

## Prediksi Harga Mobil Bekas Menggunakan Algoritma Regresi Linear Berganda

**Dea Miftahul Huda<sup>1\*</sup>, Giftera Dwilestari<sup>2</sup>, Ade Rizki Rinaldi<sup>3</sup>, Iin<sup>4</sup>**

<sup>1</sup>Program Studi Teknik Informatika, <sup>2</sup>Program Studi Sistem Informasi, <sup>3,4</sup>Program Studi Rekayasa Perangkat Lunak, STMIK IKMI CIREBON

\*Email: deamiftah456@gmail.com

### **Abstrak**

Kurangnya informasi tentang harga mobil bekas menjadikan hambatan masyarakat dalam melakukan jual beli kendaraan tidak paham dengan harga pasar yang dijadikan acuan. Informasi ini menjadi sangat penting untuk mengetahui prediksi harga dengan kisaran variabel yang mengacu dapat perhitungkan. Bertujuan untuk mengolah prediksi menggunakan algoritma yang mampu melakukan statistik dengan teknik yang tepat untuk mencari nilai prediksi. Prediksi merupakan salah satu teknik yang sangat penting dalam pengambilan keputusan. Algoritma regresi linear merupakan teknik membangun model yang digunakan untuk prediksi nilai *dataset* yang di berikan. Pada penelitian ini digunakan algoritma regresi linear berganda untuk mendapat hasil prediksi harga mobil bekas. Dari *dataset* yang digunakan untuk membuat model prediksi dengan algoritma regresi linear bersumber dari *repositori kaggle* untuk mencari data harga mobil bekas dan kemudian hasilnya di visualisasi ke *rapiminer*. Proses prediksi menggunakan perbandingan data *testing* dan data *training* dengan rasio 0.9 data *training* dan 0.1 data *testing* pada proses membangun model dan melakukan evaluasi terhadap model yang sudah dihasilkan yaitu Hasil dari proses prediksi menggunakan algoritma regresi linear di peroleh model prediksi *Price* 1637.49. Model prediksi yang didapat akan dievaluasi penilaian dari *performance* yaitu *RMSE* dan *Relative Error*. Hasil yang diperoleh dari model tersebut, pada kategori *Price* nilai *RMSE* nya adalah 1637.49 dan nilai *Relative Error* 11.89%, Dan penerapan model regresi pada data baru dari variabel *independen* yang digunakan yaitu atribut Umur (*Age*) 24 X<sup>1</sup>, Kilometer (*KM*) 783764 X<sup>2</sup>, Tenaga kuda (*HP*) 100 X<sup>3</sup>, Transmisi (*Automatic*) 0 X<sup>4</sup>, Kapasitas mesin (*CC*) 1500 X<sup>5</sup>, Berat Kendaraan (*Weight*) 1300 X<sup>6</sup>, maka hasil prediksi yang dapat diprediksikan adalah 8438.61 dan persamaan regresi dengan 6 variabel X yang digunakan berikut hasil persamaan regresi  $Y = b_1 + b_2X_1 + b_3X_2 + b_4X_3 + b_5X_4 + b_6X_5 + b_7X_6$ .

Kata Kunci: prediksi, harga, dataset, regresi linear berganda, model regresi

### **Abstract**

*The lack of information regarding used car prices creates obstacles for people in buying and selling vehicles because they don't understand the market prices that are used as a reference. This information is very important to know price predictions with the range of variables that can be considered. The aim is to process an algorithm model that is capable of carrying out statistics using appropriate techniques to make predictions. Prediction is a very important technique in decision making. The linear regression algorithm is a model building technique used to predict the value of a given dataset. In this study, a multiple linear regression algorithm was used to predict used car prices. The dataset used to create a prediction model with a linear regression algorithm was sourced from the Kaggle repository for used car prices and then the results were visualized in Rapminer. The prediction process uses a comparison of testing data and training data with a ratio of 90 training data and 10 testing data in the process of building the model and evaluating the model that has been produced. The result of the prediction process using the linear regression algorithm is a prediction model of Price 1637.49. The prediction model will be evaluated with 2 assessment indicators, namely RMSE and Relative Error. The results obtained from this model, in the Price category, the RMSE value is 1637.49 and the Relative Error value is 11.89%. And the application of the regression model to new data from the independent variables used is the attribute Age (Age) 24 X1, Kilometers (KM), 783764 X2, Horse power (HP) 100 X3, Transmission (Automatic) 0 X4, Engine capacity (CC) 1500 regression equation  $Y = b_1 + b_2X_1 + b_3X_2 + b_4X_3 + b_5X_4 + b_6X_5 + b_7X_6$ .*

