

Penerapan Model Regresi Linear Untuk Estimasi Mobil Bekas Menggunakan Bahasa Python

Mohamad Arif Abdul Syukur¹, Muhammad Faisal^{2*}

^{1,2} Jurusan Teknik Informatika, UIN Maulana Malik Ibrahim, Malang, Indonesia

*Penulis Korespondensi. Email: mfaisal@ti.uin-malang.ac.id

Abstrak

Mobil bekas memiliki nilai transaksi yang signifikan dalam pasar otomotif. Estimasi harga mobil bekas menjadi penting bagi pembeli dan penjual untuk menentukan nilai yang sesuai. Dalam penelitian ini, kami menerapkan model regresi linear menggunakan bahasa pemrograman Python untuk memperkirakan harga mobil bekas berdasarkan atribut-atribut yang relevan seperti tahun produksi, capaian kilometer, pajak mobil, konsumsi bahan bakar, dan juga jumlah mesin. Kami menggunakan dataset mobil bekas yang mengandung informasi penting untuk analisis. 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. Sehingga dapat dikatakan model regresi linear terbukti berhasil dalam kategori baik untuk mencari estimasi harga mobil bekas berdasarkan faktor tertentu menggunakan bahasa Python.

Kata Kunci: Python; Regresi; Linear; Estimasi; Harga; Data

Abstract

Used cars have a significant transaction value in the automotive market. Estimating the price of a used car is important for buyers and sellers to determine the appropriate value. In this study, we apply a linear regression model using the Python programming language to estimate the price of a used car based on relevant attributes such as year of production, mileage, car tax, fuel consumption, and number of engines. We use a used car dataset that contains important information for analysis. In using the linear regression model in this study, it was successful in obtaining an accuracy of 0.76% and the results for estimating car prices were obtained by inputting car year = 2019, car mileage = 5000, car tax = 145, fuel consumption = 30,2, and engine size = 2. Then managed to get an estimated value of 21.208,505 in Pounds and 393.608,6514549 in Rupiah units. So that it can be said that the Linear Regression model has proven successful in the good category for finding used car price estimates based on certain factors using python language.

Keywords: Python; Regression; Linear; Estimation; Price; Data

1. Pendahuluan

Mobil bekas merupakan salah satu komoditas yang memiliki nilai transaksi yang cukup signifikan. Banyak faktor yang memengaruhi harga mobil bekas, seperti tahun produksi, capaian kilometer, pajak mobil, konsumsi bahan bakar, dan lain sebagainya. Untuk memperkirakan harga mobil bekas dengan akurat, diperlukan metode yang dapat menganalisis dan memodelkan hubungan antara variabel-variabel tersebut [1].

Harga mobil bekas sering mengalami perubahan karena disebabkan oleh bannyak faktor, baik itu penurunan harga maupun kenaikan harga. Khususnya pada waktu pandemi yang mana daya beli

masyarakat menurun sehingga harga mobil bekas mengalami penurunan yang cukup signifikan. Akan tetapi pasca pandemi harga mobil mengalami kenaikan sebesar 10-30 juta rupiah[2].

Penelitian yang dilakukan oleh N. Ekawati dan A. Salamena berhasil membuktikan menggunakan Fuzzy Logic Metode Sugeno bahwa harga jual mobil bekas termasuk kedalam kriteria mahal berdasarkan faktor tertentu seperti kondisi mesin bagus, pajak mobil hidup, dan interior mobil mulus[3]. Penelitian lain yang dilakukan oleh Siti Maryam, Efori dan Edizal Hatmi pada tahun 2021 yang menggunakan metode Fuzzy Mamdani dan Tsukamoto untuk mencari estimasi harga mobil bekas berdasarkan kondisi fisik mobil, pajak STNK, mobil, tahun pembuatan, aksesoris dan minat pasar berhasil mendapatkan kesimpulan bahwa metode Tsukamoto lebih baik dari hasil mamdani karena hasil tsukamoto lebih mendekati persentase yang ditentukan sesuai dalam penentuan harga mobil bekas[4]. Penelitian Y. Tuah dan A. Anyan yang berjudul Implementasi model Regresi Linear Sederhana Untuk Prediksi Gaji Berdasarkan Pengalaman Lama Bekerja berhasil menunjukkan bahwa pengalaman bekerja memberikan pengaruh positif terhadap besarnya gaji karyawan. Hasil itu disimpulkan bahwa semakin lama pengalaman bekerja seorang karyawan maka semakin berpengaruh positif terhadap gaji yang diterima. Sehingga fungsi linear regression layak digunakan untuk memprediksi pengaruh pengalaman kerja terhadap gaji karyawan[5]. Selanjutnya, pada penelitian lain Nur Nafi'iyah mencoba untuk memprediksi harga jual mobil bekas dengan membangun sistem menggunakan tool eclipse dan database SQLite yang berfungsi untuk menyimpan tabel hasil prediksi. Metode yang digunakan adalah metode regresi linear dan menghasilkan kesimpulan bahwa metode yang digunakan bermanfaat dengan baik. Hal ini disimpulkan karena hasil korelasi dari cara pertama mendapat nilai $-0,68$ dan hasil koefisien korelasi nilai kedua adalah $0,93373$. Sehingga menunjukkan bahwa metode regresi linear dapat bermanfaat dengan baik dalam memprediksi nilai jual harga mobil[6].

