

Penerapan Regresi Linier Berganda Untuk Memprediksi Harga Laptop Dengan Menggunakan Software Python

Syaiful Nur Wardani^{1*}, Nurmalitasari²

^{1,2} Sistem Informasi

Universitas Duta Bangsa Surakarta

^{1*}202030376@mhs.udb.ac.id, ²nurmalitasari@udb.ac.id

Abstrak— Pada era revolusi industri 4.0 saat ini perkembangan produk elektronik dan digital semakin maju pesat dengan berkembangnya suatu produk elektronik terutama laptop yang mempunyai berbagai macam-varian dan komponen penyusun yang sangat berkualitas membuat masyarakat ingin membelinya. Data yang digunakan dalam penelitian ini adalah data yang diperoleh dari website ELS Computer yang berjumlah 65 data. Metode yang digunakan dalam penelitian ini adalah Regresi linier. Regresi linier adalah metode yang memodelkan variabel dependen dan variabel independen dengan menganalisis hubungan antar variabel-variabel tersebut. Penelitian ini menggunakan variabel dependennya adalah harga dengan variabel independennya adalah ukuran layar, RAM Storage, dan berat laptop. Tujuan dari penelitian ini adalah memprediksi harga laptop menggunakan metode regresi linier. Prediksi hasil evaluasi linier dalam nilai MAE, MSE, RMSE, dan MAPE menghasilkan nilai keakuratan prediksi sebesar 0.34.

Kata kunci— Harga, Laptop, Regresi Linier, Prediksi, Python

Abstract— In the era of the industrial revolution 4.0, the development of electronic and digital products is increasing rapidly with the development of electronic products, especially laptops which have various variants and high-quality components that make people want to buy them. The data used in this study are data obtained from the ELS Computer website, which totals 65 data. The method used in this research is linear regression. Linear regression is a method that models the dependent variable and independent variable by analyzing the relationship between these variables. This study uses the dependent variable price and the independent variables are screen size, RAM storage and laptop weight. The purpose of this study is to predict the price of laptops using the linear regression method. The prediction of the linear evaluation results on the MAE, MSE, RMSE, and MAPE values resulted in a prediction accuracy value of 0.34.

Keywords: Price, Laptop, Linear Regression, Prediction, Python.

I. PENDAHULUAN

Pada era revolusi industri 4.0 saat ini perkembangan produk elektronik dan digital semakin maju pesat dengan berkembangnya suatu produk elektronik terutama laptop yang mempunyai berbagai macam-varian dan komponen penyusun yang sangat berkualitas membuat masyarakat ingin membelinya. Dalam hal ini produk laptop merupakan produk elektronik yang banyak diminati oleh masyarakat khususnya para pelajar dan mahasiswa. Harga laptop yang semakin terjangkau dan memiliki spesifikasi yang mumpuni menjadi alasan yang paling kuat masyarakat untuk memiliki sebuah laptop. Kegiatan yang umumnya menggunakan peranan teknologi informasi seperti, pengetikan, pembuatan surat-menyurat, pengolahan data penjualan, proses belajar mengajar dan lain-lain dapat dilakukan dengan menggunakan laptop.

Masyarakat terutama dikalangan pelajar, mahasiswa, dan pekerja masih bimbang dalam

memilih laptop dengan harga yang sesuai dengan kebutuhan mereka. Hal ini disebabkan karena konsumen yang dihadapkan pada pemilihan beberapa spesifikasi laptop yang berbeda-beda dengan keunggulan masing-masing, selain itu konsumen juga kurang dalam mendapatkan pemahaman dan informasi yang cukup mengenai kelebihan dan keunggulan dari setiap spesifikasi laptop yang akan dibelinya. Oleh karena itu penulis membuat prediksi harga laptop menggunakan analisis regresi linier.

Batasan masalah dari penelitian ini berfokus pada : menentukan kualitas prediksi laptop yang baik, Data yang diambil adalah data dari ELS Computer, Data yang diolah dalam penelitian ini menggunakan metode analisis regresi linier. Analisis regresi linier adalah suatu teknik statistik yang digunakan untuk memodelkan dan menganalisis hubungan antara satu variabel dependen (variabel terikat) dengan satu atau lebih

variabel independen (variabel bebas) [1]. menyatakan bahwa tujuan utama penggunaan regresi linier adalah untuk menghitung hubungan antara parameter model dan memperkirakannya. Metode ini di implementasikan menggunakan pemrograman python. Python adalah bahasa pemrograman yang mudah dipelajari, dipahami, dan di implementasi oleh pelajar, bahasa python tergolong dalam bahasa yang ramah, mudah, dan memiliki komunitas yang luas, sehingga bahasa pemrograman dapat dipelajari oleh siapapun [2]. Python dapat digunakan dalam google collab, sehingga pengoperasian menjadi lebih mudah dan dilakukan secara online.

