09-2025, Revised: 10-11- 2025, Accepted: 14-11-2025

ABSTRACT

Tourist arrival forecasting is a crucial aspect of planning and decision-making in the tourism sector. Accurate predictions are essential to anticipate surges or declines in visitor numbers, design effective marketing strategies, and manage resources efficiently. This study proposes a hybrid forecasting approach that integrates traditional statistical methods with machine learning algorithms to improve the accuracy of tourist arrival forecasts. Five forecasting models are implemented: ARIMA as a representative of traditional statistical models; Random Forest and Extreme Gradient Boosting (XGBoost) as machine learning models; a simple hybrid model, which combines ARIMA and XGBoost predictions through simple averaging; a weighted hybrid model, which merges the two models using performance-based weights; and a stacking hybrid model, which utilizes a meta-model to optimize prediction combinations. Given that the dataset exhibits significant pattern changes, or structural breaks, particularly during the COVID-19 pandemic, this study employs a rolling window backtesting approach for model evaluation. This method allows the models to be tested progressively across normal, crisis, and recovery periods, providing a realistic assessment of their performance under dynamic conditions. Model performance is evaluated using three key metrics: RMSE, MAE, dan MAPE. The results demonstrate that the stacking hybrid model consistently achieves the lowest RMSE across all test periods, highlighting its ability to capture complex trends and extreme fluctuations caused by COVID-19

Keywords: Rolling Window Backtesting, Weighted Hybrid, Stacking Hybrid.

ABSTRAK

Peramalan jumlah kunjungan turis merupakan salah satu aspek penting dalam perencanaan dan pengambilan keputusan di sektor pariwisata. Ketepatan prediksi sangat dibutuhkan untuk mengantisipasi lonjakan atau penurunan kunjungan, merancang strategi pemasaran, serta mengelola sumber daya secara efisien. Penelitian ini mengusulkan pendekatan Hybrid yang mengintegrasikan metode statistik tradisional dengan algoritma *machine learning* untuk meningkatkan kinerja peramalan jumlah kunjungan turis. Lima jenis model peramalan diterapkan, yaitu ARIMA sebagai representasi model statistik tradisional, Random Forest dan Extreme Gradient Boosting (XGBoost) sebagai model *machine learning*, Hybrid model sederhana yang menggabungkan prediksi ARIMA dan XGBoost melalui rata-rata (*simple averaging*), *Weighted Hybrid* yang menggabungkan kedua model dengan bobot berbasis performa historis, serta *Stacking Hybrid* yang memanfaatkan meta-model untuk mengoptimalkan kombinasi prediksi. Penelitian ini menggunakan metode evaluasi *rolling window backtesting* karena dataset yang digunakan memiliki perubahan pola yang signifikan atau structural break, seperti yang terjadi selama pandemi COVID-19. Dengan metode ini memungkinkan model diuji secara bertahap dari periode normal ke periode krisis dan pemulihan. Evaluasi performa dilakukan menggunakan tiga metrik utama, yaitu RMSE, MAE, dan MAPE. Hasil menunjukkan bahwa *Stacking Hybrid* secara konsisten memiliki nilai RMSE terendah, membuktikan kemampuannya menangkap tren dan fluktuasi ekstrem selama COVID-19.

Keywords: Rolling Window Backtesting, Weighted Hybrid, Stacking Hybrid

PENDAHULUAN

Pariwisata merupakan salah satu sektor penting dalam mendukung pertumbuhan ekonomi nasional melalui peningkatan devisa, penyerapan tenaga kerja, dan pemberdayaan masyarakat [1]. Destinasi wisata yang dikelola dengan baik dapat menjadi daya tarik bagi wisatawan mancanegara serta memperkuat citra positif suatu negara di kancah global [2]. Daya tarik pariwisata berperan besar dalam mendorong jumlah kunjungan wisatawan dari berbagai negara, sehingga dibutuhkan kebijakan dan program pengembangan pariwisata yang terarah [3].

Dalam pengelolaan sektor pariwisata, kemampuan untuk memprediksi jumlah kunjungan wisatawan merupakan komponen penting dalam perencanaan kebijakan dan pengambilan keputusan strategis. Peramalan yang akurat dapat membantu pemerintah maupun pelaku usaha dalam mengantisipasi perubahan permintaan wisata, kapasitas layanan, serta dampak fluktuasi musiman terhadap perekonomian daerah [4].

Pendekatan statistik klasik seperti *Autoregressive Integrated Moving Average* (ARIMA) dan *Holt-Winters Exponential Smoothing* telah banyak digunakan untuk peramalan data deret waktu jumlah wisatawan karena mampu menangkap pola tren dan musiman secara efisien [5]. Namun, pola data kunjungan wisatawan sering kali mengalami perubahan mendadak atau struktural (*structural break*) akibat peristiwa besar seperti pandemi COVID-19, yang menyebabkan ketidakstabilan pola data. Kondisi ini membuat model statistik tradisional cenderung kurang adaptif terhadap dinamika baru karena asumsi stasioneritas dan linearitas tidak selalu terpenuhi [6].

