

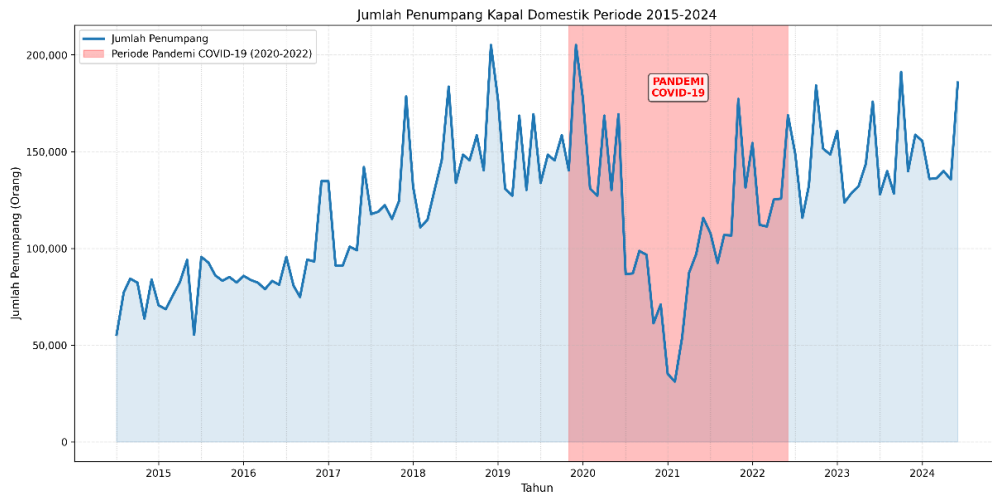
BAB I

PENDAHULUAN

A. Latar Belakang

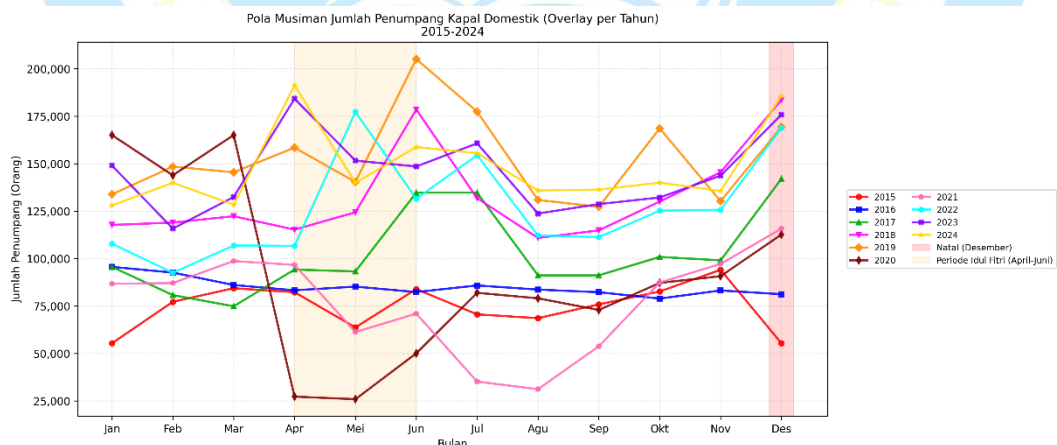
Sebagai negara kepulauan (*archipelagic state*) terbesar di dunia, dengan garis pantai sepanjang 81.000 km, yang memiliki lebih dari 17.000 pulau [1]. Indonesia sangat bergantung pada transportasi laut sebagai tulang punggung konektivitas antarwilayah, baik untuk distribusi logistik maupun mobilitas penduduk. Provinsi Kepulauan Riau merupakan salah satu provinsi dengan karakteristik kepulauan yang ekstrem, menjadikan transportasi laut menjadi kebutuhan pokok bagi masyarakat dalam mengakses layanan publik dan kegiatan ekonomi. Salah satu simpul utama transportasi laut di Provinsi Kepulauan Riau adalah Pelabuhan Sri Bintan Pura yang terletak di Tanjungpinang. Pelabuhan ini merupakan fasilitas strategis yang melayani rute pelayaran domestik antarpulau, yaitu Batam, Tanjung Balai Karimun, Lingga, Anambas, Natuna dan Riau Daratan (Dumai, Bengkalis, Selat Panjang, Tembilahan). Selain itu, Pelabuhan ini juga melayani rute pelayaran internasional, yakni ke Singapura dan Malaysia [2]. Tingginya mobilitas ini menjadikan Pelabuhan Sri Bintan Pura sebagai salah satu simpul transportasi laut tersibuk di Indonesia, yang bahkan tercatat menduduki peringkat kelima sebagai pelabuhan terpadat nasional pada periode mudik Lebaran 2026 [3].