**Keywords:** prediction, price, dataset, multiple linear regression, regression mode.

## PENDAHULUAN

Perkembangan bisnis *otomotif* saat ini sangat pesat dengan harga tingkat fitur persaingan antar perusahaan yang semakin tinggi dan penggunaan teknologi yang semakin canggih. Meningkatnya jumlah produksi kendaraan terbaru dari pabrikan dengan *trend* yang terus berkembang. Kendaraan mobilitas ini sebagai peran dalam masyarakat membantu mempermudah perjalanan ekonomi. Keadaan tersebut membuat perusahaan berfikir keras untuk mempertahankan kelangsungan hidup perusahaan. Sehingga munculnya kendaraan baru ini berdampak dengan banyaknya kendaraan mobil bekas bisa dengan sebutan mobil *second*.

Dalam jual beli kendaraan bekas memiliki harga jual yang menurun secara periodik, dengan demikian harga jual mobil bekas padan tahun yang sudah berjalan dan bisa jadi lebih rendah dari harga jual dengan kota lain harga pasaran kendaraan baru jauh lebih tinggi daripada tahun yang keluar lebih awal, dengan memperhatikan kondisi kendaraan yang sama dan hal yang memiliki kecenderungan harga jual yang mempengaruhi oleh parameter seperti tahun keluar, merk harga jual dan kondisi mobil. Maka bisa bisa disimpulkan harga mobil bekas dapat diprediksi *system* ini mampu membantu penjual untuk memprediksi harga mobil bekas tiap periodik.

Penelitian yang dilakukan sebelumnya dari (Nur Wahyudin, 2020) dalam penelitian yang digunakan 2 variabel independen yaitu cabang (X1) dan *type* (X2) dan 1 variabel dependen yaitu Penjualan (Y). Berdasarkan perhitungan Algoritma Regresi Linear Berganda menggunakan *tools* SPSS24 ditemukan hasil nilai F Hitung sebesar 48,657 dengan nilai signifikansi 0,000 yang berarti bahwa variabel X1 dan X2 mempengaruhi dari variabel Y. Hasil penelitian ini mendapatkan estimasi penjualan sebanyak 12.223. Dan penelitian dari (Arif and Syukur, 2023) Dalam menggunakan model regresi linear pada penelitian ini berhasil mendapatkan akurasi sebesar 0,76% dan untuk hasil estimasi harga mobil yang didapatkan dengan inputan tahun mobil = 2019, KM mobil = 5000, pajak mobil = 145, konsumsi BBM = 30,2, dan ukuran mesin = 2. Maka berhasil mendapatkan nilai

estimasi sebesar 21.208,505 dalam satuan *Pound* dan 393.608,6514549 dalam satuan Rupiah pada penelitian ini model regresi baik digunakan dalam prediksi harga. (Hasibuan, 2022) algoritma Regresi Linier akan digunakan ini untuk memprediksi harga mobil bekas. Proses prediksi menggunakan perbandingan data uji dan data latih pada proses pengujian model yang dibangun memiliki tingkat akurasi sebesar 76%.