Untuk mengatasi keterbatasan tersebut, berbagai penelitian terkini memanfaatkan pendekatan *machine learning* (ML) yang mampu mempelajari pola nonlinier dan beradaptasi terhadap perubahan karakteristik data. Model seperti Support Vector Regression (SVR), Random Forest (RF), dan Artificial Neural Network (ANN) telah diterapkan dalam peramalan wisata untuk menangani pola yang berubah secara signifikan [7] [8]. Model berbasis ensemble learning seperti XGBoost mampu menyesuaikan bobot antarperiode ketika terjadi pergeseran pola data [9]. Meskipun demikian, setiap model memiliki keterbatasan. Model statistik lebih baik dalam menangani data musiman dan jangka pendek, sedangkan model *machine learning* unggul dalam menangkap pola nonlinier jangka panjang. Oleh karena itu, pendekatan *hybrid forecasting* dikembangkan untuk menggabungkan keunggulan keduanya, di mana model statistik digunakan untuk menangani komponen tren dan musiman, sementara *machine learning* menangani residu atau komponen nonlinier [10].

Dengan mempertimbangkan adanya perubahan pola signifikan (*structural break*) pada data kunjungan wisatawan, pendekatan Hybrid dipilih karena lebih fleksibel terhadap ketidakstabilan pola data. Oleh karena itu, penelitian ini menerapkan metode hybrid antara model statistik dan *machine learning* untuk meningkatkan akurasi peramalan jumlah kunjungan wisatawan mancanegara, khususnya dalam kondisi pasca-pandemi di mana pola data tidak lagi mengikuti tren historis yang stabil [11].

METODE

Penelitian ini dilakukan melalui enam tahap seperti pada Gambar 1, yaitu pengumpulan data jumlah kunjungan wisatawan mancanegara dari periode 2017 hingga 2024 yang diperoleh dari sumber resmi seperti Badan Pusat Statistik (BPS) atau Kementerian

Pariwisata. Data ini menjadi dasar dalam proses analisis dan pelatihan model. Tahap kedua adalah *preprocessing* dan *feature engineering*, yang meliputi identifikasi dan penanganan nilai hilang (*missing values*), transformasi data dari bentuk lebar (*wide format*) menjadi bentuk panjang (*long format*), penambahan fitur baru seperti Bulan_num dan Periode, serta pengujian stasioneritas menggunakan *Augmented Dickey-Fuller* (ADF) Test.

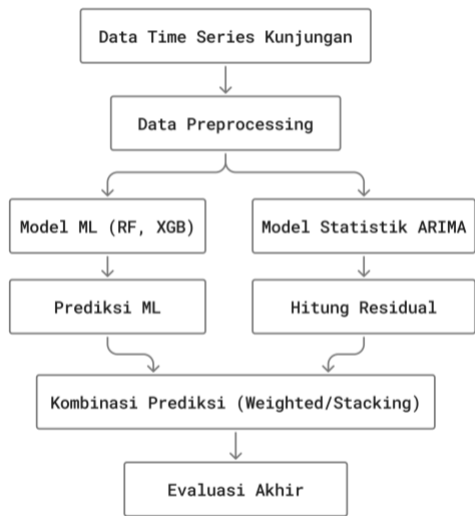


Gambar 1. Tahapan Penelitian Peramalan Jumlah Kunjungan Turis

Tahap ketiga yaitu pemodelan dengan pendekatan hybrid yang menggabungkan model statistik dan *machine learning* dalam menangani data deret waktu. Model statistik yang digunakan adalah ARIMA yang berfungsi untuk menangkap komponen linier, tren, dan musiman pada data. Setelah diperoleh hasil prediksi ARIMA ($\hat{Y}_{stat}$), dilakukan perhitungan residu ($e_t = Y_t - \hat{Y}_{stat}$) yang merepresentasikan komponen nonlinier yang tidak dapat dijelaskan oleh model statistik. Residu tersebut kemudian dimodelkan menggunakan dua algoritma *machine learning*, yaitu *Random Forest* (RF) dan *Extreme Gradient Boosting* (XGBoost), yang mampu mempelajari hubungan nonlinier dan interaksi kompleks antarwaktu. Hasil pemodelan residu dari masing-masing algoritma digunakan untuk membentuk dua model hybrid dasar, yaitu Hybrid-1 (ARIMA+RF) dengan rumus $\hat{Y}_{hyb1} = \hat{Y}_{stat} + \hat{e}_{RF}$ dan Hybrid-2 (ARIMA+XGB) dengan rumus $\hat{Y}_{hyb2} = \hat{Y}_{stat} + \hat{e}_{XGB}$. Dengan demikian, model hybrid dasar ini menggabungkan hasil peramalan dari ARIMA dengan prediksi residu dari model *machine learning* untuk mendapat hasil yang lebih akurat. Tahap berikutnya pembentukan model ensemble lanjutan *Weighted Hybrid* dan *Stacking Hybrid*. Pendekatan *Weighted Hybrid* menggabungkan hasil prediksi dari ketiga model dasar (ARIMA, RF, dan XGB) menggunakan bobot proporsional terhadap tingkat akurasi masing-masing model. Bobot (w_i) dihitung berdasarkan nilai kebalikan dari RMSE setiap model sehingga model dengan kesalahan prediksi lebih kecil akan memiliki bobot lebih besar. Formula perhitungan prediksi akhir ditulis sebagai: $\hat{Y}_{weighted} = w_1\hat{Y}_{ARIMA} + w_2\hat{Y}_{RF} + w_3\hat{Y}_{XGB}$. Sementara itu, pendekatan *Stacking Hybrid* menggunakan *meta-learner* (dalam hal ini linear regression) untuk mempelajari hubungan terbaik antarhasil prediksi dari ketiga model dasar. Hasil prediksi ARIMA, RF, dan XGB dijadikan sebagai variabel input ($X_{meta} = [\hat{Y}_{ARIMA}, \hat{Y}_{RF}, \hat{Y}_{XGB}]$), sedangkan nilai aktual Y_t menjadi target pelatihan.

