

ESTIMASI HARGA MOBIL BEKAS TOYOTA YARIS MENGGUNAKAN ALGORITMA REGRESI LINIER

Mohammad Aji Saputra¹, Martanto², Umi Hayati³

^{1,3}Teknik Informatika, ²Manajemen Informatika, STMIK IKMI Cirebon
Jalan Perjuangan No. 10B Karyamulya Kec. Kesambi Kota Cirebon, Jawa Barat 45131
majisaputra464@gmail.com

ABSTRAK

Industri otomotif mengalami perkembangan pesat di era modern ini, didorong inovasi teknologi dan persaingan antar perusahaan mobil untuk menghasilkan produk berkualitas. Mobil saat ini sudah menjadi kebutuhan bagi berbagai kalangan konsumen sehingga memerlukan kreativitas dalam pengembangan produk. Pesatnya pertumbuhan sektor otomotif dan teknologi menyebabkan diperkenalkannya berbagai jenis kendaraan, berdampak signifikan terhadap kenaikan harga mobil baru. Hal ini menyebabkan penjualan dan minat terhadap mobil bekas, sehingga menjadi alternatif yang lebih terjangkau bagi konsumen. Penelitian ini bertujuan mengembangkan model estimasi harga mobil bekas toyota yaris dengan menggunakan algoritma regresi linier. Penerapan data mining, khususnya regresi linier, digunakan untuk mengeksplorasi pola dan informasi tersembunyi dalam dataset mobil bekas toyota yaris. Hasil penelitian menunjukkan bahwa model regresi linier mampu memberikan estimasi harga mobil bekas toyota yaris dengan tingkat akurasi yang baik. Estimasi harga mobil berdasarkan inputan seperti tahun = 2022, transmisi = 0, jarak tempuh = 1113 km, jenis bahan bakar = 0, pajak = 145, konsumsi bahan bakar = 47,9 km per galon, dan ukuran mesin = 1.5, estimasi harga sebesar 220.973.121,60 dalam satuan rupiah. Evaluasi model menghasilkan nilai *MSE* sebesar 886,503, *RMSE* 941,543, dan *R2-Score* 0,861, menunjukkan bahwa model regresi linier dapat digunakan dalam estimasi harga jual mobil bekas toyota yaris berdasarkan spesifikasi.

Kata kunci: Industri Otomotif, Estimasi, Harga, Regresi Linier, Data Mining

1. PENDAHULUAN

Industri otomotif modern mengalami kemajuan pesat dalam inovasi teknologi dan persaingan antar produsen mobil untuk mengembangkan produk berkualitas tinggi yang menarik bagi konsumen. Mobil kini sudah menjadi kebutuhan berbagai kalangan dan banyak digunakan dalam industri otomotif dan transportasi. Oleh karena itu, dalam merancang produk, produsen mobil harus terus mempertimbangkan kreativitas sebagai pertimbangan utama dalam keputusan pembelian konsumen[1].

Dengan pesatnya perkembangan mobil dan teknologi, berbagai jenis kendaraan telah diperkenalkan untuk memenuhi kebutuhan masyarakat. Namun peningkatan fitur-fitur canggih pada mobil baru menyebabkan kenaikan harga kendaraan yang signifikan. Inovasi mobil baru meningkatkan penjualan mobil bekas, dan mobil bekas umumnya lebih terjangkau dan tersedia melalui dealer mobil bekas, media sosial, atau pemilik langsung. Dalam jual beli mobil bekas, menentukan nilai penilaian menjadi salah satu kendalanya. Memperkirakan harga mobil bekas penting dilakukan untuk memastikan konsumen mendapatkan harga beli yang diinginkan dan sesuai[2]. Keberhasilan penjualan mobil bekas sangat bergantung pada harga yang optimal. Harga yang terlalu tinggi dapat meningkatkan persaingan dan mempersulit penjualan, dan harga yang terlalu rendah dapat berdampak negatif terhadap keuntungan. Oleh karena itu, informasi harga mobil bekas di pasaran menjadi alat bantu yang sangat diperlukan bagi konsumen untuk membuat estimasi

harga berdasarkan spesifikasi seperti *price* (harga), *year* (tahun), *transmission* (transmisi), *mileage* (jarak tempuh), *fuelType* (jenis bahan bakar), *tax* (pajak), *mpg* (miles per gallon), dan *engineSize* (ukuran mesin) [3].

