

## KOMPARASI ALGORITMA MACHINE LEARNING DALAM MEMPREDIKSI HARGA RUMAH

Cep Haryanto, Nining Rahaningsih, Fadhil Muhammad Basysyar

Program Studi Teknik Informatika S1, STMIK IKMI Cirebon

Jalan Perjuangan No. 10B Majasem, Kota Cirebon, Indonesia

*haryanto.cep@gmail.com*

### ABSTRAK

Manusia memiliki kebutuhan mendasar akan sebuah tempat yang bisa disebut sebagai rumah atau tempat berteduh. Dengan banyaknya pertumbuhan penduduk di Indonesia dari tahun ke tahun membuat tidak semua golongan masyarakat memiliki tempat tinggal yang layak huni. Maka dari itu agar setiap keluarga dapat memiliki tempat tinggal pribadi yang layak diperlukan perencanaan yang matang. Dalam perencanaan tersebut diperlukan perencanaan atau prediksi harga di masa mendatang. Salah satu metode yang bisa digunakan adalah dengan menggunakan algoritma *machine learning*. Adapun beberapa faktor yang mempengaruhi harga rumah diantaranya luas lahan, luas bangunan berdiri, banyaknya kamar tidur, kamar mandi, dan garasi. Metode yang digunakan dalam penelitian ini adalah *multiple linear regression* dan *random forest regression*. Tujuan dari penelitian ini adalah mencari hasil prediksi terbaik antara *multiple linear regression* dan *random forest regression*. Untuk mendapatkan nilai prediksi yang tinggi penelitian dilakukan secara berulang kali dengan 80% dataset untuk training dan 20% dataset untuk *testing* guna mencapai hasil prediksi yang tinggi. Hasil dari penelitian ini diperoleh bahwa algoritma *random forest regression* merupakan algoritma terbaik dalam penelitian ini dengan nilai akurasi sebesar 81,6%.

**Kata Kunci:** *Multiple Linear Regression, Random Forest Regression, Prediksi, Harga Rumah*

### 1. PENDAHULUAN

Rumah merupakan salah satu kebutuhan utama manusia yang berguna sebagai tempat tinggal, tempat bersantai, berkumpul dengan keluarga, tempat untuk berlindung, serta untuk beristirahat setelah melakukan aktifitas sehari-hari[1]. Seperti emas, rumah juga bisa menjadi investasi di masa mendatang dikarenakan harganya yang bisa berubah sewaktu-waktu dan semakin banyak orang yang membutuhkan rumah, terlebih lagi apabila lokasi tersebut dekat dengan sekolah, perkantoran, pusat perbelanjaan, dan sarana transportasi yang mudah pasti akan mempengaruhi harga rumah tersebut dengan cepat[2].

Selain menjadi tempat tinggal dan berkumpul bagi keluarga, rumah juga berfungsi sebagai tempat singgah bagi wisatawan yang berkunjung dari daerah yang jauh[3]. Wisatawan adalah orang yang melakukan perjalanan untuk berlibur dan melakukan rekreasi. Perjalanan wisata terkadang bisa dilakukan dalam waktu sehari-hari jika tempat yang dikunjungi cukup jauh sehingga diperlukan rumah untuk singgah.

Seiring berjalannya waktu, kebutuhan masyarakat semakin berkembang, salah satunya yaitu kebutuhan akan tempat tinggal atau rumah. Pengusaha properti akan berlomba-lomba untuk membangun atau membeli rumah sebagai sarana investasi[4]. Hal ini akan membuat masyarakat yang ingin membeli rumah berfikir apakah rumah yang akan dibeli memiliki nilai keuntungan yang baik atau tidak, dikarenakan harga rumah yang semakin naik setiap waktu[5].