Tahap keempat adalah *rolling window backtesting* untuk menguji reliabilitas model dengan pendekatan *time-series cross-validation*. Tahap kelima yaitu evaluasi model, dilakukan dengan menggunakan tiga metrik yaitu *Root Mean Square Error* (RMSE), *Mean Absolute Error* (MAE), dan *Mean Absolute Percentage Error* (MAPE), dengan RMSE

dijadikan acuan utama karena lebih sensitif terhadap kesalahan besar. Tahap keenam adalah analisis dan rekomendasi model terbaik berdasarkan hasil evaluasi di mana model dengan akurasi tertinggi khususnya *Stacking Hybrid* direkomendasikan sebagai sistem peramalan yang paling efektif untuk mendukung perencanaan kebijakan pariwisata di masa mendatang.



Gambar 2. Flowchart Model Hybrid

HASIL DAN PEMBAHASAN

Dataset yang digunakan dalam penelitian ini diperoleh dari BPS yang berisi data runtun waktu bulanan jumlah wisatawan mancanegara ke Indonesia yang dikelompokkan berdasarkan negara asal wisatawan. Periode data yang digunakan dari tahun 2017-2025 dan akan digabung dalam satu tabel untuk peramalan kunjungan turis per negara. Tabel 1 adalah sample dataset tahun 2017 dimana data tersebut memerlukan pembersihan dan transformasi.

Tabel 1. Sample Dataset Kunjungan Turis 2017

		Jumlah Kunjungan Wisatawan Mancanegara per bulan Menurut Kebangsaan (Kunjungan)						
		2017						
		Januari	Februari	Maret	April	...	Desember	Tahunan
1	Brunei Darussalam	1713	1440	2591	1993	...	3421	-
2	Malaysia	149488	142045	172097	171536	...	228056	-
3	Philippines	24536	23163	24054	27855	...	27439	-
4	Singapore	118433	91255	135441	135619	...	205706	-
5	Thailand	8677	8785	9673	14420	...	11182	-
6	Vietnam	5062	5217	5423	6323	...	6055	-
7	Laos	255	229	237	334	...	350	-
8	Kamboja	466	452	417	692	...	444	-
9	Myanmar/Burma	3817	3244	3546	5762	...	4434	-
10	Indonesia	22339	16982	20553	19733	...	20386	-
11	A S E A N	334786	292812	374032	384267	...	507473	-
...		...	...	...	...	...	...	...

Tahap *preprocessing* dilakukan untuk memastikan bahwa data yang digunakan dalam proses peramalan memiliki kualitas tinggi, bebas dari nilai agregat, kesalahan, maupun ketidakkonsistenan. Langkah pertama adalah melakukan penghapusan entitas yang bersifat agregat, yaitu negara atau wilayah yang merupakan gabungan dari beberapa negara lain dan bukan unit individu. Beberapa entitas yang dihapus antara lain ASEAN, TOTAL ASIA (Excl. ASEAN), TOTAL MIDDLE EAST, TOTAL EUROPE, TOTAL AMERICA, TOTAL OCEANIA, TOTAL AFRICA, dan GRAND TOTAL.