Berdasarkan hasil penelitian yang sudah, maka pada penelitian yang akan dilakukan penulis akan menggunakan model regresi linear untuk mencari estimasi harga mobil bekas. Dalam pencarian prediksi estimasi diperlukan menggunakan metode atau model. Salah satu model yang bisa digunakan adalah model atau metode regresi linear. Maka untuk memudahkan dalam menentukan estimasi harga mobil bekas penelitian ini akan menggunakan model regresi linear. Alasan utama menggunakan model regresi linear adalah untuk membantu menganalisis koleksi pengamatan harga berdasarkan variabel atau fitur yang ditentukan. Regresi linear adalah salah satu cara yang menggunakan garis linear atau lurus untuk menggambarkan hubungan antara variabel satu ke variabel yang lain. Sehingga dapat membantu menggambarkan hubungan antara dua variabel atau lebih[7].

Meskipun banyak penelitian lain yang menggunakan metode yang baik untuk memprediksi estimasi harga mobil bekas. Akan tetapi masih terdapat kekurangan yang diperlukan untuk meningkatkan performance model estimasi seperti identifikasi dan analisis variable input yang mempengaruhi keberhasilan nilai estimasi harga mobil bekas dalam konteks yang berbeda yang dikembangkan oleh peneliti. Oleh karena itu peneliti mencoba melanjutkan untuk menggunakan model regresi linear dalam memprediksi estimasi harga mobil bekas dengan menambahkan tampilan input aplikasi berbasis web menggunakan bahasa Python.

2. Metode Penelitian

Proses menyelesaikan permasalahan dalam memprediksi estimasi harga mobil bekas pada penelitian kali ini menggunakan metode regresi linear. Dalam regresi linear terdapat dua jenis regresi, yaitu Regresi Linear Sederhana dan Regresi Linear Berganda. Pada dasarnya kedua regresi tersebut sama-sama melibatkan variabel pemberi pengaruh. Yang membedakan adalah regresi linear sederhana hanya melibatkan satu variabel saja untuk pemberi pengaruh sedangkan regresi linear berganda melibatkan lebih dari satu variabel pemberi pengaruh[8]. Regresi linear merupakan salah satu perhitungan metode kuantitatif dimana waktu digunakan sebagai dasar prediksi estimasi. Berikut ini adalah rumus persamaan dasar dari metode regresi linear :

$$Y = a + bX \quad (1)$$

dengan :

- Y = Variabel terikat
- a = Konstanta (intercept)
- b = Koefisien regresi
- X = Variabel bebas

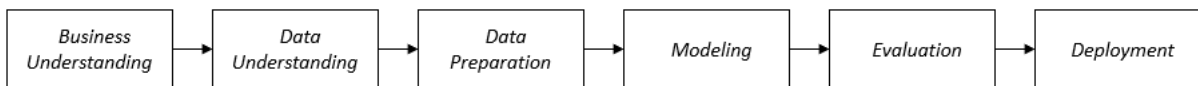
Jika untuk pencarian estimasi melibatkan banyak variabel pemberi pengaruh maka yang digunakan adalah persamaan regresi linear berganda karena untuk mengatasinya diperlukan penggunaan matematis regresi linear berganda[9]. Rumus persamaan dari regresi linear berganda adalah sebagai berikut:

