

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era digital yang terus berkembang, industri penerbangan semakin mengandalkan teknologi dan analisis data untuk meningkatkan efisiensi operasional, termasuk dalam memprediksi harga tiket pesawat yang bersifat dinamis dan dipengaruhi oleh berbagai faktor seperti waktu pemesanan, musim perjalanan, permintaan pasar, serta kebijakan maskapai. Pemanfaatan teknologi web dalam sistem pemesanan tiket pesawat menjadikan web sebagai media informasi yang dinamis, sehingga mempermudah calon penumpang dalam mengakses dan melakukan pemesanan secara lebih praktis dan efisien (Fhonna, 2021). Oleh karena itu, pengembangan model prediktif yang akurat menjadi krusial agar maskapai, agen perjalanan, dan konsumen dapat mengambil keputusan dengan lebih tepat.

Harga tiket pesawat selalu berubah-ubah dan dipengaruhi oleh berbagai faktor seperti waktu pemesanan, musim liburan, permintaan pasar, dan kebijakan maskapai. Misalnya, harga tiket untuk penerbangan di akhir pekan atau saat liburan biasanya lebih mahal dibandingkan hari biasa, dan semakin dekat dengan tanggal keberangkatan, harga tiket cenderung naik karena meningkatnya permintaan. Maskapai menggunakan sistem *dynamic pricing*, yaitu strategi penentuan harga tiket berdasarkan analisis data historis dan tingkat permintaan, yang menyebabkan harga dapat berubah dalam hitungan jam atau bahkan menit. Bagi konsumen, mengetahui kapan waktu terbaik untuk membeli tiket sangat penting agar bisa

mendapatkan harga lebih murah, sementara bagi maskapai dan agen perjalanan, sistem prediksi harga yang akurat dapat membantu dalam menetapkan strategi harga yang lebih kompetitif. Berdasarkan berbagai studi di bidang ini, penggunaan *Machine learning*, *Artificial Intelligence*, dan teknik *Deep Learning* memungkinkan estimasi harga tiket pesawat pada waktu tertentu dilakukan hanya dalam hitungan detik (Mukane dkk., 2022). Oleh karena itu, dibutuhkan model prediksi berbasis *Machine learning* yang mampu menganalisis pola perubahan harga tiket secara lebih akurat, sehingga dapat memberikan perkiraan harga di masa depan dan membantu pengguna dalam mengambil keputusan pembelian tiket dengan lebih baik. Tantangan utama dalam analitik data adalah meningkatkan performa model agar memberikan hasil optimal. Dengan metode yang tepat, penerapan teknik ini dalam analisis harga tiket dapat membantu berbagai pihak mengoptimalkan strategi penentuan tarif dan perencanaan perjalanan secara lebih efisien.

Machine learning menjadi semakin krusial dalam berbagai aspek kehidupan. Kemampuannya untuk menangani tugas-tugas kompleks yang sulit dilakukan secara manual menjadikannya solusi yang sangat dibutuhkan (Cantt, 2023). Mengingat keterbatasan manusia dalam mengakses dan mengolah data dalam jumlah besar, sistem komputer berbasis pembelajaran mesin hadir sebagai pendekatan yang efektif untuk menyederhanakan proses pengolahan data dan mendukung pengambilan keputusan secara lebih efisien.

Extreme Gradient Boosting (XGBoost) merupakan salah satu algoritma *Machine learning* berbasis pohon keputusan yang populer karena memiliki