Sebelum proses penghapusan, total data observasi berjumlah 2.241 baris, dan setelah penghapusan agregat, jumlah data berkurang menjadi 2.169 baris. Langkah berikutnya adalah penanganan nilai hilang (*missing values*) lalu dilakukan pemeriksaan total kunjungan per negara untuk mengidentifikasi entitas dengan jumlah kunjungan yang tidak representatif. Ditemukan enam negara dengan total kunjungan sama dengan nol, yaitu Alaska, Cabinda, Guam, Siberia, Walis & Futuna, dan Zomba. Enam negara tersebut dihapus dari dataset karena tidak memberikan informasi berarti terhadap pola kunjungan wisatawan. Sementara itu, terdapat 22 negara dengan total kunjungan sangat kecil (kurang dari 100 tetapi lebih dari 0) yang dianggap tidak cukup signifikan untuk diperlakukan sebagai entitas mandiri. Negara-negara kecil tersebut kemudian digabungkan ke dalam satu kategori baru bernama “UNO” (*Unspecified or Other Nations*). Negara-negara yang digabungkan ke kategori UNO meliputi Bermuda, British Indian Ocean Territory, Burkinabe, Cook Islands, Falkland Islands, Guadeloupe, Guiana, Guinea Equatorial, Kaledonia, Martinique, Maseru, Niue, Polynesia, Puerto Rico, Sahara, Sao Tome & Principe, Scotlandia, St. Christopher, Stateless, Volta Atas, Wales, Yugoslavia, dan Zaire. Jumlah negara awalnya 241 sekarang menjadi 213. Setelah seluruh proses pembersihan, penghapusan, dan penggabungan selesai dilakukan, jumlah keseluruhan data yang siap digunakan untuk tahap selanjutnya menjadi 2.115 baris.

Tabel 2. Hasil Penghapusan Baris Agregat

Negara	Januari	Februari	Maret	...	Desember	Tahunan	Tahun
Brunei	1713	1440	2591	...	3421	23455	2017
Darussalam							
Malaysia	149488	142045	172097	...	228056	2121888	2017
Philippines	24536	23163	24054	...	27439	308977	2017
Singapore	118433	91255	135441	...	205706	1554119	2017
Thailand	8677	8785	9673	...	11182	138235	2017
...	...	...	...	...	...	...	...

Penambahan fitur Bulan_num dan Periode untuk membantu model ML Random Forest dan XGBoost dalam mengenali pola musiman (*seasonality*) dan memahami urutan waktu (*temporal ordering*). Dataset ini berbentuk tabel yang lebar (*wide format*) dimana banyak kolom tahun yang sebenarnya mewakili hal sama sehingga perlu diubah ke dalam bentuk yang panjang (*long format*). agar struktur data jadi lebih konsisten, lebih mudah dianalisis, divisualisasikan, dan diolah dengan berbagai metode statistik

Tabel 3. Dataset *Long Format*

No	Negara	Tahun	Bulan	Jumlah	Bulan_num	Periode
0	Afganistan	2017	Januari	14	1	01-01-17
1	Afganistan	2017	Februari	33	2	01-02-17
2	Afganistan	2017	Maret	73	3	01-03-17
...	...	...	...	...	...	...
25377	Zimbabwe	2025	Oktober	0	10	01-10-25
25378	Zimbabwe	2025	November	0	11	01-11-25
25379	Zimbabwe	2025	Desember	0	12	01-12-25

Transformasi dari *wide format* ke *long format* menyebabkan replikasi baris berdasarkan variabel waktu yang sebelumnya disimpan sebagai kolom sehingga jumlah baris menjadi jauh lebih banyak. Setiap kombinasi “negara–bulan” menjadi satu baris. Hasilnya: 25.379 baris = 213 negara × 119 bulan (2017–2024). Dalam penelitian ini dipilih lima negara dengan jumlah kunjungan tertinggi. Banyak studi pariwisata juga menerapkan pendekatan serupa dengan hanya meneliti pasar utama *atau top origin countries* karena tersebut paling relevan terhadap perencanaan kebijakan pariwisata, pemasaran destinasi, serta pengembangan infrastruktur pendukung [13].

Tabel 4. Lima Negara dengan Total Kunjungan Turis Tertinggi

Negara	Tahunan
Malaysia	14451198
Singapore	9113046
China	8753728
Other Asia	8333810
Australia	7962408

Berdasarkan data tahunan kunjungan wisatawan mancanegara dari tahun 2017-2025, terlihat Malaysia menempati posisi teratas dengan kunjungan sebesar 14.451.198. Gambar 3 grafik asli menunjukkan tren jumlah turis dari lima negara dengan kunjungan tertinggi ke Indonesia, yaitu Australia, China, Malaysia, Singapura, dan kelompok Other Asia dalam periode 2017 hingga 2025. Secara umum, terlihat pola fluktuasi yang signifikan pada semua negara, dengan penurunan tajam antara tahun 2020–2021. Penurunan ini merupakan dampak dari pandemi COVID-19, yang menyebabkan pembatasan perjalanan internasional dan penurunan drastis jumlah wisatawan mancanegara. Setelah periode tersebut tampak tanda pemulihan kunjungan wisatawan meskipun belum sepenuhnya kembali ke level sebelum pandemi. Fenomena ini menciptakan anomali data atau *structural break*, di mana pola historis sebelum 2020 tidak sepenuhnya relevan untuk memprediksi periode saat pandemi maupun pasca-pandemi. Menurut Hyndman & Athanasopoulos [14], *structural break* seperti ini dapat mengurangi akurasi model time series jika tidak ditangani dengan tepat. Untuk mengatasinya penelitian ini menggunakan pendekatan *time-series cross-validation* dengan *rolling window backtesting*. Pendekatan ini bekerja dengan cara memperluas periode pelatihan (*training set*) secara bertahap sehingga model dapat belajar dari data sebelum pandemi, diuji pada masa pandemi, dan dievaluasi kembali pada masa pemulihan pasca-pandemi[15]

