

BAB I

PENDAHULUAN

Bab ini menyajikan landasan penelitian yang mencakup latar belakang, rumusan masalah, tujuan, manfaat, dan batasan penelitian. Setiap sub-bab dirancang untuk memberikan pemahaman menyeluruh mengenai relevansi dan kontribusi penelitian sebagai landasan pembahasan selanjutnya.

1.1. Latar Belakang

Pariwisata merupakan salah satu sektor strategis yang memberikan kontribusi signifikan terhadap perekonomian Indonesia melalui penciptaan lapangan kerja, peningkatan kesejahteraan masyarakat, dan pembangunan daerah [1]-[2]. Sektor ini menunjukkan pertumbuhan yang konsisten dengan kontribusi Produk Domestik Bruto (PDB) meningkat dari 4% pada 2011 menjadi 5,5% pada 2019, disertai pertumbuhan devisa dari \$8,55 miliar menjadi \$16,91 miliar [3]. Meskipun mengalami penurunan akibat pandemi COVID-19, sektor pariwisata menunjukkan pemulihan yang kuat dengan kontribusi PDB 3,9% pada 2023, kemudian meningkat menjadi 4% pada 2024 dengan devisa mencapai \$16,7 miliar [4]. Kunjungan wisatawan mancanegara (wisman) tidak hanya menjadi indikator daya tarik Indonesia di mata dunia, tetapi juga berkontribusi langsung terhadap penerimaan devisa yang mendukung perdagangan internasional [5]-[6].

Dinamika sektor pariwisata Indonesia mengalami guncangan hebat akibat pandemi COVID-19. Berdasarkan data Badan Pusat Statistik (BPS), kunjungan wisman yang sebelumnya mencatat pertumbuhan rata-rata 9,14% per tahun selama periode 2008-2019, mengalami penurunan drastis hingga 61,82% pada periode Januari hingga November 2021 [7]-[10]. Meskipun telah memasuki fase pemulihan dengan peningkatan kunjungan dari 5,89 juta pada 2022 menjadi 13,90 juta pada 2024, angka tersebut masih belum mencapai rekor pra-pandemi sebesar 16,11 juta pada 2019 [11]. Kondisi ini mengindikasikan perlunya strategi pengelolaan yang lebih adaptif dan responsif terhadap perubahan kondisi global.

Dalam konteks pemulihan dan pengembangan sektor pariwisata, pemahaman mendalam terhadap karakteristik kunjungan wisman menjadi fundamental untuk

perencanaan strategis. Karakteristik ini mencakup pola spasial dan temporal yang dapat dianalisis dari jalur masuk dan waktu kunjungan. Wisman memiliki tiga pilihan jalur masuk ke Indonesia yaitu udara (bandara), laut (pelabuhan), dan darat (PLBN) [12]. Jalur udara secara konsisten mendominasi sebagai pintu masuk utama dengan rata-rata 64,98% hingga 67,39% dari total kunjungan [10]-[11]-[13]. Dominasi ini dipilih karena perannya yang strategis sebagai penghubung vital dalam aktivitas sosial ekonomi dan perdagangan internasional, terutama bagi negara kepulauan seperti Indonesia [14]. Sementara itu, dari aspek temporal, analisis data historis menunjukkan pola musiman yang konsisten dengan puncak kunjungan pada Juli hingga September (bersamaan dengan liburan musim panas Eropa-Amerika) dan Desember (libur Natal dan Tahun Baru) [10]-[11]. Pola-pola ini mengonfirmasi pengaruh signifikan faktor musiman dan preferensi akses transportasi terhadap fluktuasi kunjungan wisman.