Penelitian serupa telah dilakukan sebelumnya oleh Susanti dan Sussolaikah [4] mengenai penerapan metode regresi linier untuk memprediksi harga jual yaris dan jazz bekas di wilayah DKI Jakarta. Metode yang digunakan adalah regresi linier. Penelitian ini juga bertujuan untuk mengevaluasi keakuratan prediksi harga jual mobil bekas dengan menguji model regresi, uji F-statistik, dan uji multikolinearitas. Berdasarkan penelitian, faktor-faktor seperti warna, model mobil, transmisi, dan tahun model berpengaruh positif terhadap harga jual mobil yaris dan jazz bekas di wilayah DKI Jakarta. Mobil dengan transmisi otomatis cenderung memiliki harga jual yang lebih tinggi dibandingkan mobil dengan transmisi manual, dan tahun pembuatan mobil tersebut juga berdampak positif terhadap harga jual. Hasil uji F menunjukkan bahwa seluruh variabel independen berpengaruh signifikan terhadap harga jual mobil.

Penelitian selanjutnya yang dilakukan oleh Nafi'iyah [5] tentang Prediksi Penjualan di Toko Makmur Jaya Elektronik bertujuan untuk memperkirakan jumlah persediaan di bulan berikutnya. Pendekatan ini menggunakan regresi linier berganda dengan jenis produk dan bulan sebagai variabel input, sementara persediaan sebagai variabel output. Dalam penelitian ini, diterapkan sistem peramalan penjualan menggunakan regresi linier

berganda dengan dua faktor prediksi untuk kulkas dan TV dari total 72 sampel. Penelitian yang dilakukan oleh Pradito dan Purnia [6] membahas perbandingan efektivitas Algoritma Regresi Linier dan Neural Network dalam memprediksi nilai kurs mata uang. Fokus dari penelitian ini adalah untuk mengevaluasi kinerja kedua algoritma tersebut dalam meramalkan nilai tukar dan untuk menentukan algoritma yang paling optimal dalam mencapai tujuan tersebut. Penelitian ini melibatkan prediksi nilai tukar dengan menggunakan operator window, serta melakukan analisis data menggunakan kedua algoritma tersebut untuk mendapatkan tingkat akurasi prediksi. Hasil pengujian menunjukkan bahwa Algoritma Regresi Linier menunjukkan kinerja yang sedikit lebih baik dalam memprediksi nilai tukar mata uang.

Berdasarkan penelitian sebelumnya, penelitian ini bertujuan untuk mengembangkan model estimasi harga mobil bekas toyota yaris dengan menggunakan algoritma regresi linier. Perbedaan penelitian ini dengan penelitian sebelumnya adalah bahwa penelitian ini berfokus pada pengembangan model estimasi harga mobil bekas Toyota Yaris dengan menggunakan algoritma regresi linier berganda untuk mengidentifikasi pengaruh atribut-atribut berdasarkan spesifikasi mobil bekas toyota yaris dengan menggunakan nilai evaluasi *MSE (Mean Squared Error)*, *RMSE (Root Mean Squared Error)*, dan *R2-Score*. Metode Regresi linier berganda merupakan metode yang digunakan untuk mengidentifikasi ada atau tidak pengaruh atribut yang signifikan dua atau lebih variabel bebas terhadap variabel terikat pada data daftar mobil toyota yaris [7]. Hasil model regresi linier ini digunakan sebagai evaluasi hasil estimasi harga mobil bekas toyota yaris berdasarkan spesifikasi.

