

# BAB 1

## PENDAHULUAN

---

### 1.1 Latar Belakang

Industri pariwisata global telah menunjukkan pertumbuhan yang luar biasa selama beberapa dekade terakhir, dengan peningkatan jumlah wisatawan internasional dari hanya 25 juta pada tahun 1950 menjadi 1,4 miliar pada tahun 2019 . Pertumbuhan ini didorong oleh berbagai faktor seperti peningkatan pendapatan global, kemajuan teknologi transportasi, serta kemudahan akses informasi melalui internet. Industri pariwisata yang berkembang pesat ini telah mengubah dinamika sektor perhotelan, menuntut penyesuaian berkelanjutan pada harga, ketersediaan kamar, dan kualitas layanan untuk memenuhi ekspektasi pelanggan yang semakin tinggi. Misalnya, penelitian menunjukkan bahwa setiap peningkatan 10% dalam jumlah wisatawan dapat meningkatkan pendapatan hotel hingga 5% . Namun, pertumbuhan ini juga menghadirkan tantangan baru bagi manajer hotel, terutama dalam hal mengelola pembatalan reservasi yang dapat berdampak signifikan pada pendapatan dan efisiensi operasional [1].

Dalam konteks perkembangan pariwisata yang pesat, hotel-hotel di seluruh dunia menghadapi tuntutan untuk terus berinovasi dan menyesuaikan fasilitas serta layanan mereka. Antisipasi dan preferensi pelancong yang terus berkembang menuntut adanya perubahan dalam manajemen hotel. Faktor-faktor utama seperti harga kamar, ketersediaan, dan kualitas layanan kamar menjadi semakin dinamis. Menurut data dari World Tourism Organization, peningkatan volume wisatawan internasional dari 25 juta pada tahun 1950 menjadi 1,4 miliar pada tahun 2019 tidak hanya mendorong pertumbuhan industri perhotelan tetapi juga menuntut penyesuaian berkelanjutan dalam strategi manajemen untuk mempertahankan profitabilitas dan kepuasan pelanggan [2].

Proses manual dalam manajemen hotel, terutama dalam hal pengelolaan reservasi dan penentuan harga, sering kali menjadi hambatan signifikan di tengah dinamika

industri yang cepat berubah. Metode tradisional ini cenderung tidak responsif terhadap fluktuasi permintaan pasar, menyebabkan ketidakefektifan dalam pemanfaatan kapasitas dan distribusi sumber daya. Sebagai contoh, studi telah mengungkapkan bahwa hotel yang masih menggunakan metode manual untuk manajemen reservasi mengalami rata-rata tingkat pembatalan sebesar 8% dan pada pemesanan daring sebesar 17%, sehingga berdampak negatif pada pendapatan dan efisiensi operasional [3]. Oleh karena itu, diperlukan pendekatan yang lebih canggih dan adaptif untuk meningkatkan akurasi prediksi dan efisiensi manajemen hotel, guna menghadapi tantangan ini.

Dalam kerangka ini, pengembangan aplikasi model prediksi pembatalan hotel dengan menggunakan Machine Learning (ML) menjadi solusi yang menjanjikan. Aplikasi ini akan membantu pengelola hotel untuk mengantisipasi pembatalan dengan lebih baik, mengoptimalkan ketersediaan kamar, dan meningkatkan responsibilitas terhadap perubahan dinamika pasar. Model ML seperti Boosting models, yang mencakup algoritma seperti Gradient Boosting, AdaBoost, dan XGBoost, dapat digunakan untuk menganalisis data historis dan memprediksi kemungkinan pembatalan reservasi. Dengan implementasi model ini, hotel dapat mengambil tindakan proaktif untuk mengurangi dampak pembatalan, seperti penawaran khusus untuk mengisi kembali kamar yang dibatalkan atau penyesuaian strategi harga untuk mengoptimalkan pendapatan. Inovasi ini diharapkan dapat memberikan kontribusi positif dalam mengatasi tantangan terkait manajemen pembatalan di industri perhotelan, sehingga meningkatkan efisiensi operasional dan kepuasan pelanggan.