Menurut data Badan Pusat Statistik (BPS) Provinsi Kepulauan Riau, jumlah penumpang kapal domestik di Pelabuhan Sri Bintan Pura mencapai 1.775.162 penumpang pada tahun 2024 [4], menunjukkan tren pemulihan secara bertahap dibandingkan dengan tahun sebelumnya yang mencatat 1.746.391 penumpang pada 2023 [5] dan 1.520.405 penumpang pada 2022 [6]. Sebaliknya, pada periode 2020 dan 2021, mobilitas penumpang mengalami penurunan drastis dengan angka 1.101.741 penumpang pada 2020 [7] dan menyentuh titik terendah sebesar 921.915 penumpang pada tahun 2021 [8]. Padahal, pada masa sebelum pandemi di tahun 2019, jumlah penumpang yang terdata mencapai 1.835.357 penumpang [9], meningkat dari tahun 2018 yang berada di angka 1.593.773 penumpang [10]. Berikut merupakan grafik jumlah penumpang di Pelabuhan Sri Bintan Pura dari tahun Januari 2015-Desember 2024.



Gambar 1. Jumlah Penumpang Kapal Domestik (2015-2024)

Penurunan drastis pada tahun 2020 hingga 2021 merupakan akibat langsung dari intervensi pandemi Covid-19 dan kebijakan pembatasan mobilitas. Selain dinamika historis dan dampak pandemi tersebut, sebagai pelabuhan dengan aktivitas yang padat, Pelabuhan Sri Bintan Pura dihadapkan pada tantangan dinamika jumlah penumpang yang berfluktuasi secara signifikan. Volume penumpang domestik sangat dipengaruhi oleh pola musiman (*seasonality*), seperti hari raya keagamaan seperti Hari Raya Idul Fitri dan Hari Raya Natal, yang dapat dilihat pada Gambar 2.

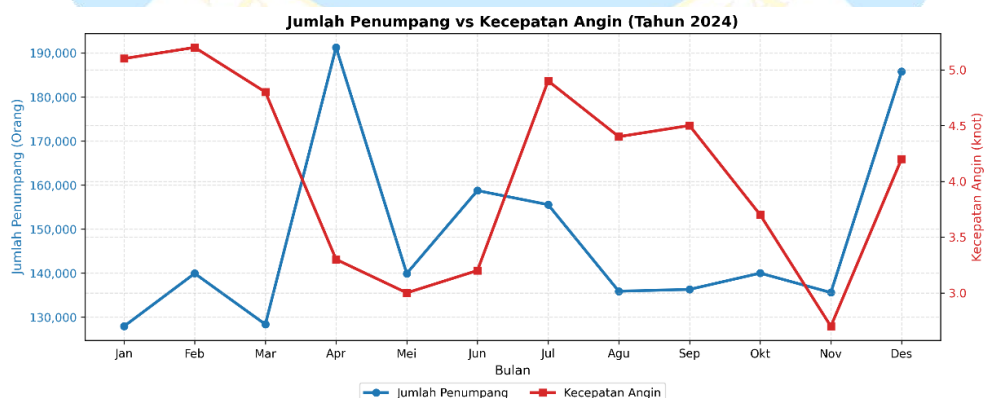


Gambar 2. Hubungan jumlah penumpang dengan hari libur (2015-2024)

Sebagai contoh, pada momentum arus mudik Hari Raya Idul Fitri tahun 2024 lalu, tercatat lonjakan penumpang bulanan yang signifikan hingga mencapai

191.232 penumpang, jauh melampaui rata rata pergerakan pada bulan operasional biasa yang hanya berkisar di angka 139.818 penumpang. Selain hari libur, kondisi meteorologi dan cuaca maritim juga berpengaruh terhadap jumlah penumpang kapal. Kondisi ekstrem yang dipicu oleh anomali pergerakan angin monsun serta tingginya kecepatan angin menjadi faktor alam yang menghambat mobilitas transportasi laut. Hal ini sejalan dengan beberapa penelitian, yang menegaskan bahwa kecepatan angin ekstrem secara langsung memaksa dihentikannya operasional pelayaran [11][12]. Lebih lanjut, fenomena ini diperkuat secara spesifik oleh penelitian yang menyimpulkan bahwa faktor utama penyebab penurunan jumlah penumpang yaitu diakibatkan oleh cuaca buruk, sehingga penumpang takut untuk berpergian dan kapal tidak berlayar [13]. Hal ini sesuai dengan kondisi di Pelabuhan Sri Bintan Pura, angin kencang, gelombang tinggi dan hujan deras membuat sejumlah kapal tujuan Dumai, Karimun, Anambas dan Batam mengalami keterlambatan berlayar [14]. Rangkaian kejadian empiris tersebut membuktikan bahwa musim angin dan kecepatan angin merupakan parameter krusial yang menyebabkan ketidakstabilan volume penumpang pelabuhan.