Tujuan dari penelitian ini adalah untuk mengembangkan model prediksi harga mobil yang akurat dengan menggunakan metode regresi linier berganda. Dengan model ini, diharapkan dapat mengetahui variabel X yang mempengaruhi harga mobil dan memprediksi harga mobil berdasarkan karakteristik dan kondisi pasar yang relevan supaya dapat menemukan hasil prediksi yang bisa ditemukan dalam penelitian ini. Perbedaan Penelitian ini dan penelitian sebelumnya dari menggunakan algoritma Regresi Linier untuk memprediksi harga mobil bekas dan mengimplementasikannya pada website. Sementara itu, prediksi harga mobil bekas menggunakan *RapidMiner* melibatkan penggunaan platform *RapidMiner* untuk melakukan analisis data dan membangun model prediksi harga mobil bekas dengan berbagai algoritma pembelajaran mesin yang tersedia dalam platform tersebut .

## TINJAUAN PUSTAKA

### 2.1 Prediksi

Prediksi adalah tentang sesuatu yang akan terjadi di masa depan. Prediksi adalah salah satu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil (Adiguno, Syahra and Yetri, 2022).

### 2.2 Harga

Harga adalah sebuah alat tukar atau uang yang digunakan untuk ditukar oleh konsumen kepada produsen atas produknya (Kemal, 2021). Harga merupakan jumlah uang yang wajib dibayar oleh pembeli untuk mendapatkan suatu produk/jasa. Karena

faktanya sebagian besar keuntungan yang diterima perusahaan dari penjualan produknya, baik itu barang atau jasa, harga dianggap sebagai satu-satunya prediktor terbesar keberhasilansuatu usaha (Susanti and Sussolaikah, 2022)

### 2.3 Data mining

Data mining ialah proses pengumpulan serta pengelolahan informasi yang bertujuan untuk mengekstrak data penting pada data (Karlina, 2023). DataMining adalah langkah analisis terhadap proses penemuan pengetahuan di dalam basis data atau *knowledge discovery in database (KDD)*. Dalam Data Miningterdapat beberapa teknik salah satunya adalah estimasi/prediksi (Rizky, Dkk, 2019). Datamining sering dikatakan pula *Knowledge Discovery in Databases (KDD)*, yang menggambarkan semua proses mengubah sekumpulan data biasa menjadi pengetahuan yang bermanfaat (Akmal, Faqih and Dikananda, 2023).

### 2.4 Regresi linier Berganda

Metode Regresi Linear Berganda adalah metode peramalan yang menggunakan lebih dari dua faktor yang dapat mempengaruhi hasil sehingga dapat menemukan hasil yang maksimal (Triyanto Ervan, 2019). Regresi linier berganda adalah sebagai kajian terhadap ketergantungan satu variabel dengan variabel lain dengan tujuan membuat estimasi rata-rata nilai variabel yang sudah diketahui. Regresi linier berganda Regresi linier berganda adalah hubungan secara linier antara dua atau lebih variabel *independen* ( $x_1, x_2, x_3, \dots, x_n$ ) dengan variabel *dependent* ( $y$ ) (Rizky dkk., 2019). Bentuk rumus umum dalam regresi linier berganda dapat dinyatakan secara statistik sebagai berikut:

$$Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

Keterangan:

$Y$  = Variabel Terikat

$X$  = Variabel Bebas

$a, b_1, b_2, \dots, b_n$  = Parameter Regresi/  
Koefisien Regresi  
Kelebihan

Berikut persamaan regresi linier berganda:

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n \quad (\text{Ginanto, 2021})$$

Persamaan umum metode regresi linear ini adalah:

$$Y = a + b(X)$$

Keterangan:

$a$  = Konstanta

$b$  = Koefisien regresi

$Y$  = Variabel *dependen* (variabel tak bebas)

$X$  = Variabel *independen* (variabel bebas) (Ginting, Buulolo and Siagian, 2019).