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 \dots \dots + b_n X_n \quad (2)$$

dengan :

- Y = Variabel terikat
- a = Konstanta (intercept)
- $b_1 \dots n$ = Jumlah Koefisien regresi
- $X_1 \dots n$ = Jumlah Variabel bebas

Penelitian ini mengadaptasi metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang merupakan standard untuk melakukan pengembangan data mining. Metode ini dikembangkan oleh 300 organisasi hingga dipublikasikan pada tahun 1999. Pada awalnya metode ini hanya dilakukan untuk data mining yang spesifik akan tetapi standard ini juga relevan digunakan untuk lingkup lain seperti perkembangan *machine learning*[10]. Dalam pembuatan model regresi linear terdapat enam tahap yaitu *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation* dan *deployment*[11].



Gambar 1. Metode penelitian

2.1 Business Understanding

Tahap *business understanding* yaitu tahapan yang meliputi aktivitas untuk memahami kasus dari data yang digunakan. Data yang didapatkan akan dijadikan sebagai dasar dalam menentukan model yang dapat dibangun untuk memprediksi estimasi harga mobil bekas[10]. Dalam penelitian ini memerlukan data untuk digunakan dalam penentuan model. Sehingga perlu adanya dataset harga mobil bekas dan dataset yang digunakan adalah data public harga mobil bekas toyota.csv yang bersumber dari platform Kaggle (<https://www.kaggle.com/datasets/adityadesai13/used-car-dataset-ford-and-mercedes?select=toyota.csv>). Pemilihan dataset ini didasari dengan kesesuaian fitur yang terdapat pada dataset toyota.csv dengan kenyataan yang ada dipasar jual beli mobil bekas.

2.2 Data Understanding

Tahap yang kedua adalah tahap *data understanding* yaitu tahap pemahaman data yang harus dilakukan seperti mengumpulkan data awal, mendeskripsikan data, mengeksplorasi data, dan memverifikasi kualitas data[12]. Tahap ini digunakan untuk pemeriksaan data agar potensi permasalahan pada dataset yang digunakan teridentifikasi dan segera dibersihkan atau diperbaiki.

Untuk melakukan *data understanding* perlu tindakan yang teliti, cermat dan tidak tergesa-gesa karena dapat menyebabkan tidak ditemukannya kesimpulan atau kehilangan data. Jika tahap ini lambat maka akan berpengaruh kepada tahap-tahap yang selanjutnya, baik itu tahap *preparation* ataupun *modeling*. Maka dari itu pada tahap ini harus dilakukan dengan jeli untuk mengamati visualisasi data dan segera menemukan data yang akan dijadikan kesimpulan.

2.3 Data Preparation

Tahap data *preparation* yaitu penerapan transformasi terhadap dataset sehingga seluruh input yang diberikan pada model adalah data bertipe numerik sehingga dapat dieksekusi oleh sistem dengan lebih mudah. Maka data yang bukan tipe numerik tidak dapat dieksekusi oleh sistem dan akan dihilangkan selama proses pada tahap ini. Tahap ini juga membagi data latih dan data uji dengan proporsi yang telah ditentukan[10]. Dengan demikian tahap *preparation* melakukan pengecekan data kembali untuk kebenaran yang terdeteksi benar-benar baik. Sehingga tahap ini adalah tahap yang paling menguras waktu karena untuk menemukan model yang akurasiya baik berawal dari proses tahap *preparation* yang baik dan teruji berulang-ulang. Tahap ini juga memberlakukan data *missing* dan juga inkonsistensi.

2.4 Modeling

Tahap *modeling* adalah tahap yang mengaplikasikan algoritma atau metode yang akan dibangun untuk mencari, mengidentifikasi, dan menampilkan pola. Pemilihan metode berdasarkan tipe data karena dari tipe data kita bisa melihat apakah data tersebut cocok diestimasi, prediksi, klasifikasi, clustering atau melihat hubungan asosiatif[13]. Pada penelitian ini metode yang akan digunakan adalah metode regresi linear karena kita akan mencari estimasi harga mobil bekas pada dataset yang digunakan. Pada model regresi terdapat nilai positif dan negatif terhadap jumlah estimasi maka terdapat pengaruh yang signifikan berdasarkan faktor-faktor tertentu. Sehingga pada tahap ini perlu menggunakan machine learning atau metode statistika untuk model yang akan diterapkan. Jika sudah menemukan metode maka selanjutnya adalah menerapkan metode tersebut dengan menggunakan dataset yang sudah ditentukan. Jika kemudian terdapat penyesuaian kembali terhadap data maka tahapan dapat kembali ke tahap data *preparation*.