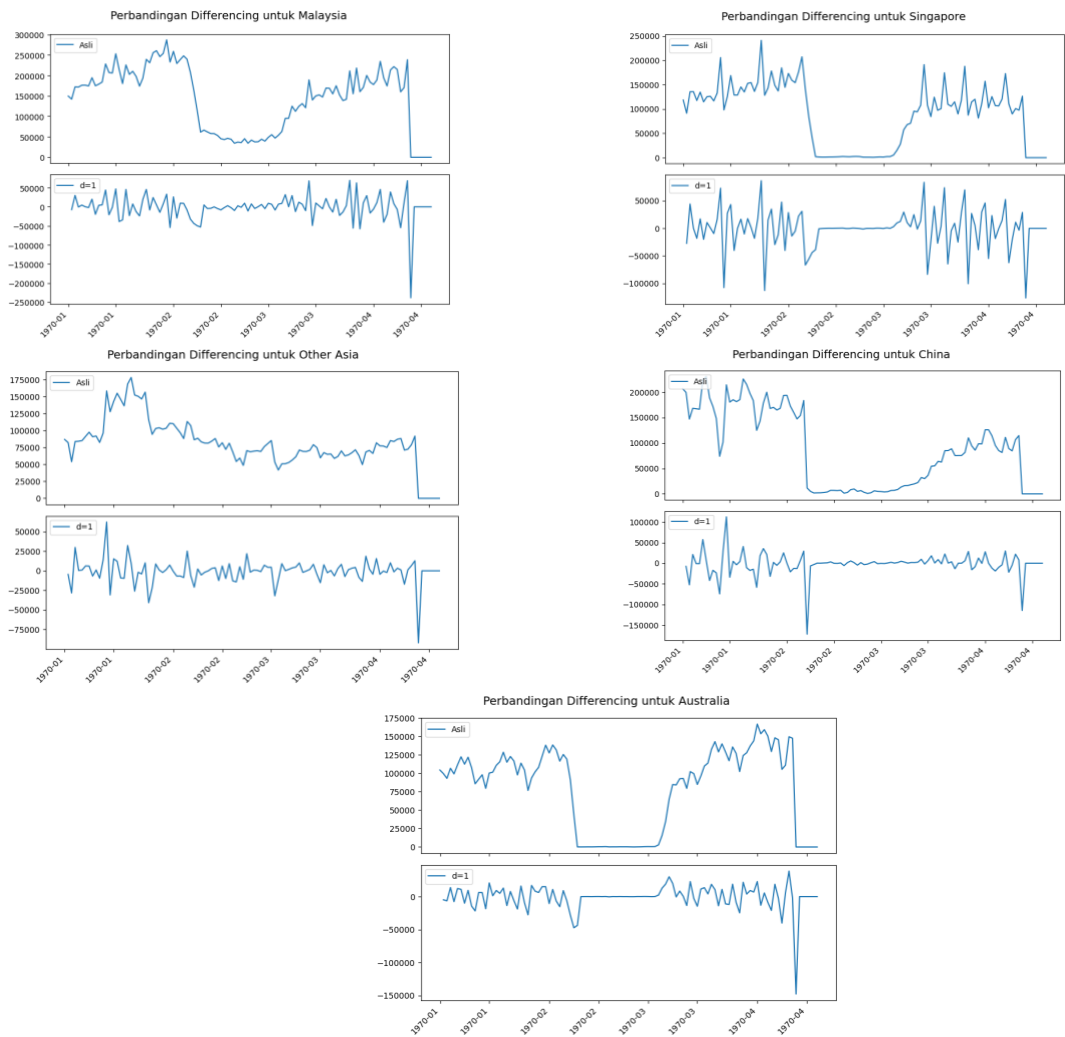
Sebelum pemodelan, dilakukan uji stasioneritas data menggunakan ADF Test untuk mendeteksi keberadaan akar unit pada data deret waktu. Akar unit merupakan kondisi di mana nilai suatu variabel sangat dipengaruhi oleh nilai sebelumnya secara permanen sehingga

perubahan kecil di masa lalu, misalnya penurunan sementara jumlah kunjungan akibat faktor musiman, kebijakan visa, atau pembatasan perjalanan seperti pada masa pandemi COVID-19, dapat memberikan dampak jangka panjang terhadap pola kunjungan pada periode berikutnya [16]. Dengan kata lain, data yang memiliki akar unit bersifat tidak stasioner, karena nilai rata-rata dan variannya berubah seiring waktu [17].