Berdasarkan kompleksitas dan volatilitas yang tinggi dalam sektor pariwisata pasca-pandemi, Kementerian Pariwisata dan Ekonomi Kreatif melalui Permenparekraf No. 11 Tahun 2022 menekankan pentingnya penerapan kebijakan berbasis data dan informasi yang akurat (*data-driven policy*) [15]. Dalam konteks ini, peramalan kunjungan wisman menjadi kebutuhan strategis yang dapat dijadikan landasan bagi pemerintah dalam menyusun perencanaan dan merumuskan kebijakan sektor pariwisata yang tepat waktu dan efisien [16]. Peramalan yang akurat memungkinkan optimalisasi alokasi sumber daya, persiapan infrastruktur, dan antisipasi terhadap fluktuasi permintaan layanan pariwisata.

Tantangan dalam peramalan kunjungan wisman terletak pada kompleksitas pola data yang mengandung unsur musiman, tren jangka panjang, dan fluktuasi non-linear yang dipengaruhi faktor eksternal. Model statistik tradisional seperti *Seasonal Autoregressive Moving Average* (SARIMA) telah terbukti efektif dalam menangani pola musiman, seperti yang ditunjukkan penelitian Febrian dan Wijayanto (2024) dengan model SARIMA(0,1,1)(1,0,1)¹² yang mencapai RMSE 10,223 untuk prediksi kunjungan wisman ke Bandara Kualanamu [17]. Namun, SARIMA memiliki keterbatasan dalam menangani hubungan non-linear dan mengintegrasikan variabel eksternal yang kompleks. Di sisi lain, algoritma *machine learning* seperti *eXtreme Gradient Boosting* (XGBoost) menunjukkan keunggulan

dalam memodelkan data non-linear dan tahan terhadap *overfitting*, sebagaimana dibuktikan dalam penelitian Prayuda dan Pratama (2024) dengan RMSE 671.935,2 dan MAE 648.139,1 dalam memprediksi kedatangan wisman berdasarkan pintu masuk udara [9]. Meskipun XGBoost menunjukkan keunggulan dalam memodelkan hubungan non-linear dan tahan terhadap *overfitting*, model ini kurang optimal dalam menangkap pola musiman yang konsisten.

Kombinasi model SARIMA-XGBoost menjadi solusi yang relevan dengan memadukan keunggulan masing-masing model untuk menghasilkan peramalan yang lebih akurat. Pendekatan ini telah diterapkan dalam studi kasus berbeda oleh Man et al. (2023), yang menggunakan kombinasi SARIMA-XGBoost untuk memprediksi penyakit *Hand, Foot, and Mouth Disease* (HFMD) di Xinjiang, China. SARIMA digunakan untuk menangkap tren musiman dan pola linear, sedangkan XGBoost diterapkan pada residual SARIMA untuk menangani pengaruh non-linear dari variabel eksogen. Hasil penelitian menunjukkan bahwa model SARIMA-XGBoost mampu memberikan akurasi prediksi yang lebih tinggi dibandingkan dengan model tunggal, terutama di wilayah dengan pola kompleks dan heterogenitas spasial yang tinggi [18]. Untuk mengoptimalkan performa terbaik dari model SARIMA-XGBoost, proses *hyperparameter tuning* dilakukan menggunakan Optuna yang memanfaatkan algoritma *Tree-structured Parzen Estimator* (TPE) dan strategi *pruning* cerdas untuk mempercepat proses eksplorasi tanpa mengorbankan kualitas hasil [19].

Penelitian ini menerapkan model SARIMA-XGBoost dengan optimasi Optuna untuk meramalkan kunjungan wisman melalui jalur udara. Pendekatan ini menggabungkan kelebihan SARIMA dalam menangkap pola musiman dan tren temporal dengan kemampuan XGBoost dalam memodelkan hubungan non-linear, sehingga mengatasi keterbatasan penelitian sebelumnya yang hanya menggunakan model tunggal. Optuna digunakan untuk mengoptimalkan *hyperparameter* guna meningkatkan akurasi prediksi. Model ini diharapkan dapat memberikan prediksi yang lebih robust terhadap volatilitas data kunjungan wisman pasca-pandemi. Hasil penelitian ini diharapkan berkontribusi pada pengembangan metodologi peramalan pariwisata yang lebih akurat untuk mendukung kebijakan berbasis data di sektor pariwisata Indonesia.

