

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Menurut laporan Rumah123 edisi April 2024 yang dikutip oleh Kompas.com, harga rumah di Indonesia mengalami kenaikan tahunan sebesar 0,7% pada Maret 2024 dibandingkan Maret 2023. Di kawasan Jabodetabek, Bogor mencatat kenaikan harga hunian tertinggi, yaitu 5,6%, dengan pertumbuhan yang stabil. Kenaikan ini dipengaruhi oleh perkembangan infrastruktur, tingginya minat masyarakat terhadap proyek perumahan, serta harga yang masih lebih terjangkau dibandingkan daerah lain di Jabodetabek. Selain itu, lingkungan yang asri dan akses mudah ke pusat kota turut meningkatkan permintaan hunian di Bogor, menjadikannya pilihan menarik untuk investasi properti, terutama bagi masyarakat urban yang mencari hunian nyaman di luar Jakarta.

Rumah sebagai salah satu aset properti yang signifikan, sering dipengaruhi oleh kemajuan infrastruktur dan tingginya permintaan hunian di lokasi strategis (Bastian et al., 2023). Kawasan Bogor menjadi salah satu pilihan utama bagi pembeli dan investor yang mencari peluang properti yang menguntungkan. Namun, fluktuasi harga rumah tetap menjadi tantangan utama, dengan variasi harga yang dipengaruhi oleh faktor seperti luas lahan, dimensi bangunan, dan total kamar. Oleh karena itu, prediksi harga rumah di kawasan ini menjadi aspek penting dalam pengambilan keputusan bagi pembeli, penjual, dan pengembang properti. Bagi pembeli, prediksi yang akurat membantu merencanakan keuangan serta menentukan harga yang sesuai, sementara bagi penjual dan pengembang, prediksi harga berperan dalam merancang strategi pemasaran yang kompetitif (Nuris, 2024).

Metode prediksi harga rumah seringkali terbatas dalam menghadapi kompleksitas data dan dinamika pasar yang terus berubah. Penggunaan algoritma yang lebih mutakhir, seperti *Random Forest* dan *Gradient Boosting*, memungkinkan hasil prediksi harga rumah dengan akurasi yang lebih tinggi. *Random Forest* dikenal efektif dalam mengidentifikasi hubungan *non-linear* antar variabel, seperti luas tanah, luas bangunan, dan jumlah kamar tidur. Selain itu, algoritma ini juga lebih tahan terhadap *overfitting*. Sementara itu, *Gradient Boosting* merupakan metode pembelajaran mesin yang membentuk model dengan

cara bertahap, yang dimana setiap model yang dibuat memiliki tujuan memperbaiki kekurangan atau kesalahan dari model yang dibuat sebelumnya. Pendekatan seperti ini umum digunakan dalam tugas-tugas regresi dan klasifikasi guna meningkatkan ketepatan prediksi (Fitri, 2023).

Penelitian terdahulu menunjukkan keunggulan masing-masing algoritma dalam memprediksi harga properti. Pada penelitian Warjiyono et al., (2024) menggunakan *Random Forest* dengan data dari *Kaggle* dan mencapai akurasi 75,10%, dengan penurunan harga rata-rata sebesar 16%. Selanjutnya, Ardiansyah, (2024) menunjukkan bahwa model *Random Forest* sangat efektif untuk melakukan prediksi harga properti. Pemodelan ini memiliki akurasi tinggi dengan nilai evaluasi yang unggul nilai *MAE (Mean Absolute Error)* 15,000, dan *R-squared* atau  $R^2$  sebesar 0,90. Kemudian Penelitian Fitri (2023) menggunakan tiga metode, yaitu *Gradient Boosted Trees Regression*, *Linear Regression* dan *Random Forest Regression*, untuk memprediksi harga hunian rumah. Penelitian ini memberikan hasil metode *Random Forest Regression* keluar dengan akurasi prediksi paling tinggi, yaitu sebanyak 81,5%.

Berdasarkan Penelitian sebelumnya menunjukkan perlunya pendekatan yang lebih optimal guna meningkatkan akurasi dalam estimasi harga properti. Penggunaan dua algoritma, yaitu *Random Forest* dan *Gradient Boosting* dengan metode *Hyperparameter Tuning* seperti *Grid Search* dan *RandomizedSearchCV* untuk mengoptimalkan kinerja model. *Random Forest* mempunyai kelebihan dalam mengolah data-data yang lebih kompleks dan bervariasi, sementara *Gradient Boosting* efektif dalam meningkatkan akurasi dengan memperbaiki kesalahan secara bertahap. Dengan memanfaatkan kedua algoritma ini, prediksi harga rumah dapat dilakukan dengan lebih tepat, terutama jika faktor seperti luas lahan, dimensi bangunan, dan total kamar dipertimbangkan dengan cermat. Pendekatan ini dapat mendukung pengambilan keputusan yang lebih efektif bagi pembeli, penjual, maupun pengembang properti.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, rumusan masalah dalam penelitian ini adalah

1. Seperti apa kinerja algoritma *Random Forest* dan *Gradient Boosting* dalam memprediksi harga rumah berdasarkan metrik *RMSE (Root Mean Squared Error)* dan  $R^2$  (*R-squared*) ?
2. Faktor apa saja yang paling memengaruhi prediksi harga rumah berdasarkan kedua algoritma tersebut ?

### 1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Membandingkan kinerja algoritma *Random Forest* dan *Gradient Boosting* dalam memprediksi harga rumah menggunakan metrik RMSE dan  $R^2$ .
2. Mengidentifikasi faktor-faktor utama yang memengaruhi harga rumah.

### 1.4 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah:

1. Bagi Penulis : Memberikan pengalaman langsung dalam menerapkan ilmu dari mata kuliah *Machine Learning* dan *Data Mining* pada masalah nyata di bidang properti.
2. Bagi Masyarakat : Memberikan informasi yang dapat digunakan sebagai panduan untuk memahami faktor-faktor yang memengaruhi harga rumah, sehingga dapat mendukung pengambilan keputusan yang lebih baik dalam jual beli rumah.