Tabel 5. Hasil Uji ADF

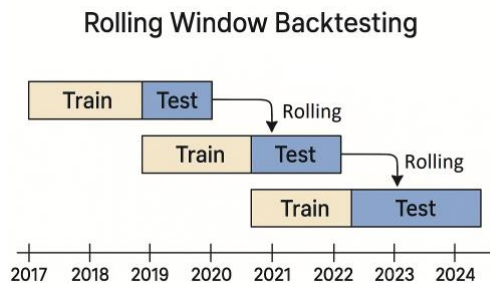
Negara	ADF Statistic	p-value	Kesimpulan
Malaysia	-2.199381	0.20647	Tidak stasioner
Singapore	-1.953578	0.307291	Tidak stasioner
Other Asia	-1.786738	0.387065	Tidak stasioner
China	-2.171816	0.216633	Tidak stasioner
Australia	-1.6367	0.463986	Tidak stasioner

Berdasarkan hasil uji ADF pada Tabel di atas, seluruh negara memiliki p-value > 0.05, sehingga tidak stasioner pada level orde-0 ($d=0$). Parameter $d=0$ menunjukkan bahwa data masih pada level aslinya tanpa dilakukan proses *differencing*. Dengan demikian maka dilakukan *differencing* orde-1 ($d=1$) dengan menghitung selisih antara nilai saat ini dan nilai sebelumnya untuk menghilangkan tren linier.



Gambar 3. Pola Kunjungan Lima Negara Sebelum dan Sesudah *Differencing*

Proses *differencing* orde-1 pada penelitian ini cukup untuk menstabilkan rata-rata data deret waktu tanpa menghilangkan informasi penting mengenai pola musiman dan fluktuasi (Box et al., 2016). Dengan demikian, data hasil *differencing* menjadi lebih mendekati stasioner dan siap digunakan dalam pemodelan ARIMA sedangkan model ML (RF dan XGB) tidak memerlukan prasyarat stasioneritas sehingga tetap menggunakan data asli.



Gambar 4. Skema *Rolling Window Backtesting*

Berdasarkan Gambar 4, pada periode pertama data tahun 2017–2019 digunakan sebagai *training set*, sedangkan tahun 2020 digunakan sebagai *testing set*. Model dilatih menggunakan data 2017–2019 kemudian diuji untuk memprediksi data pada tahun 2020. Pada langkah berikutnya, jendela pengujian (*rolling window*) digeser ke depan sehingga *training set* mencakup data 2017–2020 dan *testing set* menjadi tahun 2021. Proses ini dilakukan secara berulang, misalnya train 2017–2021 untuk memprediksi 2022, dan seterusnya hingga tahun 2024. Pendekatan ini memungkinkan model untuk terus belajar dari data terbaru dan mengevaluasi performa prediksi pada periode yang belum pernah dilihat sebelumnya.

Sebelum proses *rolling window backtesting*, data ditransformasi dari format lebar (*wide format*) menjadi format panjang (*long format*), di mana setiap baris mewakili kombinasi negara dan periode waktu (bulan). Hasil transformasi yaitu 25.379 observasi dari 213 negara selama periode 2017–2024. Jumlah ini menjadi keseluruhan dataset yang digunakan dalam pembangunan model, yang kemudian dibagi secara dinamis ke dalam subset pelatihan dan pengujian sesuai dengan mekanisme *rolling window* di atas. Setelah seluruh proses *preprocessing* selesai, tahap berikutnya adalah pembangunan model peramalan.

Hasil pemodelan menunjukkan bahwa model ARIMA mampu menangkap pola tren dan musiman, tetapi kurang responsif terhadap fluktuasi tajam akibat faktor eksternal seperti pandemi COVID-19. Sebaliknya, model ML (RF dan XGB) lebih adaptif terhadap perubahan mendadak, namun cenderung kehilangan konsistensi pada pola jangka panjang. Untuk mengatasi keterbatasan masing-masing, dibentuklah model Hybrid, yaitu penggabungan ARIMA dengan RF atau XGB berdasarkan pola residu. Dari proses ini diperoleh dua model hybrid dasar, yakni ARIMA+RF dan ARIMA+XGB, yang kemudian dikembangkan lebih lanjut menjadi model *Weighted Hybrid* dan *Stacking Hybrid*.

Weighted Hybrid menggabungkan hasil prediksi ARIMA, RF, dan XGB dengan bobot berdasarkan performa model (RMSE). Pendekatan ini memberikan kontribusi lebih besar pada model yang memiliki kesalahan prediksi kecil, sehingga hasil akhirnya lebih stabil. Sementara itu, *Stacking Hybrid* menggunakan meta-learner berbasis regresi linier untuk

mempelajari hubungan optimal antarhasil prediksi ketiga model dasar. Pendekatan ini memungkinkan model beradaptasi secara dinamis terhadap perubahan pola data.