### 2.5 Evaluasi

*RMSE (Root Mean Square Error)* merupakan satu dari berbagai cara untuk mengevaluasi model regresi dengan mengukur tingkat akurasi hasil perkiraan *RMSE (Root Mean Square Error)*. Terdapat beberapa perhitungan dalam menghitung *eror* atau kesalahan. Salah satunya yang disebut *RMSE (Root Mean Square Error)*. *Root Mean Squared Error (RMSE)*, merupakan nilai rata-rata dari jumlah kuadrat kesalahan, juga dapat menyatakan ukuran besarnya kesalahan yang dihasilkan oleh suatu model prediksi. *RMSE* diperoleh dari (Bramasto and Khairiani, 2022) dengan perhitungan berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$$

Keterangan:

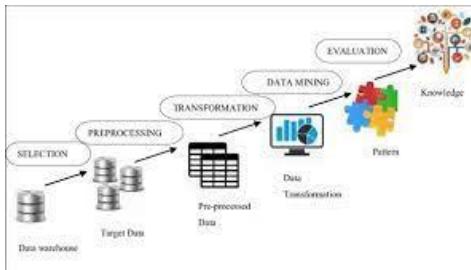
$N$  = Jumlah data

$At$  = Nilai Aktual

$Ft$  = Nilai Prediksi.

## METODE PENELITIAN

Pada metodologi ini memaparkan terkait tahap-tahap yang digunakan untuk prediksi jumlah peminat mata kuliah pilihan dengan regresi linier. Untuk tahapannya, meliputi 5 langkah yang harus dilakukan, Gambar 1. berikut merupakan langkah-langkah mengenai metodologi yang akan digunakan. Tahapan Metode Penelitian:



Gambar 1 Metode *knowledge discovery in databases* (KDD)

Berikut merupakan hal – hal yang perlu dilakukan dalam penelitian berdasarkan tahapan *knowledge discovery in databases*:

- 1) Data *Selection* adalah salah satu proses dimana pengambilan informasi/ pengumpulan informasi, pemilihan informasi serta pengecekan informasi.
- 2) Data *Selection & Data Cleaning* merupakan proses Pembersihan data yang bertujuan untuk menghilangkan data yang tidak memiliki nilai (*null*). Pembersihan data juga akan mempengaruhi performasi dari *system* data mining karena data yang akan ditangani akan berkurang jumlah dan kompleksitasnya.
- 3) Data *transformation* bertujuan untuk memudahkan koordinasi data yang diproses oleh algoritma dan alat yang digunakan dalam penelitian.
- 4) Data *Mining* adalah proses pengolahan data berdasarkan algoritma yang konsisten dengan alur *data mining*.
- 5) *Evaluation* Pada tahap evaluasi digunakan untuk mengevaluasi hasil dari algoritma Regresi linear untuk mengetahui kesimpulan dari proses data mining. Nilai *performance* yang digunakan adalah *RMSE* (*Root Mean Squared Error*).

## HASIL DAN PEMBAHASAN

### 4.1 Hasil Penelitian

Pada penelitian ini, *dataset* yang digunakan tentang Harga mobil bekas dengan beberapa atribut di data sebanyak 1435 data dari *update* bulan ini menggunakan 1 atribut Y menjadi label dan 6 atribut X menjadi jenis dan spesifikasi. Pengumpulan data ini diunduh dari *dataset* yang bersumber dari situs *Kaggle*. *Kaggle* merupakan suatu layanan *website*

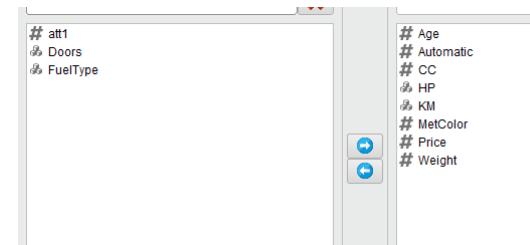
penyedia beragam *dataset* dengan format.csv berupa tabel. Adapun *dataset* sampel yang ditampilkan pada Gambar 2.