2. TINJAUAN PUSTAKA

2.1. Estimasi

Estimasi adalah proses menentukan perkiraan yang digunakan untuk berbagai tujuan meskipun data masukan tidak lengkap, tidak pasti, atau tidak stabil. Perkiraan di sini mencakup perkiraan, penilaian, dan opini. Secara kuantitatif, perkiraan adalah suatu pengukuran yang menggunakan hasil data numerik, dan keakuratannya dapat diukur dengan menggunakan angka. Dalam praktiknya, estimasi melibatkan penggunaan estimator untuk menghasilkan nilai parameter, yang menggambarkan pengukuran berdasarkan hasil kuantitatif dan presisi yang diukur secara numerik [3].

2.2. Metode KDD

Knowledge Discovery in Database (KDD) merupakan pendekatan yang memanfaatkan pengetahuan dan analitik untuk mengekstrak informasi dari database yang ada [6]. Tahapan dalam metodenya yaitu:

a. Data Selection

Data Selection adalah proses memilih data dari suatu kumpulan untuk diproses sebagai tujuan.

b. Data Preprocessing

Data preprocessing adalah tahap awal pengolahan data yang mempersiapkan data untuk diproses guna menghindari data yang kosong atau *missing value*.

c. Data Transformation

Transformation merupakan adaptasi data yang akan diolah dengan model atau algoritma yang digunakan pada tahap pengolahan data.

d. Data Mining

Data Mining proses data mining melibatkan pencarian dan penggalian pengetahuan untuk mendapatkan model yang berguna dan bermakna.

e. Evaluation

Tahap evaluasi melibatkan penyajian hasil model yang diperoleh dan menguji keakuratan dan kesesuaiannya.

2.3. Regresi Linier Berganda

Proses dalam mengestimasi harga mobil bekas toyota yaris pada penelitian kali ini menggunakan metode regresi linear berganda. Metode regresi linier berganda digunakan untuk menguji pengaruh beberapa variabel independen terhadap variabel dependen. Persamaan regresi linier berganda [8].

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX \quad (1)$$

Keterangan:

Y = variabel terikat

a = konstanta

b_1, b_2 = koefisien regresi

X_1, X_2 = variabel bebas

2.4. MSE

Mean squared error (MSE) merupakan rata-rata nilai error kuadrat antara nilai aktual dan prediksi. Nilai *MSE* biasa digunakan untuk menentukan nilai estimasi error dari suatu nilai prediksi. Nilai *MSE* yang mendekati nol atau nilai terendah berarti hasil estimasi sesuai dengan nilai sebenarnya dan model dapat digunakan untuk menentukan nilai estimasi. Persamaan untuk menghitung nilai *MSE* [9].

$$MSE = \frac{\sum_{t=1}^n (y'_t - y_t)^2}{n} \quad (2)$$

Y'_t = nilai data sebenarnya

Y_t = nilai sebenarnya

n = banyaknya data yang diuji

2.5. RMSE

Root mean square error (RMSE) adalah metode penghitungan yang menghitung selisih antara nilai aktual dan nilai prediksi. Nilai *RMSE* merupakan hasil akar kuadrat dari perhitungan *MSE* pada rumus. Nilai

RMSE dengan nilai lebih rendah pasti bisa lebih akurat dibandingkan metode prediksi dengan nilai RMSE lebih tinggi. Persamaan untuk menghitung nilai RMSE [10].

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (Y't - Y_t)^2}{n}} \quad (3)$$

Y't = nilai data sebenarnya
 Yt = nilai sebenarnya
 n = banyaknya data yang diuji

2.6. Koesfisen Determinasi

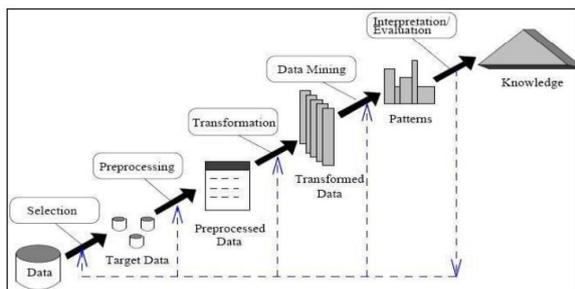
Koefisien Determinasi (R^2) adalah rasio variabilitas nilai yang dimodelkan terhadap variabilitas nilai data asli. Umumnya R^2 digunakan sebagai informasi tentang model fit. Regresi linier menggunakan R^2 sebagai ukuran seberapa baik garis regresi mendekati nilai dari data model asli. Persamaan menghitung nilai R^2 [11].