2.5 Evaluation

Tahap *evaluation* adalah tahap yang digunakan untuk membantu pengukuran tingkat presisi pada model. Tahap ini berfungsi sebagai alat validasi terhadap model yang dikembangkan. Keakuratan model dapat ditingkatkan melalui siklus CRISP-DM[14]. Dan jika sudah terbukti baik dan tepat maka akan dilanjutkan untuk tahap yang selanjutnya dan model yang sudah dibuat sudah bisa digunakan. Regresi linear sederhana merupakan analisis regresi yang melibatkan hubungan antara satu variabel bebas yang dihubungkan dengan satu variabel tak bebas yang mana juga merupakan metode statistik untuk menguji sejauh mana sebab akibat variabel penyebab terhadap variable akibatnya.

2.6 Deployment

Tahap *Deployment* adalah tahap yang merancang model agar dapat dimanfaatkan dengan kebutuhan estimasi secara waktu nyata[10]. Berdasarkan penelitain yang akan dibuat maka model yang sudah dibuat akan dijadikan model yang bertipe .sav agar bisa dieksekusi oleh bahasa Python. Setelah terbuat file type .sav maka buatlah file baru bertipe .py dengan bahasa Python untuk membuat source code tampilan aplikasi estimasi harga mobil bekas. Jika sudah maka file .sav dipanggil kedalam source code .py untuk dieksekusi. Tahapan penelitian ini berdasarkan pada model CRISP-DM dan langkah yang dilakukan untuk penelitain ini dimulai dari pengumpulan data sampai akhirnya

menghasilkan tampilan web untuk mencari estimasi harga mobil bekas. Sehingga perlu mempelajari data yang akan digunakan dan memeriksa data yang kosong atau tidak ada nilainya dan juga memeriksa batasan nilai. Pada penelitian kali ini data yang digunakan untuk pembuatan model adalah data testing sehingga langsung untuk pengujian model. Setelah model sudah terbuat maka langkah selanjutnya adalah membuat source code untuk tampilan yang akan digunakan sebagai inputan serta hasil estimasi harga mobil bekas. Pembuatan tampilan ini menggunakan bahasa Python dengan melibatkan fitur-fitur dari data yang sudah diseleksi untuk dijadikan sebagai inputan dalam mencari estimasi harga mobil bekas. Terdapat beberapa fitur yang digunakan untuk inputan pencarian estimasi harga mobil bekas yang didapatkan dari fitur-fitur dari dataset yang digunakan. Untuk memudahkan penampilan tampilan yang dibuat kita juga perlu import library seperti pickle dan streamlit.

3. Hasil dan Pembahasan

3.1 Business Understanding

Dataset yang digunakan adalah data mobil toyota yang bertipe .csv. data toyota.csv merepresentasikan model, tahun produk, harga, transmission, capaian kilometer, bahan bakar, pajak, konsumsi bahan bakar dan ukuran mesin. Nilai harga yang direpresentasikan adalah nilai harga dalam satuan Pond dimana satu Pond kira-kira kurang lebih adalah 19.000 Rupiah. Jadi untuk melihat dalam satuan rupiah kita harus mengkonversikan atau mengalikan dengan 19.000. jumlah total baris dataset toyota.csv ada 6735 baris. Dan terdapat 9 kolom atau fitur.

```
In [2]: df = pd.read_csv('toyota.csv')
df
```

```
Out[2]:
```

	model	year	price	transmission	mileage	fuelType	tax	mpg	engineSize
0	GT86	2016	16000	Manual	24089	Petrol	265	36.2	2.0
1	GT86	2017	15995	Manual	18615	Petrol	145	36.2	2.0
2	GT86	2015	13998	Manual	27469	Petrol	265	36.2	2.0
3	GT86	2017	18998	Manual	14736	Petrol	150	36.2	2.0
4	GT86	2017	17498	Manual	36284	Petrol	145	36.2	2.0
...
6733	IQ	2011	5500	Automatic	30000	Petrol	20	58.9	1.0
6734	Urban Cruiser	2011	4985	Manual	36154	Petrol	125	50.4	1.3
6735	Urban Cruiser	2012	4995	Manual	46000	Diesel	125	57.6	1.4
6736	Urban Cruiser	2011	3995	Manual	60700	Petrol	125	50.4	1.3
6737	Urban Cruiser	2011	4495	Manual	45128	Petrol	125	50.4	1.3