Tujuan dari penelitian ini adalah Untuk menganalisis harga laptop sebagai variabel dependen dan variabel independennya adalah Ukuran Layar, RAM, Storage dan Berat. Selain itu juga untuk memprediksi harga laptop agar konsumen dapat menemukan target harga yang sesuai. Manfaat dari penelitian ini yaitu untuk membantu dan mempermudah suatu konsumen pada prediksi harga laptop yang baik dan sesuai dengan kebutuhan, selain itu juga untuk mengetahui spesifikasi laptop itu sendiri dengan menampilkan price pada laptop.

II. METODOLOGI PENELITIAN Metodologi

penelitian berisi tentang teori-teori yang berkaitan dengan analisis regresi dan metode dalam pengolahan data.

A. Pengumpulan Data

Data yang dikumpulkan oleh penulis termasuk ke dalam data primer, data primer adalah data yang dikumpulkan langsung dari sumber asli untuk keperluan penelitian tertentu dan tidak bersumber dari pihak ke dua dalam penelitian [4]. Pada penelitian ini menggunakan dataset yang diperoleh dari ELS Computer yang merupakan dataset publik. Penulis menggunakan salah satu metode data primer yaitu observasi harga dan spesifikasi laptop dari web dan katalog ELS Computer.

B. Variabel Data

Variabel yang digunakan dalam penelitian ini menggunakan dua jenis variabel yaitu variabel dependen dan variabel independen. Variabel independen yaitu variabel yang menjadi sebab terjadinya atau terpengaruhnya variabel terikat [3]. Variabel independen yang dipakai dalam penelitian ini adalah ukuran layar, RAM, storage, dan berat. Variabel dependen adalah variabel terikat yang dipengaruhi karena adanya variabel bebas. Variabel dependen dalam penelitian ini adalah harga [3].

C. Metode Analisis Data

Metode analisis data yang digunakan dalam penelitian ini adalah regresi linier berganda. Berikut langkah-langkah penerapan teknik regresi linier [8] :

Menentukan variabel independen (x) dan variabel dependen (y).

Melakukan pengelompokan data agar data yang ada tidak mengalami gangguan selama melakukan proses data yang disebabkan oleh permasalahan pada data tersebut.

Mencari dan menentukan konstanta (a) dan koefisien (b) dapat dilakukan dengan rumus sebagai berikut :

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{(n)(\sum X^2) - (\sum X)^2} \quad (1)$$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{(n)(\sum X^2) - (\sum X)^2} \quad (2) \quad [6]$$

Dengan a adalah konstanta, b adalah koefisien atau intersep, x adalah variabel independen, dan y adalah variabel dependen. Melakukan prediksi dengan regresi linier menggunakan rumus :

$$Y = a + Bx \quad (3)$$

dengan a adalah konstanta dan b adalah koefisien atau intersep. [5]

III. HASIL DAN PEMBAHASAN

A. Mempersiapkan Data

Hasil observasi disimpan dalam bentuk file excel dan dilakukan penghapusan data yang tidak digunakan. File excel disimpan ke dalam bentuk csv agar bisa terbaca dengan mudah di dalam *software python*. Dataset memiliki total 65 data.

1	Merk Laptop	Ukuran Layar (Inch)	RAM (GB)	Storage (GB)	Berat (Kg)	price
2	Asus Vivobook K513EA	14	4	512	1.5	6999000
3	MSI Modem 14-C12M	16	8	512	2.2	6399000
4	Asus Vivobook A1405ZA	14	4	256	1.5	6599000
5	Asus Vivoobok 14X	14	8	512	1.5	9799000
6	Asus Zenbook Flip	14	8	512	1.5	13999000
7	Asus Vivobook 14 A14	14	4	256	1.5	5799000
8	Asus Vivobook A1502ZA	14	8	512	1.5	11049000
9	Asus ROG Flow	16	16	512	2	19499000
10	Asus TUF Gaming	16	16	512	2	13499000
11	Asus Zephyrus	16	16	1000	2	34799000
12	Lenovo Ideapad Slim 5	14	16	512	1.5	11279000
13	Lenovo Ideapad slim 3	14	8	256	1.5	6099000
14	Lenovo Ideapad Slim 1	14	4	256	1.5	5999000
15	Lenovo Thinkpad X260	14	8	256	1.5	6999000
16	Lenovo Yoga	13	8	512	1.4	12999000
17	Lenovo Yoga	14	16	512	1.5	15399000
18	Lenovo Legion	16	16	512	2	24999000
19	Lenovo Legion	16	16	512	2	24999000
20	Lenovo Legion	16	16	512	2	24999000
21	Acer Aspire A514	14	8	512	1.5	6849000
22	Acer Aspire A314	14	8	512	1.5	7599000
23	Acer Aspire 3	14	8	512	1.5	6449000
24	Acer Predator Helios	16	16	1000	2	35999000
25	Acer Aspire 5 Gaming	16	16	512	2	12999000
26	Acer Aspire Nitro	16	16	512	2	15699000

