

# **BAB I**

## **PENDAHULUAN**

Bab pendahuluan ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan yang akan digunakan dalam dokumen skripsi ini yang berjudul Klasifikasi Keterlambatan Kedatangan Pesawat Terbang dengan *Stacking Ensemble Learning*.

### **1.1. Latar Belakang**

Keterlambatan kedatangan pesawat adalah masalah signifikan dalam industri penerbangan yang berdampak luas terhadap berbagai aspek, termasuk kepuasan penumpang, efisiensi operasional maskapai, dan kinerja bandara. Data dari Bureau of Transportation Statistics menunjukkan bahwa sekitar 20% dari seluruh penerbangan mengalami keterlambatan lebih dari 15 menit setiap tahunnya (Dhone dkk., 2021).

Amerika Serikat mencatat kerugian ekonomi lebih dari 33 miliar USD per tahun akibat keterlambatan penerbangan, yang disebabkan oleh biaya tambahan operasional dan kompensasi bagi penumpang (Mtimkulu & Maphosa, 2023). Sedangkan di Eropa, rata-rata keterlambatan penerbangan mencapai 17,5 menit per penerbangan pada tahun 2023, dengan hanya 70,6% penerbangan yang tiba tepat waktu atau lebih awal dari jadwal yang ditentukan (Eurocontrol, 2024).

Kawasan Asia, khususnya China dan India, menghadapi lonjakan jumlah penumpang yang signifikan sejak tahun 2016 hingga 2019, sebelum terjadinya pandemi COVID-19. Selama periode tersebut, pertumbuhan tahunan penumpang udara mencapai dua digit, mendorong peningkatan tekanan terhadap infrastruktur dan sistem pengelolaan lalu lintas udara di kedua negara (Gui dkk., 2020). Beberapa bandara utama yang terdampak di China antara lain Beijing Capital International Airport, Shanghai Pudong International Airport, dan Guangzhou Baiyun International Airport, sementara di India meliputi Indira Gandhi International Airport (Delhi), Chhatrapati Shivaji Maharaj International Airport (Mumbai), dan Kempegowda International Airport (Bangalore). Bandara-bandara tersebut mengalami rata-rata keterlambatan penerbangan melebihi 30 menit, yang mencerminkan beban operasional dan kompleksitas lalu lintas udara yang terus meningkat.

Dampak keterlambatan penerbangan dirasakan oleh berbagai pihak. Penumpang mengalami ketidaknyamanan akibat perubahan jadwal, pembatalan penerbangan, serta biaya tambahan yang mungkin timbul, seperti pengeluaran akomodasi dan konsumsi yang tidak terduga. Maskapai penerbangan menghadapi peningkatan biaya operasional, risiko kehilangan loyalitas pelanggan, serta kewajiban kompensasi yang besar, sebagaimana diatur dalam regulasi penerbangan di berbagai negara. Bandara dan pengelola lalu lintas udara juga terdampak, karena harus menangani antrean panjang, penjadwalan ulang slot penerbangan, serta peningkatan kepadatan lalu lintas udara yang dapat memicu gangguan operasional lebih lanjut (Reddy dkk., 2023).

Berbagai strategi manajerial dan kebijakan operasional telah diterapkan untuk mengurangi dan memprediksi keterlambatan penerbangan. Studi yang dilakukan oleh Eurocontrol mengusulkan sistem manajemen slot yang lebih ketat guna mengoptimalkan kapasitas bandara dan mengurangi kemacetan lalu lintas udara. Hasilnya menunjukkan penurunan keterlambatan rata-rata sebesar 8,5% di bandara dengan volume lalu lintas tinggi (Eurocontrol, 2024). Strategi lain yang dikembangkan oleh Federal Aviation Administration (FAA) memanfaatkan analisis pola penerbangan untuk memberikan rekomendasi penjadwalan yang lebih efisien, yang mampu mengurangi keterlambatan rata-rata hingga 10 menit per penerbangan (FAA, 2024).