Perkembangan pesat industri pariwisata telah memberikan dampak besar pada sektor perhotelan, menuntut adanya penyesuaian berkelanjutan dalam manajemen harga, ketersediaan kamar, dan kualitas layanan. Tantangan dalam pengelolaan pembatalan hotel memerlukan pendekatan yang lebih canggih dan adaptif. Pengembangan aplikasi model prediksi pembatalan hotel dengan menggunakan Machine Learning menawarkan solusi inovatif untuk mengantisipasi dan mengelola pembatalan dengan lebih baik, mengoptimalkan ketersediaan kamar, dan meningkatkan responsibilitas terhadap perubahan dinamika pasar. Dengan

demikian, diharapkan dapat meningkatkan profitabilitas dan kepuasan pelanggan di industri perhotelan yang semakin kompetitif.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, fokus permasalahan yang muncul dalam industri perhotelan terkait dengan prediksi pembatalan hotel dapat dirumuskan sebagai berikut:

### **Rumusan Masalah:**

1. Bagaimana cara meningkatkan akurasi prediksi pembatalan reservasi hotel?
2. Apa saja faktor-faktor utama yang mempengaruhi pembatalan reservasi hotel?

## **1.3 Tujuan**

Penelitian ini dilakukan dengan tujuan untuk mengatasi tantangan yang dihadapi oleh manajer hotel dalam mengelola pembatalan reservasi dengan memanfaatkan teknologi Machine Learning. Tujuan yang ingin dicapai meliputi:

1. Meningkatkan akurasi prediksi pembatalan reservasi hotel dengan menggunakan metode Machine Learning yang efektif dan validasi yang tepat. Mengembangkan model prediksi pembatalan yang mampu mengidentifikasi kemungkinan pembatalan dengan tingkat akurasi yang tinggi.
2. Mengidentifikasi faktor-faktor utama yang mempengaruhi pembatalan reservasi melalui analisis data yang mendalam dan pengembangan aplikasi Machine Learning yang user-friendly, yang mencakup modul untuk EDA, pelatihan dan validasi model, serta prediksi real-time.

## **1.4 Batasan Masalah**

Penelitian ini memiliki batasan tertentu untuk memastikan fokus dan kelayakan pelaksanaan proyek. Berikut adalah batasan-batasan dalam ruang lingkup kerja/magang dan pelaksanaan solusi yang diidentifikasi:

1. Fokus pada eksplorasi data historis reservasi hotel melalui Exploratory Data Analysis (EDA) untuk memahami distribusi data, korelasi antar fitur, dan identifikasi pola serta anomali dalam data.
2. Melakukan analisis statistik dasar untuk menggali insight yang relevan terkait dengan pembatalan reservasi.
3. Pemilihan dan penerapan algoritma Machine Learning berbasis Boosting Model untuk prediksi pembatalan reservasi.w
4. Melakukan pelatihan dan validasi model untuk memastikan tingkat akurasi yang optimal dalam prediksi pembatalan.
5. Mengembangkan aplikasi berbasis web atau desktop yang user-friendly untuk digunakan oleh manajer hotel, mencakup fitur visualisasi hasil EDA, pelatihan dan validasi model, serta prediksi pembatalan secara real-time.
6. Aplikasi akan fokus pada prediksi pembatalan reservasi hotel berdasarkan data historis yang tersedia, tanpa mencakup aspek manajemen operasional hotel secara langsung seperti pengelolaan staf, inventaris fisik, atau pelayanan pelanggan.
7. Dataset yang digunakan terbatas pada data historis yang telah diberikan oleh pihak hotel, tanpa pengumpulan data tambahan dari sumber eksternal.
8. Model prediksi hanya akan menggunakan fitur-fitur yang tersedia dalam dataset yang ada, tanpa memperhitungkan faktor eksternal seperti tren pasar global atau kebijakan pemerintah terkait pariwisata.
9. Penelitian ini tidak akan mencakup pengembangan atau pembaruan infrastruktur teknologi yang digunakan oleh hotel, seperti sistem manajemen hotel (PMS) atau perangkat keras, dan tidak akan melibatkan perubahan langsung dalam prosedur operasional hotel.

### **1.5 Metode Pengerjaan**

Dalam penelitian ini, metode pengerjaan yang digunakan adalah Cross-Industry Standard Process for Data Mining (CRISP-DM). Metode ini dipilih karena fleksibilitas dan kerangka kerja yang terstruktur untuk proyek-proyek data mining dan machine learning. Model CRISP-DM terdiri dari enam tahapan utama yang berupa:

### 1. Business Understanding

Pada tahap ini, tujuan bisnis didefinisikan dengan jelas. Penelitian ini bertujuan untuk mengembangkan model prediksi pembatalan reservasi hotel menggunakan algoritma machine learning berbasis boosting. Selain itu, dilakukan identifikasi dan pemahaman terhadap masalah bisnis yang dihadapi oleh manajer hotel, seperti tingginya tingkat pembatalan reservasi yang dapat mempengaruhi pendapatan dan operasional hotel.