Prediksi merupakan teknik untuk memperkirakan suatu nilai pada masa yang akan datang dengan memperhatikan data atau informasi pada masa lalu

maupun data atau informasi saat ini [6]. Kemampuan untuk melakukan prediksi dengan tingkat akurasi yang dapat dipertanggung jawabkan mampu membantu sebuah perusahaan atau instansi untuk membuat keputusan yang tepat. Pada penelitian lain berjudul “Prediksi harga rumah di kota Bandung timur dengan menggunakan model algoritma *moving average*” menjelaskan bahwa target capaian yaitu menganalisis sistem prediksi harga rumah di sekitar kota Bandung dengan kesimpulan pengujian model algoritma *moving average* mendapatkan nilai rata-rata nilai akurasi 70-90% dan nilai error sekitar 10-20% [7]

Harga setiap rumah setiap tahun semakin meningkat, kenaikan harga ini bisa diukur melalui aspek atau fitur yang dimiliki oleh rumah tersebut[8]. Harga dapat selalu berubah dan tidak dapat terprediksi secara akurat membuat pembeli rumah membutuhkan sistem untuk memperkirakan harga rumah berdasarkan fitur yang dimiliki[9]. Penelitian ini menggunakan algoritma regresi dalam memperhitungkan harga rumah dengan menggunakan model *multiple linear regression* dan *random forest regression* dengan menganalisis data untuk membangun model *machine learning* dengan tujuan memprediksi harga rumah.

### 2. TINJAUAN PUSTAKA

#### 2.1. Penelitian Terdahulu

Pada penelitian Reza Fahlepi dan Andreas Widjaja pada Jurnal Strategi Vol 1 Nomor 2 November 2019 yang berjudul “Penerapan Metode Multiple Linear Regression dalam Memprediksi Harga Sewa Kamar Kost”. Dalam penelitian ini mendiskusikan tentang bagaimana cara memprediksi atau menentukan

biaya sewa untuk tiap kamar kos. Fitur yang terdapat pada kamar kos tersebut juga mempengaruhi harga yang akan ditentukan sehingga tidak merugikan pemilik kos ataupun penyewa kos. Metode analisis yang dilakukan terhadap model membuktikan bahwa model memiliki koefisien korelasi 0,70 dan lebih baik dibandingkan beberapa pengujian lainnya dengan hasil RSME sebesar 334.578,63 [10].

## 2.2. Multiple Linear Regression

*Multiple linear regression* merupakan perluasan dari model linier sederhana[11]. Sementara model regresi linier sederhana hanya terdiri dari variabel bebas dan satu variabel terikat, regresi linear berganda memiliki lebih banyak variabel independen daripada satu variabel dependen. Setelah menjumlahkan variabel bebas, maka bentuk umum persamaan regresi linier yang memuat dua atau lebih variabel bebas adalah sebagai berikut.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

Keterangan:

Y = Variabel terikat atau variabel *response*.

X = Variabel bebas atau variabel *predictor*.

$\alpha$  = Konstanta.

$\beta$  = *Slope* atau Koefisien *estimate*.

## 2.3. Random Forest Regression

*Random forest* merupakan model ansambel yang terdiri dari beberapa model pohon keputusan dalam penerapan metode regresi dan klasifikasi *bootstrap aggregating (bagging)* dan pemilihan fitur acak [12]. *Decision tree* atau pohon keputusan adalah bagan alir yang dirancang sebagai pohon dengan simpul akar yang digunakan untuk mengumpulkan informasi. *Decision tree* mengklasifikasikan sampel data yang kelasnya tidak diketahui ke dalam kelas yang ada. Tujuan dari *decision tree* adalah untuk menghindari *overfitting* kumpulan data dan mencapai akurasi maksimum.

*Random forest* adalah kombinasi dari pohon(*tree*) yang kemudian digabungkan menjadi satu model. *Random forest* bergantung pada nilai vektor acak yang memiliki distribusi yang sama di semua pohon(*tree*), dengan setiap *decision tree* memiliki kedalaman maksimum. *Random forest* adalah pengklasifikasi yang berbentuk pohon  $\{h(\mathbf{x}, \theta_k), k = 1, \dots\}$  dimana  $\theta_k$  adalah vektor acak yang terdistribusi secara independen dan setiap pohon unit memilih kelas paling populer dari input  $\mathbf{x}$ .

## 2.4. Jupyter Notebook

Jupyter Notebook adalah singkatan dari tiga bahasa pemrograman, yaitu Ju (Julia), Py (Python) dan R. Ini adalah alat ilmuwan data populer untuk pemrosesan data yang memungkinkan integrasi kode dan keluaran secara interaktif ke dalam satu dokumen. termasuk kode langsung, persamaan, visualisasi, dan teks naratif yang kaya. Kemudahan menulis dan berbagi teks dan kode dengan program ini juga menjadikannya bagus untuk kolaborasi.