Gambar 1. Pengumpulan data excel

B. Mengolah Data

Menampilkan data yang telah di konversi dalam bentuk file CSV dan di import ke dalam software python dan Google Colab. Untuk library yang digunakan adalah numpy, pandas, seaborn, dan matplotlib [7].

```

[4]
Merk Laptop  Ukuran Layar (Inch)  RAM (GB)  Storage (GB)  Berat (Kg)  price
0  Asus Vivobook K513EA           14         4         512         1.5  6999000
1  MSI Modem 14-C12M             16         8         512         2.2  6399000
2  Asus Vivobook A1405ZA         14         4         256         1.5  6599000
3  Asus Vivoobok 14X             14         8         512         1.5  9799000
4  Asus Zenbook Flip             14         8         512         1.5  13999000

df.describe()

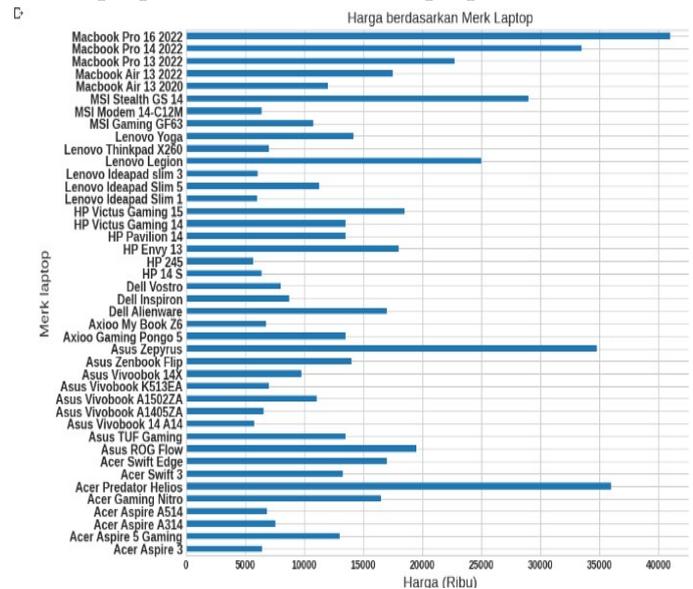
```

	Ukuran Layar (Inch)	RAM (GB)	Storage (GB)	Berat (Kg)	price
count	64.000000	64.000000	64.000000	64.000000	6.400000e+01
mean	14.437500	11.250000	502.125000	1.654688	1.427447e+07
std	1.206793	4.666667	176.873688	0.262424	8.515639e+06
min	13.000000	4.000000	256.000000	1.400000	5.699000e+06
25%	13.750000	8.000000	512.000000	1.500000	6.999000e+06
50%	14.000000	8.000000	512.000000	1.500000	1.299900e+07
75%	16.000000	16.000000	512.000000	2.000000	1.712400e+07

Gambar 2. Menampilkan data dalam python

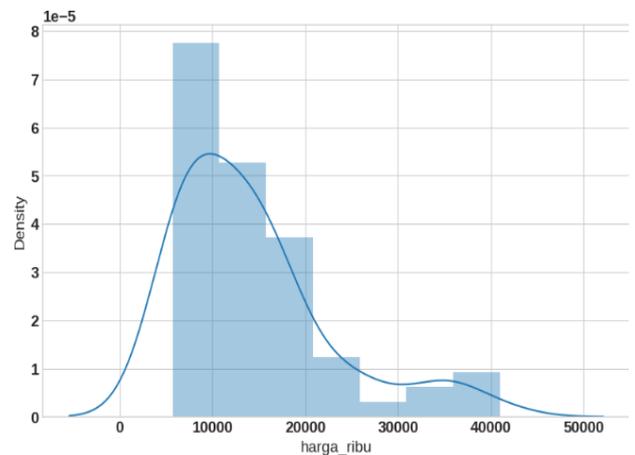
C. Visualisasi data

Berikut grafik yang menunjukkan, harga laptop berdasarkan merk laptop.