Meskipun pendekatan kebijakan dan operasional tersebut memberikan dampak positif, kompleksitas sistem transportasi udara yang semakin tinggi membutuhkan solusi yang lebih adaptif dan presisi. Dalam hal ini, penerapan *machine learning* menjadi salah satu pendekatan utama yang semakin diandalkan untuk memprediksi keterlambatan penerbangan. Kemampuan *machine learning* dalam mengolah data dalam jumlah besar dan mengidentifikasi pola-pola kompleks yang sulit dideteksi secara manual, mampu memberikan prediksi yang lebih akurat dan dapat disesuaikan dengan kondisi operasional yang dinamis. Oleh karena itu, pengembangan model prediktif berbasis *machine learning* menjadi langkah penting dalam mendukung pengambilan keputusan yang lebih cepat, tepat, dan berbasis data dalam industri penerbangan.

Beberapa penelitian terdahulu telah membuktikan efektivitas algoritma *machine learning* dalam memprediksi keterlambatan penerbangan. Hatipoğlu dan Tosun (2024) menunjukkan bahwa Extreme Gradient Boosting (XGBoost) memberikan akurasi tertinggi

sebesar 80% pada data operasional penerbangan di Turki. Gilbert dan How (2024) menemukan bahwa Decision Tree menghasilkan nilai *precision-recall* sebesar 0,975, mengungguli Random Forest dan Deep Neural Network dalam memprediksi keterlambatan di Bandara JFK. Borse dkk. (2020) menunjukkan bahwa Naive Bayes mampu menghasilkan *precision* hingga 75% dan tingkat kesalahan sekitar 15–18% saat menangani data cuaca. Sementara itu, Alsulamy (2024) menemukan bahwa Light Gradient-Boosting Machine (LightGBM) unggul dalam memprediksi keterlambatan proyek konstruksi, dengan akurasi dan sensitivitas tertinggi pada semua kategori.

Model *ensemble*, khususnya *stacking*, telah terbukti efektif dalam meningkatkan akurasi dan stabilitas prediksi. Yi dkk. (2021) mencatat bahwa model *stacking* yang dikombinasikan dengan SMOTE dan seleksi fitur Boruta mampu menghasilkan performa tinggi dengan evaluasi metrik *accuracy*, *precision*, *recall*, dan *f1-score* yang lebih baik dibandingkan model tunggal. Selain itu, Mao dkk. (2025) menunjukkan bahwa optimasi *hyperparameter* menggunakan Optuna meningkatkan akurasi model hingga 8,45% pada beberapa *dataset* publik, sekaligus meningkatkan interpretabilitas model melalui analisis SHAP.

Berdasarkan hasil-hasil tersebut, penelitian ini menerapkan *stacking ensemble learning* dan optimasi *hyperparameter* menggunakan Optuna pada empat algoritma yang telah terbukti kuat dalam berbagai studi, yaitu Decision Tree, Naive Bayes, LightGBM, dan XGBoost. Penelitian ini bertujuan untuk mengevaluasi pengaruh *stacking* terhadap peningkatan performa model tunggal dalam klasifikasi keterlambatan penerbangan serta membandingkan performa keempat model tersebut setelah optimasi.

## 1.2. Rumusan Masalah

Berdasarkan permasalahan yang telah dibahas, penelitian ini berfokus pada:

1. Pengaruh *stacking ensemble learning* terhadap peningkatan performa model *machine learning* tunggal dalam klasifikasi keterlambatan penerbangan.
2. Perbandingan performa model Decision Tree, Naïve Bayes, LightGBM, dan XGBoost setelah dilakukan optimasi *hyperparameter* menggunakan Optuna.