### 2. Data Understanding

Tahap ini melibatkan pengumpulan dan eksplorasi data. Data historis reservasi hotel dikumpulkan dari berbagai sumber seperti sistem manajemen hotel (PMS), database internal, dan file CSV atau Excel. Exploratory Data Analysis (EDA) dilakukan untuk memahami distribusi data, korelasi antar fitur, serta mengidentifikasi pola dan anomali dalam data. Tools yang digunakan dalam tahap ini termasuk Python dengan library seperti pandas, numpy, dan matplotlib.

### 3. Data Preparation

Pada tahap ini, data yang telah dikumpulkan dibersihkan dan diproses lebih lanjut. Proses ini mencakup:

- Data Ingestion

Mengumpulkan data historis reservasi hotel dari berbagai sumber.

- Data Cleaning

Membersihkan data dari kesalahan, duplikasi, dan nilai yang hilang.

- Feature Engineering

Membuat fitur-fitur baru yang dapat memberikan informasi tambahan kepada model.

- Data Storage

Menyimpan data yang telah dibersihkan dan diproses ke dalam format yang sesuai, seperti database SQL atau file HDF5.

#### 4. Modeling

Pada tahap ini, model machine learning dikembangkan. Algoritma berbasis boosting seperti LightGBM, Histogram-based Gradient Boosting, AdaBoost Classifier, dan XGBoost Classifier dipilih untuk pelatihan model. Tahap ini melibatkan:

- Melatih model menggunakan data yang telah dipersiapkan.
- Menguji model untuk memastikan performa dan akurasi yang optimal menggunakan teknik cross-validation dan metrik seperti akurasi, precision, recall, dan F1-score.
- Tools yang digunakan termasuk Python dengan library scikit-learn, LightGBM, dan XGBoost.

#### 5. Evaluation

Tahap evaluasi bertujuan untuk menilai efektivitas model yang telah dikembangkan. Model dievaluasi berdasarkan kriteria yang telah ditentukan, seperti akurasi dan recall. Selain itu, dilakukan analisis untuk memastikan bahwa model dapat memberikan prediksi yang andal dan dapat digunakan untuk pengambilan keputusan oleh manajer hotel.

#### 6. Deployment

Tahap akhir adalah implementasi model ke dalam lingkungan produksi. Aplikasi berbasis web menggunakan Streamlit dikembangkan untuk menyediakan antarmuka pengguna yang user-friendly. Aplikasi ini memungkinkan manajer hotel untuk mengakses dan menggunakan hasil prediksi secara real-time, sehingga dapat mengantisipasi pembatalan dengan lebih baik dan mengoptimalkan ketersediaan kamar.

## 1.6 Jadwal Pengerjaan

Tabel 1. 1 Tabel Pelaksanaan Kerja

Bulan	Penyiapan Data	EDA	Pengembangan Model	Pelatihan dan Pengujian Model	Deployment	Pemeliharaan dan Perbaikan
Jun	Data Ingestion	Initial Analysis				
Jul	Data Cleaning					
Agu	Data Storage					
Sep		Eksplorasi Fitur	Feature Engineering			
Okt			Seleksi Model	Initial Training		
Nov				Testing Model		
Des				Tuning HPO		
Jan				Validasi Model		
Feb					Model Deployment	
Ma					User Training	
Apr						Monitoring Performa
Mei						Perbaikan Model

Tabel 1.5 merupakan referensi panduan penjadwalan dalam proyek pengembangan model prediksi pembatalan reservasi hotel selama satu tahun. Tabel ini memberikan gambaran umum tentang langkah-langkah utama yang terlibat dalam proyek, mulai dari persiapan data, eksplorasi data, pengembangan model, pelatihan dan pengujian model, hingga penerapan dan pemeliharaan model. Harap dicatat bahwa tabel ini tidak merefleksikan proses yang terjadi secara real dan hanya berfungsi sebagai panduan. Untuk histori pengerjaan proyek yang sebenarnya, dapat merujuk pada lampiran yang telah disediakan.