6738 rows × 9 columns

Gambar 2. Dataset

3.2 Data Understanding

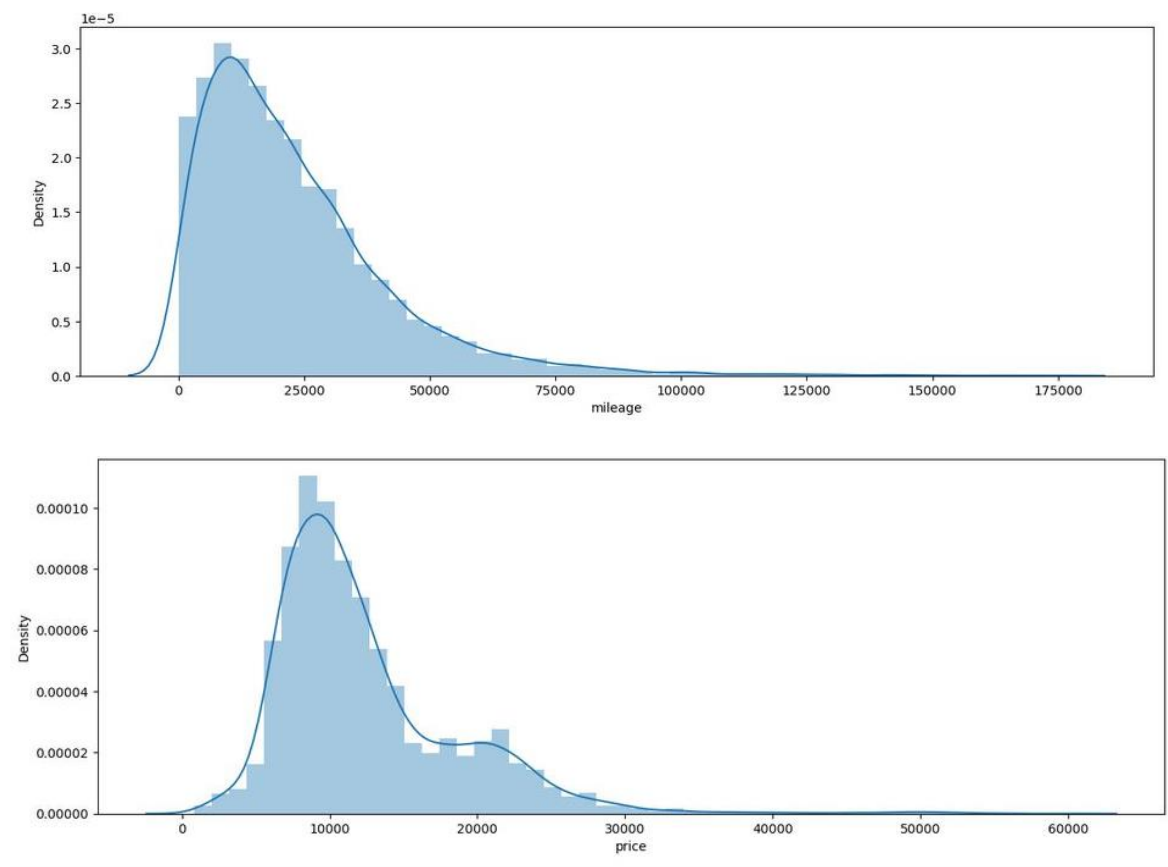
Pada tahap ini adalah analisis data statistik deskriptif data dan juga analisis distribusi data. Berdasarkan analisis yang telah dilakukan diperoleh nilai-nilai statistik deskriptif seperti jumlah data, rata-rata, standart deviasi, kuartil serta nilai minimum dan maksimum pada setiap fitur numerikal. Sedangkan untuk analisis distribusi data dilakukan pada fitur yang secara statistika bertipe data kontinu. Fitur-fitur tersebut adalah “price” dan “mileage”. Analisis distribusi ini dilakukan dengan cara memvisualisasikan setiap fitur, dimana sumbu x merepresentasikan nilai fitur itu sendiri sedangkan sumbu y merepresentasikan nilai densitinya.

```
In [4]: df.describe()
```

```
Out[4]:
```