$$R^2 = \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y}_i)^2} \quad (4)$$

R^2 = Koefisien Determinan
 Yi = Nilai Pengamatan
 \hat{Y}_i = Nilai Y
 \bar{Y}_i = Nilai Rata-rata pengama

3. METODE PENELITIAN

Dalam penelitian ini, *Knowledge Discovery in Databases (KDD)* sebagai metode utama untuk mengeksplorasi dan menganalisis data. Langkah-langkah KDD melibatkan data *selection*, data *preprocessing*, data *transformation*, data mining, dan *evaluation*, berikut penjelasan tahapannya yaitu:



Gambar 1. Metode penelitian (KDD)

Sumber: <https://www.researchgate.net/figure/DataMining-in-KnowledgeDiscovery-KDD->

3.1. Data Selection

Seleksi data adalah langkah mengumpulkan dan memilih informasi. Tahapan ini menggunakan data yang tersedia atau *open public*, khususnya data dari kaggle.com, sebuah website yang menyediakan daftar harga mobil bekas toyota. Data tersebut kemudian dipilih untuk fokus pada daftar mobil bekas satu model yaitu toyota yaris. Proses pemilihan data ini menggunakan *software microsoft excel* dengan tujuan

untuk menyederhanakan analisis dan memfokuskan pada satu model mobil bekas yaitu toyota yaris.

3.2. Data Preprocessing

Preprocessing merupakan bagian dari penyiapan informasi sebelum data mining dimulai. Sebagai persiapan sebelum estimasi, langkah dalam *preprocessing* menjadi penting. Data kosong atau nilai yang hilang *missing value* untuk memastikan kebersihan dan integritas data. Proses *preprocessing* ini dilakukan dengan google colab dan memastikan keandalan hasil estimasi karena data yang bersih dan lengkap meningkatkan akurasi dan keandalan model yang digunakan untuk analisis.

3.3. Data Transformation

Pada langkah ini, data diubah menjadi format yang sesuai untuk pemrosesan data mining. Tujuan dari proses ini adalah untuk menyederhanakan data agar dapat diproses secara efektif dengan algoritma regresi linier dan perangkat lunak penelitian yang digunakan.

3.4. Data Mining

Data mining adalah proses pengelolaan data yang menggunakan algoritma yang konsisten dan mengikuti alur data mining. Pada langkah ini, data diolah dengan menerapkan algoritma regresi linier. Algoritma regresi linier memungkinkan proses data mining untuk mengidentifikasi pola dan hubungan linier dalam data, sehingga menghasilkan estimasi harga mobil bekas toyota yaris yang lebih akurat terkait fenomena yang sedang dipelajari.

3.5. Evaluation

Pada langkah *evaluation* data, mengevaluasi hasil model estimasi harga mobil bekas Toyota Yaris menggunakan algoritma regresi linier. Mengevaluasi hasil estimasi harga mobil bekas toyota yaris berdasarkan spesifikasi dengan mengukur nilai skor MSE, RMSE, dan R2. Hasil evaluasi memberikan gambaran menyeluruh tentang seberapa baik model regresi linier dalam estimasi harga mobil bekas toyota yaris dengan tingkat akurasi dan reliabilitas tertentu.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Pada tahap *selection* data mengambil data dari website *open public kaggle.com*. Data yang diambil adalah dataset yang berisi informasi mengenai daftar mobil bekas toyota. Dalam *selection* data menggunakan microsoft excel, dengan fokus memilih dataset yang berisi daftar mobil bekas toyota dengan model yaris. Pada proses *selection* data dilakukan pemilihan atribut atau variabel yang relevan untuk membuat model regresi linier dalam proses estimasi harga. Atribut yang di *selection* adalah model karena memiliki nilai variabel yang sama. Proses *selection* data menghasilkan data mengenai daftar mobil bekas toyota yaris dengan atribut yang sudah diseleksi. Data

daftar mobil bekas toyota yaris berisi data dengan jumlah baris ada 2111 data dan 9 kolom atau atribut.