## 3. METODE PENELITIAN

Metode yang di gunakan dalam perancangan ini menggunakan metode deskriptif kuantitatif. Dimana penelitian kuantitatif adalah mengeksplorasi suatu kondisi atau keadaan yang secara sistematis memecahkan masalah berdasarkan data dan fakta yang ada, dan data kuantitatif ialah data yang berupa angka dan dapat digunakan dalam proses operasi matematika serta data ini bersumber dari fakta yang ada. Data yang digunakan yaitu data sekunder yang didapatkan dari kaggle.com.

### 3.1. Pengumpulan Data

#### 1. Observasi

Observasi adalah kegiatan penelitian yang melibatkan pengumpulan informasi yang berkaitan dengan masalah penelitian bersumber dari data publik yang didapat dari kaggle.com.

#### 2. Studi Pustaka

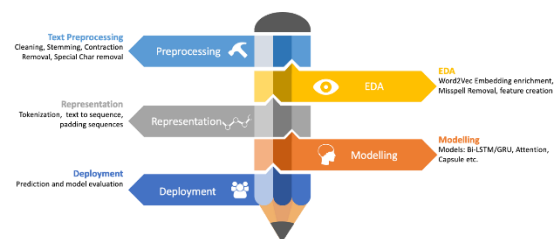
Merupakan metode yang dilakukan dengan mengambil referensi yang terpercaya karena dibutuhkan untuk laporan penelitian.

#### 3. Dokumentasi

Penggunaan literatur telah lama digunakan sebagai sumber informasi dalam penelitian, seperti halnya literatur sebagai sumber informasi digunakan untuk penelitian, interpretasi bahkan prediksi.

### 3.2. Exploratory Data Analysis (EDA)

*Exploratory Data Analysis* adalah proses penelitian pertama untuk mengidentifikasi pola, menemukan *outlier*, menguji hipotesis dan memverifikasi asumsi. EDA sangat berguna untuk pendeteksian kesalahan secara dini karena pengguna dapat menemukan anomali dalam data, mengetahui hubungan antar data dan mengekstrak faktor-faktor penting dari data tersebut. Proses EDA ini sangat berguna dalam proses analisis statistik.



Gambar 1. Alur penelitian menggunakan EDA

Dalam hal *exploratory data analysis*, terdapat beberapa teknik yang dapat digunakan dalam pengolahan data, yaitu:

### 3.3. Pre-processing

*Pre-processing* ialah melakukan pengolahan data agar data yang salah dapat diperbaiki atau dibersihkan sehingga data dapat digunakan. Dalam *pre-processing* data awal yang digunakan masih merupakan data mentah, kemudian dilakukan proses untuk menyesuaikan dengan data yang dibutuhkan.

### 3.4. Exploratory Data Analysis

*Exploratory Data Analysis* adalah cara memeriksa data yang dimiliki dan menentukan bagaimana memproses data tersebut. Pada tahap ini dilakukan pengecekan data kosong, menghapus data yang sama dan mengkonversi kategori data.

### 3.5. Representation

Pada tahap ini, data diubah dari data mentah menjadi tampilan atau visualisasi. Data dapat divisualisasikan dalam bentuk boxplot, histogram, dan bentuk lainnya.

### 3.6. Modelling

Modelling ialah penerapan model algoritma, dimana data akan diproses lebih lanjut untuk mencari kesimpulan dari data. Proses ini akan menghasilkan hasil yang berbeda tergantung dari karakteristik data.

### 3.7. Deployment / Evaluation

Pada tahap ini dilakukan penarikan kesimpulan dari hasil data *mining*. Kesimpulan akhir ditarik berdasarkan berbagai hipotesis yang dihasilkan dari proses data *mining*.