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang yang telah dipaparkan sebelumnya, penelitian ini merumuskan beberapa masalah sebagai berikut:

1. Bagaimana proses implementasi model SARIMA-XGBoost dengan optimasi Optuna dalam peramalan jumlah kunjungan wisman melalui jalur udara?
2. Bagaimana pengaruh penerapan optimasi *hyperparameter* menggunakan Optuna terhadap performa model SARIMA-XGBoost?

1.3. Tujuan Penelitian

Berdasarkan uraian rumusan masalah yang telah dijelaskan, maka tujuan dari penelitian ini adalah sebagai berikut:

1. Mengimplementasikan model SARIMA-XGBoost dengan optimasi Optuna untuk meramalkan jumlah kunjungan wisman melalui jalur udara.
2. Mengevaluasi pengaruh optimasi *hyperparameter* menggunakan Optuna terhadap peningkatan performa model SARIMA-XGBoost dalam menghasilkan hasil peramalan yang lebih akurat.

1.4. Manfaat Penelitian

Berdasarkan penjabaran tujuan penelitian yang telah diuraikan, maka manfaat dari penelitian ini adalah sebagai berikut.

A. Manfaat Teoritis

1. Memberikan kontribusi bagi pengembangan ilmu pengetahuan, khususnya dalam bidang peramalan berbasis data deret waktu dengan pendekatan yang menggabungkan metode statistik dan *machine learning*.
2. Menjadi referensi dan landasan awal bagi peneliti lain yang ingin mengembangkan model peramalan kombinasi atau menerapkan pendekatan serupa pada bidang lain seperti transportasi, ekonomi, maupun sektor publik lainnya.
3. Menjadi acuan bagi praktisi data dan analisis data dalam memilih model peramalan yang sesuai dengan data deret waktu yang memiliki karakteristik musiman dan non-linear.

B. Manfaat Praktis

1. Menyediakan model peramalan yang lebih akurat untuk mendukung pengambilan kebijakan berbasis data (*data driven policy*) dalam pengembangan sektor pariwisata nasional.
2. Memberikan informasi prediktif yang dapat digunakan untuk mengoptimalkan operasional bandara, termasuk manajemen kapasitas dan sumber daya selama periode puncak kunjungan wisman.
3. Menyediakan prediksi yang lebih akurat untuk membantu pelaku usaha pariwisata dalam menyusun strategi bisnis, pemasaran, dan penyediaan layanan yang sesuai dengan fluktuasi kunjungan wisman.

1.5. Batasan Masalah

Agar penelitian ini tetap terarah dan tidak melebar dari konteks utama, ditetapkan beberapa batasan masalah sebagai berikut:

1. Penelitian ini hanya berfokus pada jumlah kunjungan wisman yang masuk ke Indonesia melalui jalur udara.
2. Data yang digunakan merupakan data deret waktu bulanan dari tahun 2008 hingga 2024 yang diperoleh dari berbagai sumber terpercaya. Data utama diperoleh dari BPS, sedangkan data eksternal seperti musim dan periode pandemi COVID-19 diperoleh dari publikasi ilmiah dan lembaga kajian ilmiah.
3. Data kunjungan wisman yang dianalisis dibatasi pada enam bandara internasional, yaitu Bandara Ngurah Rai (Bali), Soekarno-Hatta (Jakarta), Juanda (Surabaya), Kualanamu (Medan), Sam Ratulangi (Manado) dan Minangkabau (Padang). Pemilihan keenam bandara tersebut didasarkan pada ketersediaan data historis dari tahun 2008, di mana hanya enam bandara ini yang telah memiliki catatan jumlah kunjungan wisman secara konsisten sejak awal periode penelitian.