Sebagai contoh, pada tahun 2024, saat musim barat (Februari) dengan kecepatan angin mencapai 5,20 knot, jumlah penumpang turun menjadi 139.928 penumpang. Sebaliknya pada musim timur (April) dengan kecepatan angin hanya 3,30 knot, jumlah penumpang mencapai 191.232 penumpang. Perbedaan volume penumpang antara musim angin dan jumlah penumpang ditunjukkan pada Gambar



Gambar 3. Hubungan jumlah penumpang dengan kecepatan angin (2024)

Fluktuasi yang drastis dan ketidakpastian volume penumpang ini dapat memicu masalah operasional seperti, penumpukan penumpang di terminal keberangkatan. Untuk memitigasi risiko tersebut, pihak pengelola pelabuhan dan penyedia jasa pelayaran membutuhkan sistem perencanaan sumber daya yang didukung oleh kemampuan peramalan (*forecasting*) volume penumpang secara presisi.

Dalam bidang ilmu komputasi dan analisis deret waktu (*time series*), peramalan dapat dilakukan menggunakan berbagai pendekatan algoritma. Secara historis, pemodelan statistik seperti *Autogressive Integrated Moving Average* (ARIMA), telah menjadi standar yang diadopsi secara luas. ARIMA merupakan metode peramalan yang menghasilkan ramalan-ramalan berdasarkan sintesis dari pola data secara historis [15]. Untuk mengakomodasi karakteristik data penumpang yang memiliki tren musiman dan dipengaruhi oleh faktor eksternal, varian *Seasonal Autoregressive Integrated Moving Average with eXogenous variables* (SARIMAX) merupakan salah satu metode statistik yang paling andal. SARIMAX mengintegrasikan data historis dengan variabel eksternal untuk meningkatkan akurasi peramalan [16]. SARIMAX mampu menangkap efek musiman sekaligus mengintegrasikan variabel intervensi (eksogen), seperti penanda hari libur nasional atau kondisi cuaca, ke dalam model. Meskipun tangguh dalam interfensi statistik dan pemodelan hubungan linear, metode keluarga ARIMA memiliki keterbatasan dasar dalam mempelajari pola data yang sangat kompleks, non-linear, dan rentan terhadap *outlier* yang sering ditemui pada data transportasi dunia nyata [17].

Kemajuan di bidang *machine learning*, melahirkan metode baru yang lebih kuat, yaitu *ensemble learning*. Metode ini menggabungkan banyak model sederhana menjadi satu model yang akurat. Salah satu algoritmanya adalah *Extreme Gradient Boosting* (XGBoost). Berbeda dengan SARIMAX yang berbasis persamaan statistik parametrik, XGBoost menggunakan pendekatan *decision tree* yang dibangun secara sekuensial untuk meminimalkan *error*. Algoritma ini menawarkan kekuatan utama dalam hal efisiensi waktu pelatihan dan penanganan hubungan non-linear yang kompleks melalui pendekatan *boosting* berbasis pohon. Model ini sangat efektif dalam menangani *outlier* dan data skala besar dengan

tingkat akurasi yang baik. Namun, XGBoost dapat menghadapi risiko *overfitting* jika parameter *tuning* tidak dilakukan dengan benar.[18].

Kajian literatur terdahulu telah menunjukkan bahwa SARIMAX terbukti cukup baik dalam memprediksi fluktuasi jumlah penumpang karena mampu mengintegrasikan efek variasi kalender, seperti hari libur keagamaan dan cuti bersama serta kebijakan pembatasan Covid-19, ke dalam pemodelan deret waktu [19]. Di sisi lain, pada ranah *machine learning*, XGBoost terbukti tangguh dalam mendeteksi pola dan tren data penumpang yang kompleks dan non-linear, serta memiliki ketahanan yang tinggi terhadap risiko *overfitting*. Secara spesifik, implementasi pemodelan XGBoost pada data transportasi publik massal menunjukkan bahwa algoritma ini mampu menghasilkan nilai *error* evaluasi yang semakin rendah dan presisi apabila diintegrasikan dengan data tambahan (variabel eksogen), seperti kondisi cuaca dan status hari libur [20]. Meskipun kedua algoritma tersebut telah terbukti keandalannya secara terpisah maupun komparatif pada berbagai moda transportasi darat dan udara, aplikasinya pada karakteristik data maritim khususnya pelayaran domestik antarpulau masih sangat terbatas.