kemampuan yang baik dalam menangani data dengan dimensi besar serta memberikan hasil prediksi yang akurat. Hasil penelitian mengenai deteksi penyakit diabetes dengan metode *Extreme Gradient Boosting (XGBoost)* menunjukkan bahwa pendekatan ini memiliki kinerja yang baik dan akurasi tinggi dalam mengklasifikasikan dataset penyakit diabetes (Dinanthi dkk., 2024). Namun, performa model *XGBoost* sangat bergantung pada pemilihan *Hyperparameter* yang optimal. *Hybrid XGBoost* model dengan *Hyperparameter Tuning* digunakan untuk meningkatkan akurasi dalam prediksi penyakit liver, sehingga dapat membantu dalam proses diagnosis dan pengobatannya (Dalal dkk., 2022). Pendekatan ini menggabungkan kekuatan model *XGBoost* dengan teknik penyetelan *hyperparameter* untuk meningkatkan kinerja prediksi. Dengan optimasi yang tepat, model ini mampu mengidentifikasi pola dalam data secara lebih akurat, sehingga dapat mendukung pengambilan keputusan medis yang lebih efektif dalam mendeteksi dan menangani penyakit liver. Pendekatan ini menggabungkan kekuatan model *XGBoost* dengan teknik penyetelan *Hyperparameter* untuk meningkatkan kinerja prediksi. Dengan optimasi yang tepat, model ini mampu mengidentifikasi pola dalam data secara lebih akurat.

Hyperparameter Tuning adalah proses yang dilakukan untuk mencari kombinasi parameter terbaik guna meningkatkan performa model dan menghindari *overfitting* atau *underfitting*. *Hyperparameter Tuning* sangat penting untuk mengoptimalkan kinerja dan generalisasi model pembelajaran mesin, dengan berbagai teknik pengoptimalan yang dieksplorasi (Ilemobayo dkk., 2024). Hasil penelitian (H. Darmawan dkk., 2023) menunjukkan bahwa model regresi *XGBoost*

memiliki performa lebih baik pada tahap pertama dengan rata-rata RMSE 1,2728 dibandingkan regresi GRU yang memiliki RMSE 1,551. Pada tahap kedua, regresi GRU melampaui *XGBoost* dengan RMSE 2,23 berbanding 2,28, serta dalam pengujian klasifikasi, model GRU menunjukkan keunggulan dengan nilai *F1-score* 0,88 dibandingkan *XGBoost* yang memiliki 0,86, sehingga temuan ini memperkuat pemilihan metode yang paling sesuai untuk optimasi model (H. Darmawan dkk., 2023). Temuan ini menunjukkan bahwa pemilihan metode yang tepat bergantung pada tahap dan jenis tugas yang dilakukan, sehingga kombinasi atau optimasi lebih lanjut dari kedua model dapat menjadi strategi yang efektif untuk meningkatkan akurasi dan kinerja prediksi secara keseluruhan.

Salah satu metode yang dapat digunakan untuk optimasi *Hyperparameter* adalah *Bayesian Optimization*. Optimasi *Bayesian* merupakan metode yang efektif dalam mengoptimalkan fungsi objektif yang memerlukan perhitungan mahal, dengan keberhasilannya yang telah terbukti di berbagai disiplin ilmu (*Optimization*, n.d.). Metode ini menawarkan pendekatan berbasis probabilitas dalam pencarian kombinasi *Hyperparameter* terbaik dengan jumlah iterasi yang lebih sedikit dibandingkan pencarian berbasis *Grid Search* atau *Random Search*. Dengan menganalisis tantangan keamanan dalam algoritma *Bayesian Optimization* serta menyediakan solusi algoritma yang dapat langsung diterapkan, penelitian ini berkontribusi dalam mewujudkan pembelajaran mesin yang lebih aman dan andal untuk diterapkan dalam berbagai aplikasi dunia nyata (Fiedler & Kreisköther, n.d.). Dengan *Bayesian Optimization* yang diterapkan melalui *BayesSearchCV*, pencarian

Hyperparameter dilakukan dengan lebih efisien, sehingga dapat menghemat waktu komputasi sambil tetap menghasilkan model dengan akurasi yang lebih tinggi.