## 4. HASIL DAN PEMBAHASAN

Penelitian prediksi harga rumah ini menggunakan metode *Exploratory Data Analysis* (EDA). *Exploratory Data Analysis* (EDA) adalah suatu proses dalam menganalisis sekumpulan data untuk meringkas karakteristik utamanya agar memahami kondisi dataset. Pada umumnya *Exploratory Data Analysis* (EDA) menggunakan metode visualisasi seperti Histogram, Box Plot, Violin Plot. Berikut proses penelitian ini dengan menggunakan *Exploratory Data Analysis* (EDA).

### 4.1. Import Library

Pada tahap ini diharuskan untuk melakukan *import library* sehingga dapat memanggil dan

menggunakan modul yang tersedia dalam bahasa Python. *Library* yang digunakan dapat dilihat pada gambar 2.

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
from sklearn.ensemble import RandomForestRegressor
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
from sklearn.model_selection import train_test_split
```

Gambar 2. Import Library

Adapun fungsi *library* yang diimport pada gambar 4.1 adalah sebagai berikut.

- Pandas adalah *library open source* yang ada pada bahasa pemrograman python yang sering digunakan untuk memproses data, manipulasi data, hingga melakukan analisis data.
- Numpy adalah *library* dalam bahasa pemrograman python yang berfungsi untuk proses komputasi numerik.
- Matplotlib digunakan untuk memvisualisasikan data dalam 2D dan 3D dan membuat gambar berkualitas tinggi yang dapat disimpan dalam berbagai format gambar seperti JPEG dan PNG.
- Seaborn adalah *library* yang digunakan untuk membuat grafik dan statistik. *Library* ini dibangun berdasarkan *library matplotlib* kemudian diintegrasikan dengan struktur data pandas.
- Sklearn adalah *library* dalam bahasa pemrograman Python yang berfungsi untuk membangun model *machine learning*, seperti regresi, klasifikasi, dll.

### 4.2. Data Pre processing

Pada tahap ini, dataset harga rumah akan diseleksi. Seleksi ini berfungsi untuk membuang data yang tidak digunakan. Data yang dibuang ialah data yang tidak valid, data yang tidak konsisten, dan data yang ganda.

Tabel 1. Data sebelum pre processing

NO	NAMA RUMAH	HARGA	LB	LT	KT	KM	GRS
1	Rumah Murah Hook Tebet Timur	3800000000	220	220	3	2	0
2	Rumah Modern di Tebet dekat Stasiun	4600000000	180	137	4	3	2
3	Rumah Mewah 2 Lantai	3000000000	267	250	4	4	4
4	Rumah Baru Tebet	430000000	40	25	2	2	0
5	Rumah Bagus Tebet	9000000000	400	355	0	5	3

Pada tabel 1 terdapat data yang tidak valid, yaitu data nomor sehingga data tersebut harus dilakukan

preprocessing untuk menghilangkan data tersebut sehingga menjadi seperti tabel 2.

Tabel 2. Data setelah pre processing

No	NAMA RUMAH	HARGA	LB	LT	KT	KM	GRS
1	Rumah Murah Hook Tebet Timur	3800000000	220	220	3	2	0
2	Rumah Modern di Tebet dekat Stasiun	4600000000	180	137	4	3	2
