

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Performa dari algoritma *Multiple Linear Regression*, *Gradient Boosting*, dan *Random Forest* dalam memprediksi harga rumah di Kota Tangerang Selatan, menunjukkan hasil yang cukup signifikan. Setelah dilakukan serangkaian proses mulai dari *data preparation*, *modeling* hingga *evaluation* menggunakan metrik R^2 (*R Squared*) dan RMSE (*Root Mean Squared Error*), ditemukan bahwa algoritma *Random Forest* menghasilkan performa paling optimal. Model *Random Forest* memperoleh nilai R^2 sebesar 0.8011, yang mengindikasikan hubungan yang sangat kuat antara variabel *input* dan *output*. Selain itu, model ini juga menghasilkan nilai RMSE terendah, yaitu sebesar 536,831,054, yang berarti memiliki tingkat kesalahan prediksi paling kecil di antara ketiga model. Hal ini menunjukkan bahwa *Random Forest* memiliki kemampuan yang lebih baik dalam menangkap pola hubungan *non-linier* dan kompleks dalam data properti dibandingkan *Multiple Linear Regression* dan *Gradient Boosting*, menjadikannya model yang paling akurat dan andal dalam konteks penelitian ini.

Berdasarkan hasil analisis *feature importance* yang dihasilkan dari model *Gradient Boosting* dan *Random Forest*, ditemukan secara konsisten bahwa variabel yang memiliki pengaruh paling absolut dan dominan terhadap prediksi harga adalah luas bangunan dan luas tanah. Kedua fitur ini memberikan kontribusi paling signifikan terhadap nilai akhir harga rumah, mengungguli semua variabel lainnya. Temuan ini menegaskan bahwa model berhasil menangkap esensi fundamental dari valuasi properti, di mana nilai intrinsik sebuah rumah sangat ditentukan oleh dimensi fisiknya, yaitu seberapa besar ruang yang dapat dihuni dan luas lahan yang dimiliki. Hal ini selaras dengan praktik di pasar properti, di mana spesifikasi fisik bangunan dan tanah menjadi pertimbangan kuantitatif utama yang membentuk dasar harga sebelum faktor-faktor lain diperhitungkan. Dengan demikian, penelitian ini mengonfirmasi bahwa pilar utama yang menopang sebagian besar

struktur harga properti di wilayah yang diteliti adalah kedua variabel dimensi tersebut.

Penelitian ini berhasil mengembangkan sebuah sistem berbasis *website* menggunakan *framework Flask*. Sistem ini dirancang untuk menerima *input* spesifikasi rumah dari *user*, seperti lokasi, luas tanah, luas bangunan, jumlah kamar, dan fitur tambahan lainnya, yang kemudian diproses oleh model *machine learning* yang telah dilatih sebelumnya. Hasil dari *input* tersebut akan menghasilkan prediksi harga rumah yang ditampilkan kepada pengguna secara langsung melalui *web interface*. Dengan adanya prototype ini, diharapkan sistem dapat terus dikembangkan ke arah yang lebih kompleks dan interaktif untuk kebutuhan industri properti ke depannya.

5.2 Saran

1. Memperluas cakupan data yang digunakan dalam penelitian dengan menambahkan lebih banyak variabel yang relevan, seperti faktor ekonomi, demografis (jumlah penduduk, tingkat migrasi), dan kondisi pasar properti secara umum. Variabel-variabel ini berpotensi memberikan kontribusi yang signifikan dalam meningkatkan akurasi prediksi harga properti.
2. Mengingat kondisi pasar properti yang dinamis dan cepat berubah, penggunaan data dari berbagai periode waktu sangat penting. Hal ini memungkinkan model untuk belajar dari pola-pola historis dan menjadi lebih adaptif dalam menangani fluktuasi harga di masa mendatang.
3. Meskipun algoritma Random Forest terbukti memberikan hasil terbaik dalam penelitian ini, tetap diperlukan upaya untuk memantau dan mengoptimalkan algoritma. Langkah ini bisa dilakukan dengan mengeksplorasi pengaturan hyperparameter yang lebih tepat guna meningkatkan performa model.
4. Untuk meningkatkan stabilitas dan generalisasi model, disarankan untuk melakukan pelatihan dan pengujian pada *dataset* yang lebih besar dan beragam. *Dataset* yang lebih luas akan memungkinkan model mengenali lebih banyak pola dan variasi dalam data, sehingga meningkatkan kemampuan prediksi.