Berdasarkan seluruh uraian persoalan yang telah dibahas, terdapat kesenjangan penelitian yang masih tersisa, yaitu kurangnya optimalisasi dalam pemilihan *Hyperparameter* pada model *XGBoost* untuk prediksi harga tiket pesawat. Dalam berbagai penelitian sebelumnya, metode *Machine learning* telah banyak digunakan untuk menganalisis dan memprediksi harga tiket. Namun, tantangan utama yang masih dihadapi adalah menentukan kombinasi *Hyperparameter* yang optimal. Pemilihan *Hyperparameter* yang tidak tepat dapat menyebabkan model memiliki performa yang kurang maksimal, baik dari segi akurasi prediksi maupun efisiensi komputasi. Saat ini, metode yang umum digunakan dalam optimasi *Hyperparameter* adalah *Grid Search* dan *Random Search*. Meskipun metode ini cukup populer, keduanya memiliki keterbatasan, terutama dalam hal efisiensi waktu dan sumber daya komputasi yang dibutuhkan. *Grid Search*, misalnya, melakukan eksplorasi menyeluruh terhadap seluruh kombinasi *Hyperparameter* yang telah ditentukan sebelumnya, sehingga dapat sangat memakan waktu. Sementara itu, *Random Search* mencoba kombinasi *Hyperparameter* secara acak, yang meskipun lebih cepat dibandingkan *Grid Search*, tetapi memiliki kemungkinan besar untuk melewatkan kombinasi terbaik.

Penelitian ini menggunakan *Bayesian Optimization* sebagai pendekatan yang lebih efisien dalam pencarian *Hyperparameter* optimal pada model *XGBoost*. *Bayesian Optimization* bekerja dengan memanfaatkan pendekatan berbasis probabilitas untuk mengarahkan pencarian ke area yang lebih menjanjikan dalam

ruang *Hyperparameter*, sehingga dapat mengurangi jumlah iterasi yang dibutuhkan untuk menemukan kombinasi terbaik. Dengan cara ini, diharapkan model prediksi harga tiket pesawat dapat bekerja lebih akurat dan lebih cepat dibandingkan dengan metode *Tuning* konvensional. Selain itu, penelitian ini juga akan melakukan perbandingan kinerja antara hasil *Bayesian Optimization* dan metode *Tuning* lainnya, seperti *Grid Search* dan *Random Search*, untuk mengevaluasi efektivitas dan efisiensi masing-masing pendekatan.

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pengembangan model prediksi harga tiket pesawat berbasis *Machine learning* yang lebih akurat dan efisien. Dari sisi akademik, penelitian ini akan menambah wawasan mengenai implementasi *Bayesian Optimization* dalam optimasi *Hyperparameter*, serta memberikan referensi bagi penelitian-penelitian selanjutnya yang ingin mengembangkan model prediksi serupa. Sementara itu, dari sisi industri penerbangan, hasil penelitian ini dapat membantu maskapai dan agen perjalanan dalam menyusun strategi harga yang lebih adaptif, berdasarkan prediksi harga tiket yang lebih akurat. Dengan demikian, baik akademisi maupun praktisi di bidang ini dapat memperoleh manfaat dari penelitian yang dilakukan, baik dalam aspek pengembangan teknologi maupun penerapannya dalam industri.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, terdapat beberapa permasalahan yang akan dikaji dalam penelitian ini:

1. Bagaimana performa *Bayesian Optimization* dalam meningkatkan akurasi model *XGBoost* untuk prediksi harga tiket pesawat?

2. Sejauh mana optimasi *Hyperparameter* menggunakan *Bayesian Optimization* dapat meningkatkan performa model *XGBoost* berdasarkan metrik evaluasi seperti *Mean Squared Error* (MSE), *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *R-squared* (R^2)?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menganalisis efektivitas *Bayesian Optimization* dalam meningkatkan akurasi prediksi harga tiket pesawat dengan model *XGBoost*.
2. Mengevaluasi dampak optimasi *Hyperparameter* terhadap performa model berdasarkan metrik MSE, RMSE, MAE, dan R^2 .