3	Rumah Mewah 2 Lantai	3000000000	267	250	4	4	4
4	Rumah Baru Tebet	430000000	40	25	2	2	0
5	Rumah Bagus Tebet	9000000000	400	355	0	5	3

### 4.3. Exploratory Data Analysis

*Exploratory Data Analysis* adalah cara memeriksa data yang dimiliki dan menentukan bagaimana memproses data. Pada tahap ini dilakukan pengecekan data kosong, menghapus data yang ganda

dan mengkonversi kategori data. Setelah data bersih dari data kesalahan, selanjutnya dilakukan statistik deskriptif untuk merangkum dan mendeskripsikan data sebagai berikut.

Tabel 3. Deskripsi data

	HARGA	LB	LT	KT	KM	GRS
Count	1.010000e+03	1010	1010	1010	1010	1010
Mean	7.628987e+09	276.5	237.4	4.6	3.6	1.9
Std	7.340946e+09	177.8	179.9	1.5	1.4	1.5
Min	4.300000e+08	40	25	2	1	0
25%	3.262500e+09	150	130	4	3	1
50%	5.000000e+09	216.5	165	4	3	2
75%	9.000000e+09	350	290	5	4	2
max	6.500000e+10	1126	1400	10	10	10

Pada tabel 3 dapat dilihat nilai dari data, mulai dari total (*count*), standar deviasi, nilai rata-rata (*mean*), nilai minimal (*min*), serta nilai maksimal (*max*).

Tabel 4. Informasi tipe data

#	Column	Non-Null Count	Dtype
0	NAMA RUMAH	1010 non-null	Object
1	HARGA	1010 non-null	Int64
2	LB	1010 non-null	Int64
3	LT	1010 non-null	Int64
4	KT	1010 non-null	Int64
5	KM	1010 non-null	Int64
6	GRS	1010 non-null	Int64

Pada tabel 4 dapat dilihat informasi dari tipe data yang berjumlah 1010 *record*. Kolom harga rumah memiliki tipe data *object*, sedangkan kolom harga, lb, lt, kt, km, dan grs memiliki tipe data *int64*. Karena tipe data yang akan dilakukan proses sudah numerik, jadi tidak perlu melakukan perubahan tipe data. Kemudian, dilakukan pengecekan apakah terdapat data yang kosong, seperti pada tabel 5.

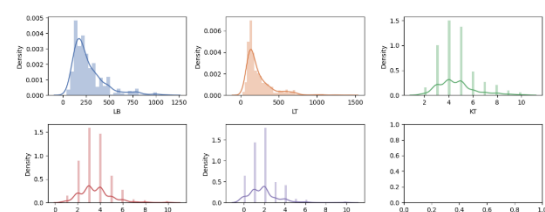
Tabel 5. Cek missing value

NAMA RUMAH	0
HARGA	0
LB	0
LT	0
KT	0
KM	0
GRS	0
Dtype : int64	

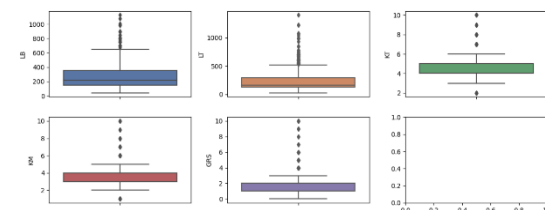
Berdasarkan tabel 5 pada pengecekan *missing value*, tidak ada data yang kosong sehingga dapat dilanjutkan ke tahapan selanjutnya.

### 4.4. Representation

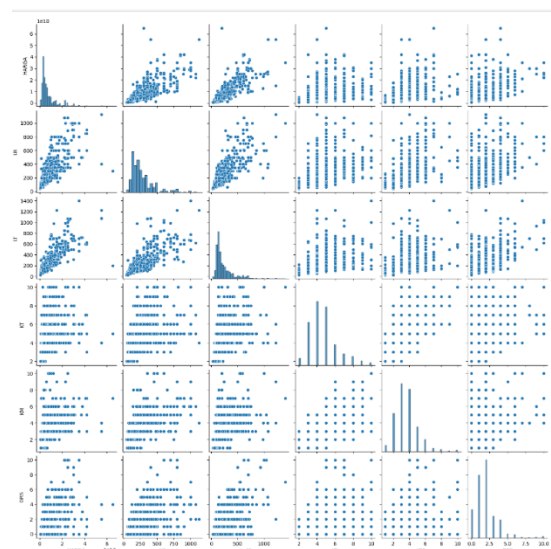
Pada tahap ini dilakukan visualisasi dari data rumah, sehingga dapat dilihat persebaran dari data seperti pada gambar 3, 4 dan 5.



Gambar 3. Distribusi data rumah (displot)



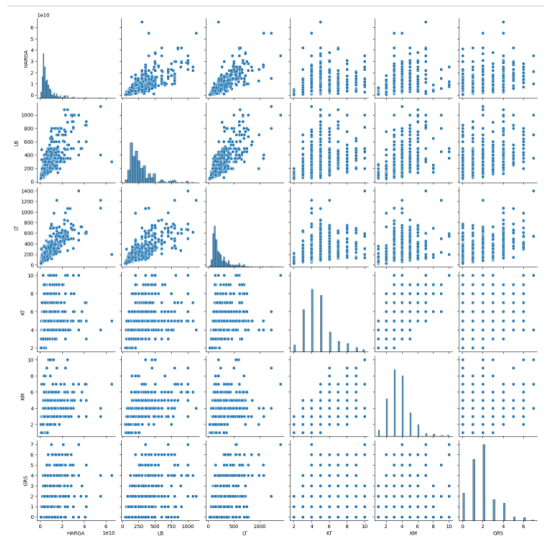
Gambar 4. Distribusi data rumah (boxplot)



Gambar 5. Persebaran data rumah (pairplot)

Pada gambar 3, 4 dan 5 dapat dilihat persebaran data dari data rumah mulai dari luas bangunan, luas tanah, kamar tidur, kamar mandi, dan garasi. Dapat dilihat terdapat beberapa *outlier* pada data yang harus dibersihkan agar data menjadi lebih bersih. Outlier

yang akan dibersihkan hanya 1% dari data rumah, sehingga data yang awalnya 100% menjadi 99%. Data yang telah dibersihkan outliernya dapat dilihat pada gambar 6 berikut.