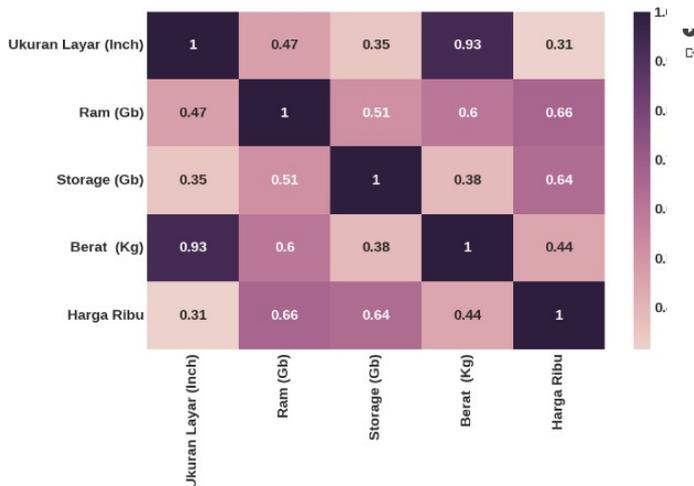


Gambar 3. Grafik harga berdasarkanMerk Laptop

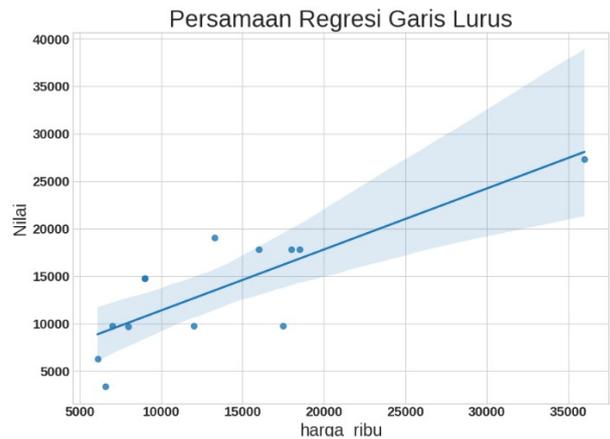
Berikut Histogram harga laptop.



Gambar 4. HistogramHarga laptop



Gambar 5. Korelasi Variabel



Gambar 8. Persamaan Garis Linier

Gambar diatas menunjukkan korelasi variabel-variabel tertentu. Variabel yang berpengaruh besar terhadap harga adalah RAM dengan angka nominal yaitu 0.66.

D. Pembuatan Model

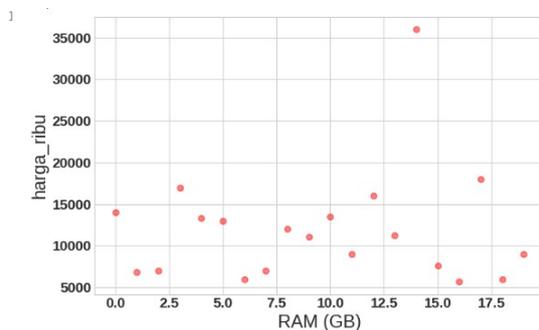
```
[79] x = df.harga.values.reshape((-1,1))
     y = df.iloc[:, 2:5]
```

```
[80] mdl.fit( x,y )
```

```
LinearRegression
LinearRegression()
```

Gambar 6. Model x dan y

Model ini digunakan untuk menentukan variabel dependen (x) dan variabel independen (y) [6].



Gambar 7. Perbandingan harga laptop dengan RAM.

E. Uji Normalitas

Uji normalitas digunakan untuk menguji sebuah data agar bisa distribusi secara normal, salah satu metode uji normalitas adalah *jarque-bera*. Nilai jarque bera yang di dapat

```
[ ] from scipy import stats
     #uji statistik dengan jarque bera
     #pengujian jarque bera
     jb_statistic, jb_p_value = stats.jarque_bera(x)

     print("Nilai p- : " , jb_p_value)
```

Nilai p- : 2.554265055862247e-06

Gambar 9. Nilai jarque bera

Nilai minimal data distribusi normal adalah 0,05 maka hasil jarque bera diatas, termasuk dalam pendistribusian data yang normal.