Tabel 1. Daftar mobil toyota yaris

No	Year	Price	Transmission	Mileage	Fueltype	Tax	Mpg	Enginesize
1	2001	1995	Automatic	89000	Petrol	160	42.9	1.3
2	2002	2695	Automatic	21000	Petrol	160	42.9	1.3
3	2004	950	Manual	104000	Petrol	200	40.4	1.5
4	2004	2695	Automatic	21000	Petrol	200	42.2	1.3
5	2004	1999	Automatic	71788	Petrol	165	42.2	1.3
6	2005	975	Manual	107000	Petrol	145	50.4	1
7	2006	2495	Automatic	100000	Petrol	160	44.8	1.3
8	2007	2495	Manual	87263	Petrol	145	50.4	1.3
9	2007	2699	Manual	64373	Petrol	145	50.4	1
10	2007	2795	Manual	48830	Petrol	145	50.4	1.3
11	2008	2395	Manual	99000	Petrol	160	47.1	1.3
12	2008	2750	Automatic	94000	Petrol	145	48.7	1.3
13	2008	2795	Manual	86000	Petrol	155	47.1	1.3
14	2008	2750	Manual	48000	Petrol	165	47.1	1.3
15	2008	2490	Manual	62000	Diesel	30	62.8	1.4
...
2102	2022	17995	Automatic	1127	Petrol	145	47.9	1.5
2103	2022	19276	Automatic	3113	Petrol	145	47.9	1.5
2104	2022	14995	Manual	1985	Petrol	145	47.9	1.5
2105	2022	12495	Manual	4895	Petrol	150	49.6	1
2106	2022	15885	Manual	1261	Petrol	145	47.9	1.5
2107	2022	15885	Manual	1192	Petrol	145	47.9	1.5
2108	2022	15885	Manual	2114	Petrol	145	47.9	1.5
2109	2022	15885	Manual	1113	Petrol	145	47.9	1.5
2110	2022	13999	Manual	1120	Petrol	145	47.9	1.5
2111	2022	16691	Manual	1230	Petrol	145	47.9	1.5

4.2. Data Preprocessing

Pada langkah *pre-processing* data, melakukan seleksi data ulang yang digunakan untuk membuat model regresi linier dalam proses estimasi harga mobil bekas toyota yaris. *Pre-processing* dilakukan untuk melihat data yang kosong atau *missing value* menggunakan bahasa pemrograman python dengan tools software google colab. Setelah melakukan proses *pre-processing* pada data daftar mobil bekas toyota yaris tidak ditemukan data yang bernilai kosong atau *missing value*.

```

yarisdf.isnull().sum()
year      0
price     0
transmission 0
mileage   0
fuelType  0
tax       0
mpg       0
engineSize 0
dtype: int64
    
```

Gambar 2. Pre-processing data

4.3. Data Transformation

Pada langkah *transformation* data daftar mobil bekas toyota yaris diproses menggunakan bahasa pemrograman python dengan tools software google colab. Proses transformation data dilakukan pada atribut *transmission* (transmisi), dan *fuelType* (jenis

bahan bakar. *Transformation* nilai variabel yang berupa object atau nominal ke nilai variabel integer atau angka. Transformation data menghasilkan nilai variabel berupa integer angka pada atribut *transmission* (transmisi) seperti 0 = *manual*, 1 = *automatic*, dan 2 = *semi-auto* dan pada atribut *fuelType* (jenis bahan bakar) seperti 0 = *petrol*, 1 = *hybrid*, dan 2 = *diesel*.

```

print(yarisdf['transmission'].value_counts())
print(yarisdf['fuelType'].value_counts())

Manual      1336
Automatic   759
Semi-Auto    16
Name: transmission, dtype: int64
Petrol      1533
Hybrid      556
Diesel       22
Name: fuelType, dtype: int64

yarisdf.replace({'transmission':{'Manual':0, 'Automatic':1, 'Semi-Auto':2},
                'fuelType':{'Petrol':0, 'Hybrid':1, 'Diesel':2}}, inplace = True)
    
```