	Price	Age	KM	FuelType	HP	MetColor	Automatic	Doors	Weight
0	13500	23	46986	Diesel	90	1	0	2000	three
1	13750	23	72937	Diesel	90	1	0	2000	3
2	13950	24	41711	Diesel	90		0	2000	3
3	14950	26	48000	Diesel	90	0	0	2000	3
4	13750	30	38500	Diesel	90	0	0	2000	3
5	12950	32	61000	Diesel	90	0	0	2000	3
6	16900	27	???	Diesel	?????		0	2000	3
7	18600	30	75889		90	1	0	2000	3
8	21500	27	19700	Petrol	192	0	0	1800	3
9	12950	23	71138	Diesel	?????		0	1900	3
10	20950	25	31461	Petrol	192	0	0	1800	3
...	...	...	...	...	...	...	...	...	...
1430	8450	80	23000	Petrol	86	0	0	1300	3
1431	7500		20544	Petrol	86	1	0	1300	3
1432	10845	72	???	Petrol	86	0	0	1300	3
1433	8500		17016	Petrol	86	0	0	1300	3
1434	7250	70	???		86	1	0	1300	3
1435	6950	76	1	Petrol	110	0	0	1600	5
									1114

Gambar 2. *Dataset* Penjualan

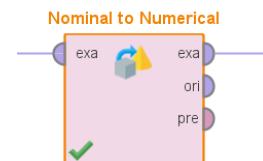
#### 4.1.1 Data Selection

Memilih atribut yang akan digunakan. Tahapan seleksi dilakukan di *rapidminer* pada proses penelitian menggunakan data penjualan mobil bekas. Tidak semua atribut digunakan, bisa dilihat pada Gambar 3.



Gambar 3. Data Selection

Mengubah tipe data *nominal* menjadi *numerik*. Atribut yang berisi data nominal diubah menjadi data *numerical*.



Gambar 4. Nominal to Numeric.

#### 4.1.2 Data Preprocessing

Dalam tahap ini dilakukan *Preprocessing* data yaitu proses *cleaning* data yang hilang atau kosong, seperti yang ditunjukkan Gambar 5.

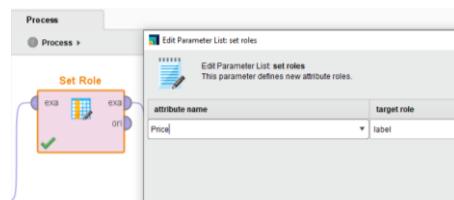
Name	Type	Missing	Statistics	Filter (7 / 7 attributes)
KM	Numeric	0	Min 0 Max 1255	Average 595.425
HP	Numeric	0	Min 0 Max 12	Average 4.681
Price	Integer	0	Min 4350 Max 32500	Average 10730.825
Age	Integer	100	Min 1 Max 80	Average 55.672
MetColor	Integer	150	Min 0 Max 1	Average 0.675
CC	Integer	0	Min 1300 Max 2000	Average 1566.828
Weight	Integer	0	Min 1000 Max 1615	Average 1072.460

Gambar 5. Data Missing Value

Pada tahap *data cleaning* merupakan proses pembersihan dari data yang akan dipakai untuk penghapusan data.

#### 4.1.3 Transformation

Pada proses transformasi data ke dalam bentuk format yang sesuai untuk diproses ke *data mining*, bisa dilihat pada Gambar 6.

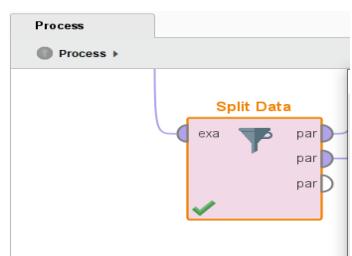


Gambar 6. Pemilihan label

#### 4.1.4 Data Mining

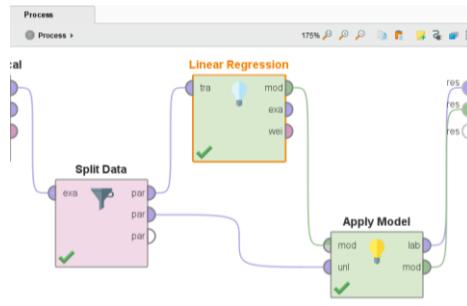
Proses data *mining* yang diterapkan dalam penelitian ini yaitu prediksi menggunakan regresi linear. Secara spesifik dan nilai evaluasi dari hasil prediksi diukur dengan *RMSE*.