Gambar 6. Persebaran data rumah setelah dibersihkan

#### 4.5. Matriks Korelasi

Matriks korelasi adalah suatu tabel yang menunjukkan hubungan linier antara dua atau lebih variabel. Matrix korelasi menggunakan koefisien korelasi Pearson sebagai ukurannya, yang bernilai antara -1 sampai +1. Nilai positif menggambarkan hubungan positif antar variabel, artinya jika nilai salah satu variabel naik, maka variabel lain juga naik. Nilai negatif menunjukkan hubungan negatif antar variabel, artinya jika nilai salah satu variabel naik, maka variabel lain akan turun. Nilai 0 menunjukkan tidak ada hubungan linier antar variabel.

	HARGA	LB	LT	KT	KM	GRS
HARGA	1.00	0.75	0.81	0.32	0.40	0.48
LB	0.75	1.00	0.74	0.44	0.53	0.49
LT	0.81	0.74	1.00	0.43	0.39	0.48
KT	0.32	0.44	0.43	1.00	0.67	0.28
KM	0.40	0.53	0.39	0.67	1.00	0.35
GRS	0.48	0.49	0.48	0.28	0.35	1.00

Gambar 7. Matriks korelasi

Pada gambar 7 dapat dilihat bahwa warna korelasi koefisien yang ungu menunjukkan hubungan linear yang positif yang kuat. Sedangkan warna putih menunjukkan hubungan linear yang negatif yang berarti tidak ada korelasi antar variabel.

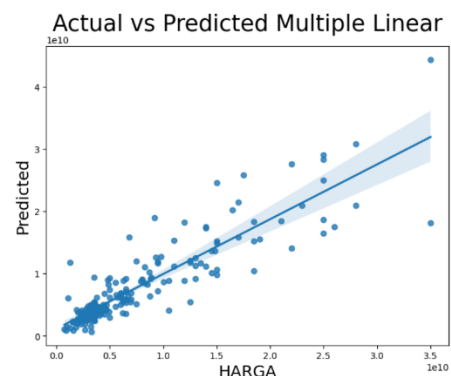
#### 4.6. Multiple Linear Regression

Perhitungan *multiple linear regression* dilakukan dengan menggunakan library sklearn. Dataset berjumlah 1010 dibagi menjadi 80% data training (808 data) dan 20% data testing (202 data). Langkah selanjutnya ialah menetapkan kolom yang menjadi variabel dependen (y) yaitu kolom harga dan kolom variabel independen (x) yaitu LB, LT, KT, KM, dan GRS. Proses *multiple linear regression* seperti pada tabel 6.

Tabel 6. Perintah Python dalam multiple linear regression

1	<code>x=data.drop(columns=['HARGA', 'NAMA RUMAH']) y = data['HARGA']</code>	Menetapkan kolom yang akan menjadi variabel dependen (y) dan variabel independen (x)
2	<code>x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=2)</code>	Membagi data menjadi data <i>training</i> dan data <i>test</i> . Data <i>training</i> sebanyak 80% dan data <i>test</i> sebanyak 20%
3	<code>lin_reg = LinearRegression()</code>	Membuat model
4	<code>lin_reg.fit(x_train, y_train)</code>	<i>linear regression</i>
5	<code>print ("Coefficient of determination : ",r2_score(y_test,y_pred)) print ("MSE: ",mean_squared_error(y_test,y_p red)) print("RMSE: ",np.sqrt(mean_squared_error(y_ test,y_pred)))</code>	Coefficient of determination : 0.7851252699343726 MSE: 9.633962972304015e +18 RMSE : 3103862589.146629
6	<code>y_pred=lin_reg.predict(x_test)</code>	Menghitung nilai prediksi