	year	price	mileage	tax	mpg	engineSize
count	6738.000000	6738.000000	6738.000000	6738.000000	6738.000000	6738.000000
mean	2016.748145	12522.391066	22857.413921	94.697240	63.042223	1.471297
std	2.204062	6345.017587	19125.464147	73.880776	15.836710	0.436159
min	1998.000000	850.000000	2.000000	0.000000	2.800000	0.000000
25%	2016.000000	8290.000000	9446.000000	0.000000	55.400000	1.000000
50%	2017.000000	10795.000000	18513.000000	135.000000	62.800000	1.500000
75%	2018.000000	14995.000000	31063.750000	145.000000	69.000000	1.800000
max	2020.000000	59995.000000	174419.000000	565.000000	235.000000	4.500000

Gambar 3. Analisis statistik deskriptif



Gambar 4. Distribusi data kontinu

3.3 Data Preparation

Untuk membuat model dengan keakuratan yang baik dan melihat tingkat kesalahan yang dihasilkan maka perlu disiapkan data yang sudah diproses untuk menjadi data training dan data testing. Data training adalah data untuk melatih algoritma model dan data testing adalah data untuk melakukan testing model yang sudah dilatih. Pada penelitian kali ini dari jumlah data 6738 maka didapatkan data training 5390 dan data testing 1348 dengan perbandingan data training 80% dan data testing 20%.

Tabel 1. Perbandingan data *training* dan *testing*

No.	Data	Jumlah	Persen
1.	<i>Training</i>	5390	80%
2.	<i>Testing</i>	1348	20%

Aktivitas selanjutnya pada tahap ini adalah membuang fitur-fitur yang tidak digunakan untuk mencari estimasi harga mobil bekas pada data. Fitur yang dibuang atau dihilangkan adalah “model”, “transmission”, “fuelType”, dan juga “price” karena fitur “price” akan kita cari estimasinya menggunakan model yang akan dibuat yaitu model regresi linear. Tahap preparation juga menyeleksi fitur yang akan dibuat sebagai inputan untuk mencari estimasi harga mobil bekas. Untuk fitur-fitur yang diseleksi adalah “year”, “milleage”, “tax”, “mpg”, dan juga “engineSize”.

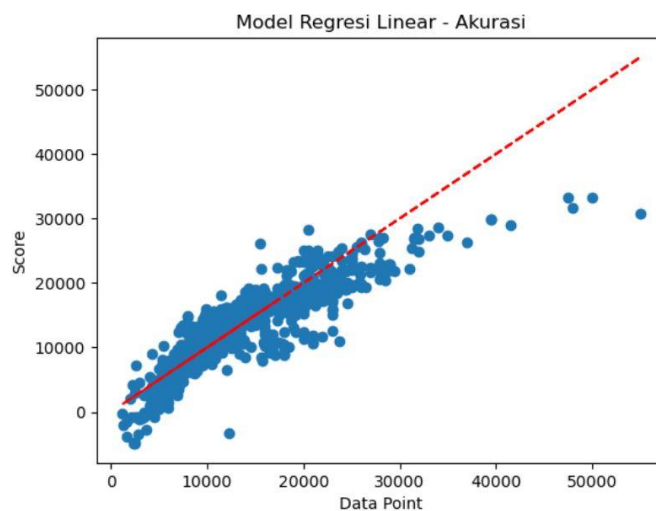
Tabel 2. Pemilihan Fitur

No.	Fitur	Type	Keterangan
1.	<i>model</i>	Object	Tidak terpilih
2.	<i>year</i>	Numerik	Terpilih
3.	<i>price</i>	Numerik	Tidak terpilih
4.	<i>transmission</i>	Object	Tidak terpilih
5.	<i>milleage</i>	Numerik	Terpilih
6.	<i>fuelType</i>	Object	Tidak terpilih
7.	<i>tax</i>	Numerik	Terpilih
8.	<i>mpg</i>	Numerik	Terpilih
9.	<i>engineSize</i>	Numerik	Terpilih