Pada proses ini menggunakan aplikasi *RapidMiner*. Variabel *dependent Y* yang merupakan atribut label, dihubungkan dengan variabel *independent X*, dengan menggunakan data (*training*) model regresi linier berganda sedangkan sisanya digunakan untuk menguji data (*testing*) kinerja model yang telah dilatih dapat dilihat pada Gambar 7.



Gambar 7. Operator Split Data

Setelah melakukan pembagian data *training* dan data *testing* maka berikutnya adalah pemodelan. model yang digunakan tersebut dapat dilihat pada Gambar 8.



Gambar 8. Proses Pemodelan

Kemudian dari enam variabel *independent* adalah variabel yang berpengaruh terhadap hasil model regresi yang telah diproses.

Pada Tabel 1 hasil dari *training* menghasilkan proses perhitungan nilai *interception* dan nilai untuk semua variabel *independent*.

Tabel 1. hasil Persamaan Regresi

<i>KM</i>	-0.17037498973614348
<i>HP</i>	-13.057.334.370.013.600
<i>Age</i>	-13.225.637.368.479.600
<i>Automatic</i>	5.615.811.110.340.700
<i>CC</i>	-33.809.557.931.962.300
<i>Weight</i>	2.575.793.661.712.080
<i>(Intercept)</i>	-36.464.733.271.828.500

Dari data Tabel 1 model regresi linear adalah Bentuk persamaan garis regresi berganda 6 variabel, hasil persamaan regresi dari data *testing* yang dapat dihasilkan berikut :

$$Y = b_1 + b_2X_1 + b_3X_2 + b_4X_3 + b_5X_4 + b_6X_5 + b_7X_6$$

$$Y = -19352,8 + 1,30325E-05X_1 - 98,23638627X_2 - 98,23638627X_3 - 153,583X_4 + 0,000275X_5 + 0,000582X_6$$

Pada tahap ini menentukan nilai korelasi, ini untuk menilai sejauh mana hubungan antara variabel *independent* dan variabel *dependent*, maka diperlukan perhitungan parsial. Proses pencarian nilai korelasi dilakukan pada baris ke 3 dan hasil korelasi ditampilkan pada Gambar 9

Attribut..	KM	HP	Price	Age	Automa...	CC	Weight
KM	1	0.251	-0.879	0.798	0.017	-0.142	-0.372
HP	0.251	1	-0.334	0.302	0.186	-0.247	-0.206
Price	-0.879	-0.334	1	-0.845	0.033	0.165	0.581
Age	0.798	0.302	-0.845	1	0.031	-0.116	-0.442
Automatic	0.017	0.186	0.033	0.031	1	-0.069	0.057
CC	-0.142	-0.247	0.165	-0.116	-0.069	1	0.651
Weight	-0.372	-0.206	0.581	-0.442	0.057	0.651	1

Gambar 9. Hasil Korelasi Matriks

Hasil data pada Gambar 9 menunjukkan korelasi setiap variabel. Variabel *weight*, variabel *CC*, dan variabel *automatic* memiliki nilai tingkat akurasi yang tinggi terhadap variabel *price dependent* dengan nilai (0.581), (0.165) dan (0.033). variabel *KM*, *HP* dan *Age* memiliki nilai akurasi yang rendah. Dapat dilihat pada tingkat akurasi, maka variabel yang digunakan pada penerapan algoritma adalah *Weight*, *CC* dan *Automatic*.

#### 4.1.5 Evaluasi

Pengujian *coefisien correlation* berguna untuk menilai sejauh mana variabel *independen* mempengaruhi *dependen* dengan tujuan menentukan hasil model regresi ini sejauh mana keakuratannya. Proses yang dapat ditampilkan pada Tabel 2.

Tabel 2. Evaluasi Performance

<i>RMSE</i>	1.637.494
<i>squared_error</i>	2.681.387.529
<i>squared_correlation</i>	0.796

Untuk mengetahui evaluasi dari hasil prediksi harga mobil pada gambar 4.11, maka dilakukan perhitungan menggunakan *RMSE*, nilai yang dihasilkan mendekati nilai terkecil maka hasil prediksi yang digunakan menunjukkan metode prediksi ini bisa digunakan dengan baik, hasil bisa dilihat pada Gambar 10.