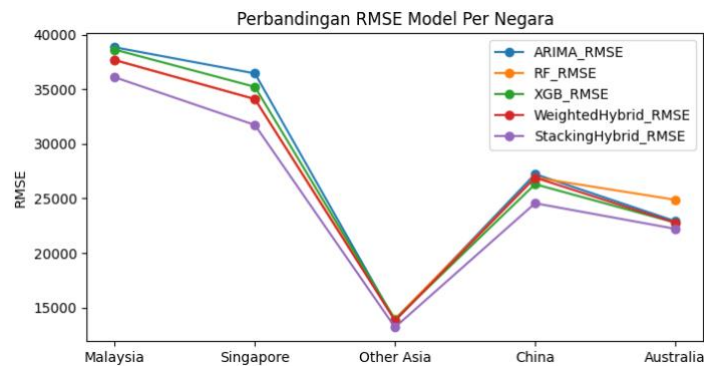
Berdasarkan hasil evaluasi pada lima negara, diperoleh hasil seperti ditunjukkan pada Tabel 6. Nilai RMSE menunjukkan bahwa model *Stacking Hybrid* secara konsisten menghasilkan kesalahan prediksi terendah dibandingkan model individual maupun *Weighted Hybrid*. Sebagai contoh, pada Malaysia, RMSE *Stacking Hybrid* sebesar 0.331 lebih rendah daripada ARIMA (0.356) maupun model *machine learning* tunggal. Hasil serupa ditemukan pada negara lain dengan pola fluktuasi serupa. Temuan ini mengindikasikan bahwa *Stacking Hybrid* mampu menangkap karakteristik data deret waktu yang kompleks dan nonlinier, serta lebih adaptif terhadap *structural break* yang muncul akibat pandemi. Kombinasi komponen linier (ARIMA) dan nonlinier (RF, XGB) yang dioptimalkan melalui meta-learner menjadikan model ini unggul dalam mengurangi kesalahan prediksi. Dengan demikian, *Stacking Hybrid* direkomendasikan sebagai model terbaik untuk peramalan jumlah kunjungan wisatawan mancanegara ke Indonesia karena memberikan hasil paling akurat, stabil, dan adaptif terhadap perubahan pola data dari tahun 2017 hingga 2024.

Evaluasi performa model dilakukan menggunakan tiga metrik utama, yaitu RMSE, MAE, dan MAPE. Dari ketiganya, RMSE dijadikan metrik utama karena memberikan penalti yang lebih besar terhadap kesalahan prediksi yang besar sehingga lebih sensitif terhadap outlier [18].

Tabel 6. Nilai RMSE Tiap Model

Negara	ARIMA	RF	XGB	WeightedHybrid	StackingHybrid
Malaysia	0.356	0.346	0.354	0.346	0.331
Singapore	0.599	0.561	0.579	0.561	0.522
Other Asia	0.214	0.217	0.216	0.214	0.205
China	0.663	0.656	0.641	0.656	0.598
Australia	0.353	0.383	0.351	0.351	0.342

Berdasarkan hasil evaluasi model menggunakan metrik RMSE pada Gambar 6 terlihat bahwa *Stacking Hybrid* unggul pada seluruh negara yang diteliti. Untuk Malaysia, RMSE *Stacking Hybrid* (0.331) lebih rendah dibandingkan ARIMA (0.356) dan model *machine learning* lainnya. Hasil serupa ditemukan di Singapore (0.522), Other Asia (0.205), China (0.598), dan Australia (0.342), di mana nilai RMSE dari *Stacking Hybrid* konsisten lebih kecil dibandingkan model individu maupun pendekatan *Weighted Hybrid*. Temuan ini menunjukkan model *stacking* mampu menangkap pola musiman dan tren yang kompleks dalam data kunjungan wisatawan, sekaligus mengurangi kelemahan yang dimiliki masing-masing model dasar. Dengan demikian, dapat disimpulkan bahwa *Stacking Hybrid* merupakan metode yang paling efektif dan akurat untuk peramalan kunjungan wisatawan internasional dalam periode penelitian 2017–2024.



Gambar 6. Perbandingan RMSE Model Tiap Negara

KESIMPULAN

Hasil evaluasi menunjukkan bahwa *Stacking Hybrid* konsisten memberikan nilai RMSE terendah pada seluruh negara (Malaysia, Singapore, Other Asia, China, Australia). Model ini mampu mengombinasikan keunggulan ARIMA, Random Forest, dan XGBoost sehingga menghasilkan prediksi yang lebih akurat dibandingkan model individu maupun hybrid berbobot. Dengan demikian, *Stacking Hybrid* direkomendasikan sebagai model terbaik untuk peramalan jumlah kunjungan wisatawan. Rekomendasi penelitian selanjutnya adalah mengembangkan model hybrid yang lebih adaptif, misalnya mengombinasikan deep learning (LSTM/GRU) dengan boosting models (XGBoost, LightGBM), atau menerapkan dynamic weighting yang dapat menyesuaikan bobot model seiring perubahan pola data. Hal ini diharapkan mampu mengatasi ketidakstabilan pola kunjungan turis akibat faktor eksternal di masa depan.