Berdasarkan tinjauan literatur, penelitian di sektor transportasi umumnya menggunakan satu metode. Belum ada penelitian yang membandingkan antara metode SARIMAX dan XGBoost untuk memprediksi jumlah penumpang domestik di Pelabuhan Sri Bintan Pura. Perbandingan ini penting untuk menemukan model dengan tingkat *error* terkecil yang valid untuk pengambilan keputusan. Berdasarkan urgensi operasional dan celah penelitian tersebut, maka penulis bermaksud melakukan penelitian untuk memodelkan dan membandingkan tingkat akurasi peramalan dari kedua algoritma tersebut. Penelitian ini dituangkan dalam bentuk skripsi dengan judul “Prediksi Jumlah Penumpang Domestik di Pelabuhan Sri Bintan Pura Menggunakan Metode SARIMAX dan XGBoost”.

B. Perumusan Masalah

Berdasarkan uraian latar belakang yang telah diuraikan di atas, permasalahan utama yang ingin dikaji dalam penelitian ini adalah bagaimana perbandingan kinerja dan tingkat akurasi prediksi dari metode SARIMAX dan

XGBoost dalam memprediksi jumlah penumpang domestik di Pelabuhan Sri Bintan Pura?.

C. Batasan Penelitian

Agar ruang lingkup penelitian ini lebih terarah, fokus, dan mendalam sesuai dengan tujuan yang ingin dicapai, maka ditetapkan batasan-batasan penelitian sebagai berikut:

1. Objek penelitian dibatasi hanya pada jumlah penumpang kapal domestik di Pelabuhan Sri Bintan Pura, Tanjungpinang.
2. Data yang digunakan merupakan data sekunder berfrekuensi bulanan (*monthly time-series*).
3. Rentang waktu data historis yang diobservasi dan dianalisis dibatasi dari Januari 2015 hingga Desember 2024, yang bersumber dari rilis resmi Badan Pusat Statistik (BPS) Provinsi Kepulauan Riau, dengan total data keseluruhan 120 data bulanan.
4. Penentuan bulan libur berdasarkan hari raya keagamaan yakni, Hari Raya Idul Fitri dan Hari Raya Natal.
5. Evaluasi tingkat akurasi serta perbandingan performa *error* antara model SARIMAX dan XGBoost menggunakan tiga metrik evaluasi, yaitu MAE, MAPE dan RMSE.

D. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, adapun tujuan dari penelitian ini adalah untuk mengetahui dan menganalisis perbandingan kinerja serta tingkat akurasi prediksi antara metode SARIMAX dan XGBoost dalam memprediksi jumlah penumpang domestik di Pelabuhan Sri Bintan Pura.

E. Manfaat Penelitian

Manfaat yang diharapkan dari pelaksanaan penelitian ini mencakup:

1. Bagi Ilmu Pengetahuan dan Teknologi (IPTEK)

- a. Memberikan kontribusi literatur dan kajian empiris di Bidang Sains Data (*Data Science*) khususnya dalam analisis pemodelan deret waktu (*time series*).

- b. Memberikan pembuktian ilmiah terkait perbandingan kinerja antara algoritma statistik klasik (SARIMAX) dan algoritma *machine learning modern* (XGBoost) ketika dihadapkan pada data transportasi laut yang memiliki fluktuasi musiman dan dipengaruhi oleh berbagai variabel.

2. Bagi Otoritas Pengelola Pelabuhan

- a. Menjadi landasan sistem pendukung keputusan yang sinergis bagi penyedia infrastruktur pelabuhan (KSOP dan PT. Pelindo).
- b. Acuan strategis untuk mengoptimalkan manajemen operasional, seperti efisiensi penjadwalan keberangkatan dan penyesuaian kapasitas armada untuk mencegah penumpukan penumpang.

3. Bagi Masyarakat

- a. Dapat memberikan informasi dan membantu masyarakat luas (calon penumpang) dalam merencanakan jadwal perjalanan yang lebih baik, dan menghindari potensi antrean panjang.
- b. Prediksi lonjakan ini dapat dimanfaatkan oleh pelaku usaha di sektor pelabuhan (seperti pedagang, penyedia jasa transportasi darat, dan sektor pariwisata) untuk melakukan perencanaan stok barang, penyediaan layanan, dan strategi usaha yang lebih optimal.