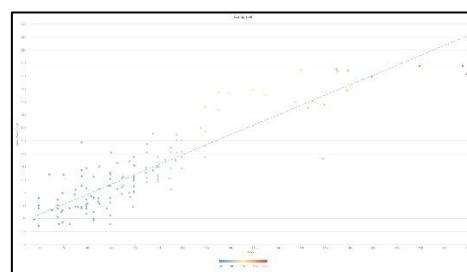
```
root_mean_squared_error
root_mean_squared_error: 1637.494 +/- 0.000

relative_error
relative_error: 11.89% +/- 11.70%
```

Gambar 10. Hasil Nilai *RMSE* dan *Relative Error*

#### 4.1.6 Knowledge

Hasil yang telah dilakukan dalam memprediksi harga mobil bekas menggunakan algoritma regresi linear berganda dengan tahapan proses *KDD*. Pada kategori prediksi harga dapat ditampilkan *plot scatter bubble* hasil prediksi dari model regresi linear. Hasil dari prediksi memiliki banyak nilai harga yang bervariasi, seperti yang ditunjukkan pada Gambar 11.

Gambar 11. *Plot Prediksi Harga*

#### 4.2 Pembahasan

Hasil nilai evaluasi yang di dapat dari perhitungan prediksi menggunakan *RMSE root\_mean\_squared\_error*: 1637.49. Pada tahap hasil *performance* dengan menggunakan evaluasi *root mean squared error* menunjukkan tingkat *error* sebesar 1637.494.

Hasil dari proses prediksi menggunakan algoritma regresi linear di peroleh model prediksi *Price* 1637.49. Model prediksi akan dievaluasi dengan 2 indikator penilaian yaitu *RMSE* dan *Relative Error*. Hasil yang diperoleh dari model tersebut, pada kategori *Price* nilai *RMSE* nya adalah 1637.49 dan nilai *Relative Error* 11.89%. Pendekatan mengukur kesalahan dengan menitng akar kuadrat dari selisih antara nilai sebenarnya dan nilai yang di prediksi, rumus yang dapat digunakan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y^i - \hat{Y}^i)^2}{n}}$$

Keterangan:

- N = Jumlah data
- At = Nilai Aktual
- Ft = Nilai Prediksi

Berdasarkan penelitian yang dilakukan sebelumnya oleh (Hasibuan dkk., 2022) tentang

prediksi harga mobil, semakin rendah nilainya maka semakin baik prediksi yang dilakukan. *RMSE* akan digunakan untuk membandingkan selisih harga *actual* dengan harga prediksi. Nilai *RMSE* adalah 3813.65 dan menurut hasil penelitian (Hamdanah & Fitrianah, 2021) nilai *RMSE* yang sangat baik adalah yang mendekati 0, lalu hasil evaluasi *Relative Error*. Dan berdasarkan (M. Sholeh dkk., 2022) penelitian yang dilakukan oleh jika *Relative Error*  $\leq 25\%$  dianggap akurat, sedangkan nilai  $> 25\%$  tetapi  $\leq 50\%$  perlu dipertimbangkan jika ingin digunakan, dan estimasi dengan nilai  $> 50\%$  dianggap tidak akurat. Berdasarkan penelitian terdahulu hasil *RMSE* dan *Relative Error* prediksi biaya asuransi termasuk model cukup, artinya adalah model yang dapat digunakan untuk memprediksi harga di bulan berikutnya.

## SIMPULAN

Berdasarkan hasil dari penelitian yang dilakukan, didapatkan kesimpulan dari penerapan *Data Mining* untuk prediksi harga mobil bekas. Ada tiga masalah utama yang dibahas dalam kesimpulan ini melibatkan proses penerapan regresi, evaluasi hasil prediksi, dan menerapkan prediksi harga menggunakan model regresi. Untuk penelitian selanjutnya dapat menambah atribut yang spesifik dan lebih banyak data dalam menentukan harga prediksi menggunakan algoritma regresi linear berganda.