### 1.3. Tujuan dan Manfaat

Penelitian ini bertujuan untuk mengevaluasi pengaruh *stacking ensemble learning* terhadap peningkatan performa model *machine learning* tunggal dalam klasifikasi keterlambatan penerbangan. Selain itu, penelitian ini juga bertujuan untuk membandingkan performa model Decision Tree, Naïve Bayes, LightGBM, dan XGBoost dengan optimasi *hyperparameter* menggunakan Optuna. Manfaat dari penelitian ini adalah memberikan kontribusi dalam pengembangan model prediktif keterlambatan penerbangan dengan pendekatan *ensemble learning* dan optimasi *hyperparameter* menggunakan Optuna. Selain itu, penelitian ini juga diharapkan dapat menjadi referensi ilmiah dalam pemilihan algoritma klasifikasi dan teknik *ensemble* yang efektif untuk permasalahan prediksi di bidang transportasi udara.

### 1.4. Ruang Lingkup

Penelitian ini diperlukan adanya ruang lingkup agar lebih terarah dan mencapai tujuan yang diharapkan. Ruang lingkup dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini dilakukan menggunakan bahasa pemrograman Python dan perkakas Google Colaboratory.
2. Data yang digunakan dalam penelitian ini bersumber dari Bureau of Transportation Statistics (BTS), yang menyediakan informasi rinci mengenai kinerja penerbangan domestik di Amerika Serikat.
3. Data penerbangan diambil dari Januari hingga Desember 2024, yang mencakup 150.000 baris data secara acak dari setiap bulannya. Maka, total data yang dianalisis berjumlah 1.800.000 baris dan 120 fitur.

### 1.5. Sistematika Penulisan

Struktur penulisan dalam laporan ini dibagi menjadi lima bab utama yang membahas penelitian Klasifikasi Keterlambatan Kedatangan Pesawat Terbang dengan *Stacking Ensemble Learning*. Pembagian ini disusun dengan tujuan agar penulisan menjadi lebih sistematis dan mudah dipahami. Berikut ini adalah ringkasan singkat dari setiap bab yang dibahas:

|         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|---------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| BAB I   | PENDAHULUAN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|         | Bab ini menyajikan latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, serta sistematika penulisan skripsi Klasifikasi Keterlambatan Kedatangan Pesawat Terbang dengan <i>Stacking Ensemble Learning</i> .                                                                                                                                                                                                                                                                          |
| BAB II  | TINJAUAN PUSTAKA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|         | Bab ini menyajikan hasil studi pustaka yang menjadi dasar teori dalam pelaksanaan dan penyusunan skripsi. Studi pustaka yang dibahas mencakup konsep dan teknik yang berkaitan langsung dengan klasifikasi keterlambatan penerbangan, yaitu <i>state-of-the-art</i> , penerbangan komersial, praproses data, klasifikasi dan <i>machine learning</i> , Decision Tree, Naïve Bayes, LightGBM, XGBoost, Optuna, <i>stacking ensemble learning</i> , <i>cross validation</i> , dan evaluasi model klasifikasi. |
| BAB III | METODE PENELITIAN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|         | Bab ini menjelaskan langkah-langkah yang dilakukan dalam penelitian Klasifikasi Keterlambatan Kedatangan Pesawat Terbang dengan <i>Stacking Ensemble Learning</i> yang meliputi pengumpulan dan eksplorasi data, praproses data, pelabelan data, pembagian data latih dan data uji, optimasi <i>hyperparameter</i> dengan Optuna pada algoritma Decision Tree, Naïve Bayes, LightGBM, dan XGBoost, metode <i>stacking</i> , dan evaluasi model.                                                             |
| BAB IV  | HASIL DAN PEMBAHASAN                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|         | Bab ini menjelaskan implementasi dan analisis hasil dari penelitian berjudul Klasifikasi Keterlambatan Kedatangan Pesawat Terbang dengan <i>Stacking Ensemble Learning</i> . Pembahasan mencakup lingkungan dan perangkat yang digunakan dalam eksperimen, skenario klasifikasi yang diterapkan, dan evaluasi performa model pada setiap skenario.                                                                                                                                                          |
| BAB V   | PENUTUP                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|         | Bab ini berisi kesimpulan dari bab-bab yang telah dijelaskan sebelumnya dan saran untuk pengembangan penelitian klasifikasi keterlambatan kedatangan pesawat terbang.                                                                                                                                                                                                                                                                                                                                       |