

## DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	ii
HALAMAN PENGESAHAN .....	iii
KATA PENGANTAR .....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI .....	v
ABSTRAK .....	vi
ABSTRACT .....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL .....	xi
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan dan Manfaat .....	3
1.4 Ruang Lingkup .....	4
1.5 Sistematika Penulisan .....	4
BAB II TINJAUAN PUSTAKA .....	6
2.1 <i>State-of-the-art</i> .....	6
2.2 Pengertian Rumah.....	9
2.3 <i>Machine Learning</i> (ML) .....	9
2.4 <i>Linear Regression</i> .....	11
2.5 <i>Decision Tree</i> .....	12
2.6 <i>Random Forest</i> .....	14
2.6.1 Konsep Dasar <i>Random Forest</i> .....	14
2.6.2 <i>Bagging</i> pada <i>Random Forest</i> .....	16
2.6.3 <i>Random Input</i> pada <i>Random Forest</i> .....	17
2.7 <i>Extreme Gradient Boosting</i> .....	17
2.8 <i>Data Preprocessing</i> .....	17
2.8.1 <i>Data Cleaning</i> .....	18
2.8.1.1 <i>Missing Values</i> .....	18
2.8.1.2 <i>Outliers</i> .....	19
2.8.2 <i>Data Integration</i> .....	20

2.8.3	<i>Data Transformation</i> .....	20
2.8.4	<i>Dimensionality Reduction</i> .....	22
2.8.5	<i>Handling Class Imbalance</i> .....	24
2.9	<i>K-Fold Cross Validation</i> .....	25
2.10	<i>Hyperparameter Tuning</i> .....	25
2.11	Evaluasi Model .....	26
2.11.1	<i>Root Mean Squared Error</i> .....	26
2.11.2	Contoh perhitungan RMSE .....	27
2.11.3	<i>Mean Absolute Error (MAE)</i> .....	28
2.11.4	Contoh Perhitungan MAE .....	28
2.11.5	<i>R-Squared</i> .....	29
2.11.6	Contoh Perhitungan <i>R-Squared</i> .....	29
2.12	<i>Tools dan Library</i> .....	30
2.12.1	<i>Visual Studio Code</i> .....	30
2.12.2	<i>NumPy</i> .....	31
2.12.3	<i>Pandas</i> .....	31
2.12.4	<i>Matplotlib</i> .....	31
2.12.5	<i>Scikit-learn</i> .....	32
BAB III METODOLOGI PENELITIAN .....		33
3.1	Alur Penelitian .....	33
3.2	Penarikan dan Pengubahan Format Data .....	35
3.3	<i>Data Preprocessing</i> .....	38
3.3.1	<i>Data Cleaning</i> terhadap <i>Missing Value</i> .....	38
3.3.2	Pembagian Data Latih dan Data Uji .....	39
3.3.3	<i>Data Cleaning</i> terhadap <i>Outliers</i> .....	42
3.3.3.1.	Duplikasi Data Latih.....	42
3.3.3.2.	Penerapan Proses PCA pada Data Latih Duplikat .....	42
3.3.3.3.	Penerapan DBSCAN pada Data Latih Duplikat.....	51
3.3.3.4.	Proses Penghapusan Baris Data Latih yang Dilabeli sebagai <i>Outliers</i> .....	56
3.3.3.5.	Proses Penghapusan Data Latih Duplikat.....	57
3.3.3.6.	Pengulangan Proses pada Sub Sub Sub Bab 3.3.3.1 hingga 3.3.3.5 .....	57
3.3.4	<i>Data Transformation</i> .....	58
3.4	Pelatihan Model Prediksi Harga Rumah Menggunakan <i>Random Forest</i> .....	60

3.5	<i>Hyperparameter Tuning</i> .....	72
3.6	Evaluasi Model Prediksi Harga Rumah Menggunakan <i>Random Forest</i> .....	73
3.7	Perbandingan Performa Model <i>Random Forest</i> , <i>Linear Regression</i> , dan XGBoost	80
3.8	Perancangan Aplikasi Prediksi .....	80
BAB IV HASIL DAN PEMBAHASAN .....		82
4.1	Lingkungan Proses Pelatihan dan Evaluasi Model.....	82
4.2	Data Penelitian dan Proses <i>Outliers Detection</i> .....	82
4.3	Implementasi Algoritma <i>Random Forest</i> untuk Model Prediksi Harga Rumah dalam Membandingkan Skenario Pengisian Nilai pada <i>Missing Value</i> .....	86
4.4	Implementasi Algoritma <i>Random Forest</i> untuk Model Prediksi Harga Rumah dalam Membandingkan Skenario pada Metode <i>Encoding</i> .....	87
4.5	Perbandingan Performa Model <i>Random Forest</i> dengan Model <i>Linear Regression</i> dan XGBoost .....	89
4.6	<i>Hyperparameter Tuning</i> .....	93
4.7	Pengujian Model <i>Random Forest</i> dalam Memprediksi Harga Rumah pada Data Uji.....	93
4.8	Implementasi Aplikasi Prediksi Harga Rumah.....	94
BAB V PENUTUP .....		97
5.1	Kesimpulan .....	97
5.2	Saran .....	97
DAFTAR PUSTAKA .....		99