7	<code>plt.figure(figsize=(8,6)) plt.title("Actual vs. predicted",fontsize=25) plt.xlabel("Actual",fontsize=18) plt.ylabel("Predicted", fontsize=18) #plt.scatter(x=test_y,y=test_predi ct) sns.regplot(x=y_test, y=ypredict) plt.show()</code>	Memvisualisasikan nilai y awal dan nilai y prediksi

Berikut adalah visualisasi dari proses regresi linear.



Gambar 8. Visualisasi data multiple linear



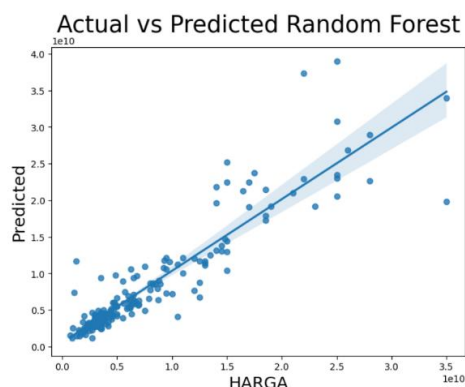
#### 4.7. Random Forest Regression

Perhitungan *random forest* dilakukan dengan menggunakan library sklearn. Dataset berjumlah 1010 dibagi menjadi 80% data training (808 data) dan 20% data testing (202 data). Langkah selanjutnya ialah menetapkan kolom yang menjadi variabel dependen (y) yaitu kolom harga dan kolom variabel independen (x) yaitu LB,LT,KT,KM, dan GRS. Proses *random forest* seperti pada tabel 7.

Tabel 7. Perintah Python dalam random forest

1	<code>x = data.drop(columns=['HARGA','NAMA RUMAH']) y = data['HARGA']</code>	Menetapkan kolom yang akan menjadi variabel dependen (y) dan variabel independen (x)
2	<code>x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=2)</code>	Membagi data menjadi data <i>training</i> dan data <i>test</i> . Data <i>training</i> sebanyak 80% dan data <i>test</i> sebanyak 20%
3	<code>random_forest = RandomForestRegressor(n_estimators=100, random_state=0)</code>	Membuat model <i>random forest</i> dengan 100 <i>tree</i> atau pohon.
4	<code>random_forest.fit(x_train, y_train)</code>	
5	<code>print ("Coefficient of determination :",r2_score(y_test,y_pred)) print ("MSE: ",mean_squared_error(y_test,y_pred)) print("RMSE: ",np.sqrt(mean_squared_error(y_test,y_pred)))</code>	Coefficient of determination : 0.8162183421218491 MSE: 8.239897201713704e+18 RMSE : 2870522113.08565
6	<code>y_pred=random_forest.predict(x_test)</code>	Menghitung nilai prediksi
7	<code>plt.figure(figsize=(8,6)) plt.title("Actual vs. predicted",fontsize=25) plt.xlabel("Actual",fontsize=18) plt.ylabel("Predicted",fontsize=18) #plt.scatter(x=test_y,y=test_y_pred) sns.regplot(x=y_test, y=y_pred) plt.show()</code>	Memvisualisasikan nilai y awal dan nilai y prediksi

Berikut adalah visualisasi dari proses *random forest*.



Gambar 9. Visualisasi data random forest

#### 4.8. Deployment / Evaluation

Pada algoritma *multiple linear regression* di dapatkan nilai akurasi sebesar 78,5% sedangkan pada algoritma *random forest regression* di dapatkan nilai akurasi sebesar 81,6%.

#### 5. KESIMPULAN DAN SARAN

Berdasarkan hasil uji pada kedua model algoritma, model *multiple linear regression* di dapatkan nilai akurasi sebesar 78,5% sedangkan pada model algoritma *random forest regression* di dapatkan nilai akurasi sebesar 81,6%. Sehingga model algoritma *random forest* lebih akurat untuk model regresi pada penelitian ini.