3.4 Modeling

Dataset toyota.csv memiliki fitur target yang bertipe numerikal sehingga pemodelan yang relevan digunakan adalah pemodelan regresi. Sehingga kita akan menggunakan algoritma regresi linear. Analisis regresi merupakan perhitungan statistik untuk menguji seberapa erat hubungan antar variabel. Dalam analisis regresi terdapat satu variabel yang terikat yang biasa ditulis dengan simbol Y dan satu atau lebih variabel bebas yang biasa ditulis dengan simbol X. Hubungan kedua antara variabel tersebut memiliki sifat linear sesuai dengan namanya [15].

Akurasi Model Regresi Linear = 0.7658867548233914



Gambar 5. Akurasi Model

3.5 Evaluation

Setelah model persamaan regresi linear didapatkan maka akan lebih mudah untuk mengestimasi harga mobil bekas. Dengan memasukan nilai-nilai yang menjadi faktor kemunculan estimasi harga seperti yang telah dilakukan seleksi fitur, maka nilai fitur itulah yang akan kita masukan untuk mendapatkan estimasi harga mobil bekas itu sendiri. Nilai yang perlu dimasukan adalah nilai tahun mobil, capaian kilometer mobil, pajak mobil, konsumsi bahan bakar mobil dan juga ukuran mesin mobil. Stelah nilai-nilai sudah kita masukan maka akan keluar hasil estimasi harga mobil bekas dalam satuan pond. Untuk merubah kepada satuan rupiah kita bisa mengalikannya dengan 19.000 karena satu pond adalah kisaran 19.000 rupiah. Dengan demikian model regresi linear terbukti menyelesaikan permasalahan untuk mencari hasil estimasi harga mobil bekas dengan akurasi 0.76 %.

3.6 Deployment

Setelah model terbuat dan terbukti akurasinya maka langkah selanjutnya adalah membuat model regresi linear ini kedalam bentuk file .sav agar bisa dieksekusi menggunakan bahasa Python. Untuk membuat file .sav kita perlu import library pickle. Setelah terbuat file .sav maka kita buat file baru dengan format file .py untuk menuliskan source code tampilan inputan pada web. Untuk menjalankan tampilan kita perlu install library streamlit terlebih dahulu di comand prompt. Setelah semuanya sudah dibuat maka kita bisa langsung mencobanya mengisikan nilai-nilai inputan yang sudah disediakan.

Prediksi Estimasi Harga Mobil Toyota Bekas

Input Tahun Mobil

 - +

Input KM Mobil

 - +

Input Pajak Mobil

 - +

Input Konsumsi BBM Mobil

 - +

Input Ukuran Mesin

 - +

Estimasi Harga

Estimasi harga mobil toyota bekas dalam satuan Pound =

0
0 21,208.5054

Estimasi harga mobil toyota bekas dalam satuan Rupiah =

0
0 393,608,651.4549

Gambar 6. Tampilan Web

4. Kesimpulan

Berdasarkan penelitian yang sudah dilakukan terbukti bahwa model atau metode regresi linear dapat mencari akurasi untuk estimasi harga mobil bekas dengan berdasarkan faktor seperti tahun produksi, capaian kilometer, harga pajak, bahan bakar dan juga ukuran mesin. Dengan data toyota .csv sebagai dataset dan regresi linear menjadi modelnya menghasilkan akurasi 0,76% sehingga tergolong sudah tinggi 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 393.608.651.4549 dalam satuan Rupiah. Untuk hasil estimasi yang dihasilkan masih menggunakan satuan pond akan tetapi bisa dikonversi kedalam satuan rupiah seperti penelitian yang sudah dilakukan ini. Untuk pembuatan model juga menggunakan tahapan yang berdasarkan CRISP-DM yaitu business understanding, data understanding, data preparation, modelling, evaluation dan deployment. Sehingga alur yang dilakukan untuk pembuatan model pada penelitian ini terlaksana dengan baik dan teratur. Selain itu untuk memudahkan pengguna dalam mencari estimasi harga mobil bekas penelitian ini juga membuat tampilan aplikasinya yang berbasis web menggunakan streamlit berbahasa pemrograman Python.