Gambar 3. Transformation data

4.4. Data Mining

Proses data mining pada penelitian ini melibatkan estimasi harga mobi bekas toyota yaris menggunakan algoritma regresi linier. Data daftar mobil bekas toyota yaris diinterpretasikan setelah melalui tahap *selection* dan *pre-processing* untuk mencari hubungan antara variabel atribut seperti *year* (tahun), *transmission* (transmisi), *mileage* (jarak tempuh), *fuelType* (jenis bahan bakar), *tax* (pajak), *mpg* (miles per gallon), dan *engineSize* (ukuran mesin) dengan variabel atribut *price* (harga). Korelasi antar

atribut dievaluasi dan model regresi linier diinisiasi dengan menentukan nama data frame yaitu features dan target price (harga). Data dibagi menjadi data train dan data test dengan nilai data train 70% dan data test 30% untuk mengevaluasi hasil model estimasi harga. Model regresi linier dibuat dengan mencari nilai coefficient dan intercept pada data train. Hasil model regresi linier menghasilkan inputan seperti tahun inputan seperti tahun = 2022, transmisi = 0, jarak tempuh = 1113 km, jenis bahan bakar = 0, pajak = 145, konsumsi bahan bakar = 47,9 km per galon, dan ukuran mesin = 1.5, estimasi harga sebesar 220.973.121,60 dalam satuan rupiah.

```
features = ['year','transmission','mileage','fuelType','tax','mpg','engineSize']
x = yarisdf[features]
y = yarisdf['price']
```

Gambar 4. Inisiasi Model

```
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_state=0)
```

Gambar 5. Data train dan data test

```
model = LinearRegression()
model.fit(x_train,y_train)
```

```
LinearRegression()
LinearRegression()
```

Gambar 6. Model regresi linier

```
print('Coefficient: ',model.coef_)
print('Intercept: ',model.intercept_)

Coefficient: [ 5.63889523e+02 1.72886923e+03 -5.22997601e-02 1.57090900e+03
 1.62351747e+00 -3.24740799e+01 2.68149242e+03]
Intercept: -1129712.1385909405
```

Gambar 7. Nilai coefficient dan intercept

```
input_data = np.array([[2022, 0, 1113, 0, 145, 47.9, 1.5]])

# Melakukan prediksi menggunakan model
prediction = model.predict(input_data)

# Membuat DataFrame untuk menampilkan input data dan hasil prediksi
datacolumns = ['Year', 'Transmission', 'Mileage', 'Fuel Type', 'Tax', 'MPG', 'Engine Size']
inputdf = pd.DataFrame(data=input_data, columns=datacolumns)

predictiondf = pd.DataFrame(data={'Estimasi Harga Mobil Bekas Toyota Yaris': prediction})

# Menampilkan input data dalam bentuk tabel dengan garis
print("Input Data:")
print(inputdf.to_markdown(index=False))

# Menampilkan hasil prediksi dalam bentuk tabel dengan garis
print("\nEstimasi Harga Mobil Bekas Toyota Yaris:")
print(predictiondf.to_markdown(index=False))
```

Year	Transmission	Mileage	Fuel Type	Tax	MPG	Engine Size
2022	0	1113	0	145	47.9	1.5

```
Estimasi Harga Mobil Bekas Toyota Yaris:
Estimasi Harga Mobil Bekas Toyota Yaris
-----:
13116.4
```