Untuk peneliti selanjutnya yang ingin meneliti dengan topik yang sama disarankan untuk melakukan penelitian ini dengan algoritma lain yang lain.

#### DAFTAR PUSTAKA

- [1] R. Mahendra Sanusi, A. Siswo, R. Ansori, R. Wijaya, and S. Si, "PREDIKSI HARGA RUMAH DI KOTA BANDUNG BAGIAN TIMUR DENGAN MENGGUNAKAN METODE REGRESI PREDICTION OF HOUSE PRICES IN THE EAST BANDUNG CITY USING THE REGRESSION METHOD." [2] U. Athiyah, A. Hananta, T. Maulidi, V. M. Eka Putra, T. F. H. Purba, and E. A. W. Bakowatun, "Sistem Pendukung Keputusan Prediksi Harga Rumah Kost untuk Mahasiswa IT Telkom Purwokerto Menggunakan Metode Fuzzy Tsukamoto Berbasis Web," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 1, no. 2, pp. 77–81, 2021, doi: 10.20895/dinda.v1i2.202.
- [3] A. Amalia, M. Radhi, S. H. Sinurat, D. R. H. Sitompul, and E. Indra, "Prediksi Harga Mobil Menggunakan Algoritma Regresi Dengan Hyper-Parameter Tuning," *J. Sist. Inf. dan Ilmu Komput. Prima(JUSIKOM PRIMA)*, vol. 4, no. 2, pp. 28–32, 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2479.
- [4] E. Febrion Rahayuningtyas, F. Novia Rahayu, Y. Azhar, and I. Artikel, "Prediksi Harga Rumah Menggunakan General Regression Neural Network," *J. Inform.*, vol. 8, no. 1, 2021, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Real>
- [5] I. I. Ridho *et al.*, "METODE NEURAL NETWORK UNTUK PENENTUAN AKURASI PREDIKSI HARGA RUMAH," 2022.
- [6] E. A. Saputri and Ekojono, "Prediksi Volume Impor Beras Nasional Menggunakan Jaringan Saraf Tiruan Metode ELM ( Extreme Learning Machine )," *Semin. Nas. Teknol. Inf. dan Apl.*, vol. 10, 2018.
- [7] A. Fadilah, A. Siswo, R. Ansori, R. Wijaya, and S. Si, "PREDIKSI HARGA RUMAH DI KOTA BANDUNG BAGIAN TIMUR DENGAN MENGGUNAKAN METODE MOVING

- AVERAGE HOME PRICE PREDICTION IN THE EAST BANDUNG CITY WITH MOVING AVERAGE.”
- [8] M. Labib Mu'tashim, S. A. Damayanti, H. N. Zaki, T. Muhayat, and R. Wirawan, “Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression,” vol. 3, p. 2021.
  - [9] A. Saiful, “Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression,” *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 41–50, 2021, doi: 10.35957/jatisi.v8i1.701.
  - [10] M. R. Fahlepi and A. Widjaja, “Penerapan Metode Multiple Linear Regression Untuk Prediksi Harga Sewa Kamar Kost,” *J. Strateg.*, vol. 1, no. November, pp. 615–629, 2019.
  - [11] K. Puteri and A. Silvanie, “Machine Learning Untuk Model Prediksi Harga Sembako Dengan Metode Regresi Linier Berganda,” *J. Nas. Inform.*, vol. 1, no. 2, pp. 82–94, 2020.
  - [12] T. Purwa, “Perbandingan Metode Regresi Logistik dan Random Forest untuk Klasifikasi Data Imbalanced (Studi Kasus: Klasifikasi Rumah Tangga Miskin di Kabupaten Karangasem, Bali Tahun 2017),” *J. Mat. Stat. dan Komputasi*, vol. 16, no. 1, p. 58, 2019, doi: 10.20956/jmsk.v16i1.6494.