DAFTAR PUSTAKA

- [1] PERPRES, *Peraturan Presiden (Perpres) Nomor 84 Tahun 2021 tentang Rencana Induk Destinasi Pariwisata Nasional Lombok-Gili Tramena Tahun 2020-2044*. 2021.
- [2] J. Bravo, R. Alarcón, C. Valdivia, and O. Serquén, "Application of Machine Learning Techniques to Predict Visitors to the Tourist Attractions of the Moche Route in Peru," *Sustain.*, vol. 15, no. 11, 2023, doi: 10.3390/su15118967.
- [3] R. Hidayat and B. Helmi Mustawinar, "Peramalan Jumlah Wisatawan Asing Dengan Model ARIMA," *Infin. J. Mat. dan Apl.*, vol. 2, no. 2, pp. 104–115, 2022.
- [4] D. Rizkya, H. Roosaputri, and C. Dewi, "Perbandingan Algoritma ARIMA, Prophet, dan LSTM dalam Prediksi Penjualan Tiket Wisata Taman Hiburan (Studi Kasus: Saloka Theme Park)," *J. Penerapan Sist. Infomatika (Komputer Manajemen)*, vol. 4, no. 3, pp. 507–517, 2023.
- [5] S. Soraya, I. Fitriana Aziza, M. Rizky, U. Juanda, G. Primajati, and P. Rahima, "Peramalan Jumlah Kunjungan Wisatawan di Provinsi Nusa Tenggara Barat (NTB) Menggunakan Metode Arima Box-Jenkins," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 6, no. 01, pp. 35–43, 2024, doi: 10.35580/variensiunm150.
- [6] J. Wang, L. Zhang, and Z. Liu, "Tourism Demand Interval Forecasting Amid COVID-19: A Hybrid Model With a Modified Multi-Objective Optimization Algorithm," no. 217, 2023, doi: 10.1177/10963480221142873.
- [7] E. Utami and A. D. Hartanto, *Sistem Basis Data Menggunakan Microsoft SQL Server*

2005. Yogyakarta: Andi Offset, 2005.
- [8] D. T. Andariesta and M. Wasesa, “Machine learning models for predicting international tourist arrivals in Indonesia during the COVID-19 pandemic: a multisource Internet data approach,” no. November, pp. 1–17, 2025, doi: 10.1108/JTF-10-2021-0239.
 - [9] N. Gunasekara, B. Pfahringer, H. Gomes, and A. Bifet, “Gradient boosted trees for evolving data streams,” *Mach. Learn.*, vol. 113, no. 5, pp. 3325–3352, 2024, doi: 10.1007/s10994-024-06517-y.
 - [10] L. B. Sina, C. A. Secco, and M. Blazevic, “Hybrid Forecasting Methods — A Systematic Review,” 2023.
 - [11] M. Cristanto and E. Mailoa, “Peramalan Jumlah Kedatangan Wisatawan Mancanegara dengan Menggunakan Model ARIMA,” *jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, 2024.
 - [12] BPS, “Jumlah Kunjungan Wisatawan Mancanegara per bulan Menurut Kebangsaan (Kunjungan).” [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/MTQ3MCMY/jumlah-kunjungan-wisatawan-mancanegara-per-bulan-menurut-kebangsaan--kunjungan-.html>
 - [13] UNWTO, “UNWTO World Tourism Barometer and Statistical Annex, November 2023,” 2023, [Online]. Available: <https://www.e-unwto.org/doi/abs/10.18111/wtobarometereng.2023.21.1.4?role=tab>
 - [14] G. Hyndman, R. J., & Athanasopoulos, *Forecasting: Principles and Practice (3rd ed.)*. OTexts, 2021.
 - [15] C. Bergmeir, R. J. Hyndman, and B. Koo, “A note on the validity of cross-validation for evaluating autoregressive time series prediction,” *Comput. Stat. Data Anal.*, vol. 120, no. April, pp. 70–83, 2018, doi: 10.1016/j.csda.2017.11.003.
 - [16] D. Gujarati, *Basic Econometrics*. New York: McGraw-Hill/Irwin, 2009.
 - [17] Walter Enders, *Applied Econometric Time Series*. NJ: John Wiley & Sons, 2015.
 - [18] X. Liu, “Tourism Demand Forecasting in Normal and Crisis Times: Combining Bootstrap-Aggregating and Bayesian Approaches,” *J. Hosp. Tour. Res.*, 2025.