Gambar 8. Inputan model estimasi harga

4.5. Evaluation

Pada tahap *evaluation* menyimpulkan hasil data dari proses data mining. Hasil evaluasi model regresi linier Dari proses data mining menghasilkan inputan hasil model estimasi harga mobil bekas toyota yaris berdasarkan atribut yang yang berpengaruh seperti

year (tahun), *transmission* (transmisi), *mileage* (jarak tempuh), *fuelType* (jenis bahan bakar), *tax* (pajak), *mpg* (miles per gallon), dan *engineSize* (ukuran mesin). Dalam proses data mining, hasil evaluasi dari model estimasi dengan regresi linier menunjukkan nilai *MSE* (*Mean Square Error*) sebesar 886,503, *RMSE* (*Root Mean Square Error*) 941,543 dan *R2-Score* 0,861. Hal ini menunjukkan bahwa evaluasi nilai *MSE*, *RMSE*, dan *R2-Score* dimanamenghasilkan nilainya mendekati 0 atau nilai terendah dan menunjukkan model estimasi menggunakan algoritma regresi linier dapat digunakan untuk proses estimasi harga mobil bekas Toyota Yaris berdasarkan spesifikasinya,

```
from sklearn.metrics import mean_squared_error, r2_score
y_test_pred = model.predict(x_test)

mse = mean_squared_error(y_test, y_test_pred)
rmse = np.sqrt(mse)
r2 = r2_score(y_test, y_test_pred)

print(f'Mean Squared Error (MSE): {mse}')
print(f'Root Mean Squared Error (RMSE): {rmse}')
print(f'R-squared (R2) Score: {r2}')
```

```
Mean Squared Error (MSE): 886503.2439166011
Root Mean Squared Error (RMSE): 941.5430122498924
R-squared (R2) Score: 0.8612613342359026
```

Gambar 9. Evaluasi nilai *MSE*, *RMSE*, dan *R2-Score*

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, model regresi linear terbukti efektif dalam memberikan estimasi harga mobil bekas toyota yaris berdasarkan spesifikasi atribut seperti seperti *year* (tahun), *transmission* (transmisi), *mileage* (jarak tempuh), *fuelType* (jenis bahan bakar), *tax* (pajak), *mpg* (miles per gallon), dan *engineSize* (ukuran mesin). Menggunakan data daftar mobil bekas toyota yaris dan algoritma regresi linear sebagai model, dilakukan estimasi harga mobil bekas toyota yaris dengan input spesifikasi seperti tahun = 2022, transmisi = 0, jarak tempuh = 1113 km, jenis bahan bakar = 0, pajak = 145, konsumsi bahan bakar = 47,9 km per galon, dan ukuran mesin = 1.5, estimasi harga sebesar 220.973.121,60 dalam satuan rupiah. Evaluasi model menggunakan algoritma regresi linear pada data mobil bekas toyota yaris menunjukkan nilai *MSE* sebesar 886,503, dimana yang mendekati nol atau nilai terendah berarti hasil estimasi sesuai dengan nilai sebenarnya dan model dapat digunakan untuk menentukan nilai estimasi.

Nilai *RMSE* sebesar 941,543, dimana nilai *RMSE* dengan nilai lebih rendah pasti bisa lebih akurat dibandingkan dengan model estimasi nilai *RMSE* lebih tinggi. *R2-Score* sebesar 0,861, dimana nilai *R2-Score* sebagai ukuran seberapa baik garis regresi mendekati nilai dari data model asli. Nilai-nilai ini mengindikasikan bahwa model regresi linear memberikan estimasi yang baik dan mampu menjelaskan estimasi harga mobil bekas secara memadai berdasarkan atribut spesifikasi yang

dimasukkan. Dengan demikian, model ini dapat dianggap sebagai model yang baik digunakan dalam melakukan estimasi harga mobil bekas toyota yaris. Untuk penelitian selanjutnya mengembangkan sistem aplikasi praktis bagi calon pembeli mobil bekas toyota yaris. Aplikasi ini memungkinkan pengguna untuk memasukkan spesifikasi kendaraannya dan mendapatkan penawaran harga yang lebih akurat berdasarkan model regresi linier. Untuk hasil yang lebih akurat, disarankan menjelajahi teknik pembelajaran machine learning lain seperti Random Forest, Gradient Boosting, atau Support Vector Machines (SVM). Algoritma ini dapat meningkatkan kinerja dalam memodelkan hubungan kompleks antara spesifikasi kendaraan dan harga, sehingga meningkatkan keandalan perkiraan. Penelitian selanjutnya dapat mengeksplorasi karakteristik yang mempengaruhi harga mobil bekas dengan mengumpulkan data yang lebih komprehensif atau menambahkan data dari berbagai sumber lain untuk mengembangkan model yang lebih kuat dan andal.