Referensi

- [1] E. Dewi, S. Mulyani, F. Mulady, D. Ramadhan, A. Ariyantono, and D. Ramdani, "Estimasi Harga Jual Mobil Bekas Menggunakan Metode Regresi Linier Berganda," *e-Jurnal JUSITI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 9, no. 1, pp. 1–8, 2020, doi: 10.36774/jusiti.v9i1.649.
- [2] I. Syahputra, "Analisis Perkiraan Harga Mobil LCGC Bekas Pasca Pandemi Covid-19 dengan Fuzzy Mamdani di Kota Langsa," vol. 3, no. 2, pp. 2–7, 2022.
- [3] N. Ekawati and A. P. Salamena, "Menentukan Harga Jual Mobil Bekas Menggunakan Fuzzy Logic Metode Sugeno Determining the Selling Price of a Used Car Using the Fuzzy Logic Sugeno Method," vol. 10, no. 2, 2022.
- [4] S. Maryam, E. Bu'ulolo, and E. Hatmi, "Penerapan Metode Fuzzy Mamdani dan Fuzzy Tsukamoto Dalam Menentukan Harga Mobil Bekas," *J. Informatics, Electr. Electron. Eng.*, vol. 1, no. 1, pp. 10–14, 2021, [Online]. Available: <https://djournals.com/jieec/article/view/54%0Ahttps://djournals.com/jieec/article/download/54/164>
- [5] Y. A. E. Tuah and A. Anyan, "Implementasi Model Regresi Linear Sederhana Untuk Prediksi Gaji Berdasarkan Pengalaman Lama Bekerja," *JUTECH J. Educ. Technol.*, vol. 1, no. 2, pp. 56–70, 2020, doi: 10.31932/jutech.v1i2.1289.
- [6] N. Nafi'iyah, "Penerapan regresi linear dalam memprediksi harga jual mobil bekas," pp. 1–5, 2015.
- [7] E. D. Sri Mulyani *et al.*, "Estimasi Pertumbuhan Penduduk Di Kabupaten Tasikmalaya Menggunakan Metode Regresi Linear Berganda," *Infosys (Information Syst. J.)*, vol. 6, no. 1, p. 1, 2021, doi: 10.22303/infosys.6.1.2021.1-11.
- [8] Y. Syahra, I. Santoso, and R. Kustini, "Implementasi Data Mining Untuk Prediksi Angka Kelahiran Bayi Pada Desa Sibolangit Menggunakan Multi Regresi," *Semin. Nas. Sains Teknol. Inf.*, no. 1, pp. 687–690, 2019.
- [9] P. Susanti and K. Sussolaikah, "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," *J. Ilm. NERO*, vol. 7, no. 2, p. 2022, 2022.
- [10] A. M. M. Fattah, A. Voutama, N. Heryana, and N. Sulistiyowati, "Pengembangan Model Machine Learning Regresi sebagai Web Service untuk Prediksi Harga Pembelian Mobil dengan Metode CRISP-DM," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 5, p. 1669, 2022, doi: 10.30865/jurikom.v9i5.5021.
- [11] M. Sholeh, E. K. Nurnawati, and U. Lestari, "Penerapan Data Mining dengan Metode Regresi

Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan RapidMiner,” *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 10–21, 2023, doi: 10.14421/jiska.2023.8.1.10-21.

- [12] 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.
- [13] A. Pambudi, “Penerapan Crisp-Dm Menggunakan Mlr K-Fold Pada Data Saham Pt. Telkom Indonesia (Persero) Tbk (Tlkm) (Studi Kasus: Bursa Efek Indonesia Tahun 2015-2022),” *J. Data Min. dan Sist. Inf.*, vol. 4, no. 1, p. 1, 2023, doi: 10.33365/jdmsi.v4i1.2462.
- [14] A. K. Rotty, T. S. Dewayana, and A. N. Habyba, “Cross-Industry Standard Process for Data Mining (CRISP-DM) Approach in Determining the Most Significant Employee Engagement Drivers to Sales at X Car Dealership,” pp. 3368–3379, 2022.
- [15] N. Almutazah, N. Azizah, Y. L. Putri, and D. C. R. Novitasari, “Prediksi Jumlah Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana,” *J. Ilm. Mat. Dan Terap.*, vol. 18, no. 1, pp. 31–40, 2021, doi: 10.22487/2540766x.2021.v18.i1.15465.