## DAFTAR PUSTAKA

- Adiguno, S., Syahra, Y. And Yetri, M. (2022) 'Prediksi Peningkatan Omzet Penjualan Menggunakan Metode Regresi Linier Berganda', *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 1(4), P. 275. Available At: <https://doi.org/10.53513/jursi.v1i4.331>.
- Akmal, K., Faqih, A. And Dikananda, F. (2023) *Perbandingan Metode Algoritma Naïve Bayes Dan K-Nearest Neighbors Untuk Klasifikasi Penyakit Stroke*, *Jurnal Mahasiswa Teknik Informatika*. Available At: [www.researchgate.net](http://www.researchgate.net).
- Arif, M. And Syukur, D. (2023) 'Penerapan Model Regresi Linear Untuk Estimasi Mobil Bekas Menggunakan Bahasa Python', *Jurnal Ilmiah Matematika, Sains Dan Teknologi*, 11(2), Pp. 182–191.
- Bramasto, S. And Khairiani, D. (2022) 'Prediksi Daya Output Sistem Pembangkit Listrik Tenaga Surya (PLTS) Menggunakan Regresi Linear Berganda', *Faktor Exacta*, 15(3).
- Ginanto (2021) 'Prediksi Penjualan Kendaraan Niaga Berdasarkan Kinerja Purnajual Dan Pertumbuhan Pasar', 14(4), Pp. 214–224. Available At: <https://doi.org/10.30998/faktorexacta.v14i4.9447>.
- Ginting, F., Buulolo, E. And Siagian, E.R. (2019) 'Implementasi Algoritma Regresi Linear Sederhana Dalam Memprediksi Besaran Pendapatan Daerah (Studi Kasus: Dinas Pendapatan Kab. Deli Serdang)', *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), Pp. 274–279. Available At: <https://doi.org/10.30865/komik.v3i1.1602>.
- Hasibuan, E. (2022) 'Implementasi Machine Learning Untuk Prediksi Harga Mobil Bekas Dengan Algoritma Regresi Linear Berbasis Web', *Jurnal Ilmiah Komputasi*, 21(4), Pp. 595–602. Available At: <https://doi.org/10.32409/jikstik.21.4.3327>.
- Karlina, D. (2023) 'Estimasi Hasil Panen Ayam Pedaging Menggunakan Algoritma Regresi Linear Berganda', *Kajian Ilmiah Informatika Dan Komputer*, 3(6), Pp. 966–976. Available At: <https://doi.org/10.30865/klik.v3i6.920>.
- Kemal (2021) *Pengaruh Minat Beli Pada Produk Mobil Wuling*.
- Nur Wahyudin, D. (2020) *Penerapan Algoritma Regresi Linear Berganda Pada Estimasi Penjualan Mobil Astra Isuzu Implementation Of Double Linear Regression Algorithm On Sales Estimation Of Astra Isuzu Car*.
- Ristamaya, W., Studi Sistem Informasi, P. And Triguna Dharma, S. (2020) 'Penerapan Data Mining Untuk Memprediksi Penjualan Sepeda Motor Pada PT Mitra Pinasthika Mustika Di Periode Yang

- Akan Datang Menggunakan Metode Regresi Linier Berganda', X. No.X.
- Rizky, Dkk, . (2019) 'Implementasi Data Mining Untuk Memprediksi Target Pemakaian Stok Barang Menggunakan Metode Regresi Linier Berganda', *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, 18(2), P. 167. Available At:  
<Https://Doi.Org/10.53513/Jis.V18i2.156>.
- Susanti, P. And Sussolaikah, K. (2022) *Penerapan Metode Regresi Linear Untuk Memprediksi Harga Jual Mobil Bekas Yaris Dan Jazz Pada Wilayah Dki Jakarta Application Of Linear Regression Method To Predictable The Selling Price Of Yaris And Jazz Brand Car In The Dki Jakarta Area, Jurnal Ilmiah Nero*.
- Triyanto Ervan, D. (2019) 'Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul', *Jurnal Teknologi dan Sistem Informasi Univrab Volume 4 No. 2, 4(2)*, pp. 73–86.