DAFTAR PUSTAKA

- [1] L. Lionarto, M. Tecoalu, and S. Wahyoedi, "Harga dan Kualitas Produk terhadap Keputusan Pembelian Mobil yang Dimediasi Kepuasan Konsumen," *J. Manag. Bussines*, vol. 4, no. 1, pp. 527–545, 2022, doi: 10.31539/jomb.v4i1.3709.
- [2] A. R. Saputra and A. P. Kusuma, "Sistem Penentuan Mobil Bekas Menggunakan Metode Analytical Hierarchy Process (Ahp)," *J. Mnemon.*, vol. 3, no. 2, pp. 1–6, 2020, doi: 10.36040/mnemonic.v3i2.2794.
- [3] A. A.-F. Nur Wahyudin, A. Primajaya, and A. S. Y. Irawan, "Penerapan Algoritma Regresi Linear Berganda Pada Estimasi Penjualan Mobil Astra Isuzu," *Techno.Com*, vol. 19, no. 4, pp. 364–374, 2020, doi: 10.33633/tc.v19i4.3834.
- [4] P. Susanti and K. Sussolaikah, "Penerapan Metode Regresi Linear Untuk Memprediksi Harga Jual Mobil Bekas Yaris Dan Jazz Pada Wilayah Dki Jakarta," *Netw. Eng. Res. Oper.*, vol. 7, no. 2, p. 133, 2022, doi: 10.21107/nero.v7i2.314.
- [5] N. Nafi'iyah, "Prediksi Jumlah Penjualan pada Toko Makmur Jaya Elektronik dengan Regresi Linier," *Res. Comput. Inf. Syst. Technol. Manag.*, vol. 2, no. 2, p. 47, 2019, doi: 10.25273/research.v2i02.5143.
- [6] B. Pradito and D. S. Purnia, "Komparasi Algoritma Linear Regression dan Neural Network Untuk Memprediksi Nilai Kurs Mata Uang," *EVOLUSI J. Sains dan Manaj.*, vol. 10, no. 2, pp. 64–71, 2022, doi: 10.31294/evolusi.v10i2.13284.
- [7] M. Adha, E. Utami, and H. Hanafi, "Prediksi Produksi Jagung Menggunakan Algoritma Apriori Dan Regresi Linear Berganda (Studi Kasus : Dinas Pertanian Kabupaten Dompu)," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 7, no. 3, pp. 803–820, 2022, doi: 10.29100/jipi.v7i3.3139.
- [8] M. Edi, E. Utami, and A. Yaqin, "Prediksi Harga pada Trading Forex Pair USDCHF Menggunakan Regresi Linear," *J. Manaj. Inform.*, vol. 13, no. 2, pp. 109–119, 2023, doi: 10.34010/jamika.v13i2.9826.
- [9] A. T. Nurani, A. Setiawan, and B. Susanto, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," *J. Sains dan Edukasi Sains*, vol. 6, no. 1, pp. 34–43, 2023, doi: 10.24246/juses.v6i1p34-43.
- [10] S. Bramasto, D. Khairiani, J. Raya, P. Serpong, and T. Selatan, "Prediksi Daya Output Sistem Pembangkit Listrik Tenaga Surya (PLTS) Menggunakan Regresi Linear Berganda," *J. Fakt. Exacta*, vol. 15, no. 3, pp. 1979–276, 2022.
- [11] A. N. Maharadja, I. Maulana, and B. A. Dermawan, "Penerapan Metode Regresi Linear Berganda untuk Prediksi Kerugian Negara Berdasarkan Kasus Tindak Pidana Korupsi," *J. Appl. Informatics Comput.*, vol. 5, no. 1, pp. 95–102, 2021, doi: 10.30871/jaic.v5i1.3184.