1.4 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat sebagai berikut:

a. Secara Teoritis:

1. Menambah wawasan dalam bidang *Machine learning* terkait dengan optimasi *Hyperparameter* menggunakan *Bayesian Optimization*.
2. Memberikan kontribusi terhadap studi-studi yang membahas efektivitas *BayesSearchCV* dalam meningkatkan performa model prediksi.

b. Secara Praktis:

1. Memberikan panduan bagi praktisi dan peneliti dalam memilih metode *Hyperparameter Tuning* yang lebih optimal untuk model *XGBoost*.
2. Membantu dalam pengambilan keputusan berbasis data dengan model prediksi yang lebih akurat dan efisien.

1.5 Batasan Masalah

Penelitian ini menetapkan batasan-batasan masalah supaya penelitian ini lebih terarah dan fokus, terdapat beberapa batasan yang ditetapkan:

1. Penelitian hanya menggunakan model *XGBoost* sebagai algoritma prediksi dan model prediktif.
2. Optimasi *Hyperparameter* hanya menggunakan pendekatan *Bayesian Optimization* dengan implementasi *BayesSearchCV*.
3. Evaluasi performa model dilakukan menggunakan metrik regresi yakni MSE, RMSE, MAE, dan R^2 , tanpa membahas metrik klasifikasi atau metrik lainnya.
4. Eksperimen dilakukan menggunakan skema validasi silang atau *cross-validation* dengan *10-fold* untuk memastikan stabilitas hasil yang diperoleh.
5. Dataset yang digunakan dalam penelitian ini merupakan data publik yang tersedia di platform *Kaggle*, dengan judul "*Flight Price Prediction*" dan dapat diakses melalui tautan berikut:
https://www.Kaggle.com/datasets/shubhambathwal/flight-price-prediction/data?select=Clean_Dataset.csv.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini disusun untuk memberikan gambaran yang jelas mengenai tahapan pelaksanaan penelitian dari awal hingga akhir. Adapun sistematika penulisan laporan tugas akhir ini adalah sebagai berikut:

Bab I Pendahuluan : Bab ini berisi latar belakang penelitian yang menjelaskan alasan dilakukannya penelitian, rumusan masalah yang ingin dijawab,

tujuan penelitian yang hendak dicapai, manfaat penelitian baik secara teoritis maupun praktis, batasan masalah yang digunakan untuk memperjelas ruang lingkup penelitian, serta sistematika penulisan laporan tugas akhir ini.

Bab II Tinjauan Pustaka : Bab ini berisi landasan teori yang berkaitan dengan topik penelitian, seperti teori mengenai *Machine learning*, *Extreme Gradient Boosting (XGBoost)*, *Hyperparameter Tuning*, serta metode *Bayesian Optimization*. Selain itu, bab ini juga mencakup penelitian-penelitian terdahulu yang relevan dan menjadi dasar dalam pengembangan penelitian ini.

Bab III Metodologi Penelitian : Bab ini menjelaskan metode dan tahapan penelitian yang digunakan, meliputi rancangan penelitian, diagram alur penelitian, tahapan pemrosesan data (*preprocessing*), pemodelan menggunakan algoritma *XGBoost*, serta proses optimasi *Hyperparameter* menggunakan *Bayesian Optimization*. Bab ini juga menjelaskan prosedur evaluasi model dengan metrik performa seperti MSE, RMSE, MAE, dan R^2 .

Bab IV Hasil dan Pembahasan : Bab ini menyajikan hasil implementasi dari penelitian yang telah dilakukan, mulai dari proses pengumpulan dan pengolahan data, hasil optimasi *Hyperparameter*, performa model sebelum dan sesudah dilakukan optimasi, serta analisis dan interpretasi terhadap hasil yang diperoleh. Pembahasan dilakukan dengan mengaitkan hasil eksperimen dengan teori dan penelitian sebelumnya.

Bab V Penutup : Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya agar dapat menjadi bahan

pengembangan lebih lanjut di bidang optimasi model *Machine learning* menggunakan *Bayesian Optimization*.