F. Uji Statistik

Uji statistik Shapiro-Wilk didasarkan pada hipotesis nol bahwa data diambil dari populasi yang terdistribusi secara normal. Metode ini menghasilkan statistik uji W dan nilai p-value yang menunjukkan seberapa baik hipotesis nol didukung oleh jumlah data yang dikumpulkan.

```
ShapiroResult(statistic=0.84095698595047, pvalue=8.816032845970767e-07)
ShapiroResult(statistic=0.7743586897850037, pvalue=1.5511805173673565e-08)
ShapiroResult(statistic=0.7484557628631592, pvalue=3.947782545310474e-09)
ShapiroResult(statistic=0.6395620107650757, pvalue=2.8983870858922955e-11)
```

Gambar 9. Uji shapiro

G. Evaluasi Linier

Evaluasi linier adalah proses mengukur dan menganalisis kualitas dan kinerja model regresi linier dalam memprediksi nilai variabel target berdasarkan variabel prediktor yang ada. Evaluasi ini bertujuan untuk mengevaluasi sejauh mana model linier dapat memberikan prediksi yang akurat dan relevan terhadap data yang ada [9]

- a) Nilai MAE mengukur rata-rata dari selisih absolut antara nilai prediksi dan nilai sebenarnya. MAE mengabaikan arah kesalahan, artinya hanya memperhitungkan besarnya selisih.
- b) Nilai MAPE diukur mengukur persentase rata-rata dari selisih absolut antara nilai prediksi dan nilai sebenarnya.
- c) Nilai MSE diukur dari persentase rata-rata dari selisih absolut antara nilai prediksi dan nilai sebenarnya.
- d) Nilai RMSE
Nilai RMSE adalah akar kuadrat dari MSE. Nilai RMSE sering digunakan karena menghasilkan skala yang serupa dengan data asli. Nilai RMSE yang lebih kecil menunjukkan tingkat kesalahan yang lebih rendah dan kualitas model yang lebih baik. [7]

```
[27] #interpretasi
#MAE
print('Nilai MAE : ',mean_absolute_error(y_test,y_pred))

Nilai MAE : 26.84704028232797

#MAPE
print('Nilai MAPE : ',mean_absolute_percentage_error(y_test,y_pred))

Nilai MAPE : 0.2370360251638835

#MSE
print('Nilai MSE : ',mean_squared_error(y_test,y_pred))

Nilai MSE : 2977.8657750585427

#RMSE
print('Nilai RMSE: ', np.sqrt(mean_squared_error(y_test, y_pred)))

Nilai RMSE: 54.56982477394025
```

Gambar 10. Hasil Nilai MAE, MAPE, MSE, dan RMSE

H. Hasil Prediksi

Setelah menemukan nilai dari MAE,MSE,MAPE,dan RMSE maka dapat memprediksi nilai keakuratan prediksi dan dihasilkan nilai keakuratan prediksi sebagai berikut [10].

```
# know score
score = mdl.score( x, y )

print("Nilai Keakuratan Prediksi : ", score)

Nilai Keakuratan Prediksi : 0.34702395063588276
```

Gambar 14. Nilai Keakuratan Prediksi

Berdasarkan hasil tersebut maka diperoleh nilai keakuratan prediksi yaitu 0.347.

IV. KESIMPULAN

Analisis regresi linier dapat digunakan untuk memprediksi harga laptop dengan variabel dependen yaitu harga laptop dan variabel independen yaitu ukuran layar, RAM, Storage dan berat laptop. Dari jumlah 65 data diperoleh hasil korelasi yang didapat adalah bahwa variabel RAM sangat berpengaruh terhadap variabel harga karena memiliki nilai tertinggi yaitu 0.66. Adanya hasil uji normalitas dengan nilai 2.5 menunjukkan bahwa data dapat distribusi dengan normal. Selain itu, berdasarkan perhitungan evaluasi linier diperoleh nilai MAE sebesar 26.84, nilai MAPE sebesar 0.23, dan nilai RMSE sebesar 54.5 sehingga

diperoleh nilai hasil prediksi keakuratan sebesar 0.34.

UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih terhadap pihak-pihak dalam membantu penulisan jurnal, terutama pada pihak dosen yang senantiasa membantu dalam konsultasi penulisan.

REFERENSI

- [1] Gujarati, D.N. and Porter, D.C. (2009) Basic Econometrics. 5th Edition, McGraw Hill Inc., New York.
- [2] Lutz, M. (1996). Programming Python. O'Reilly Media.
- [3] Kothari, C.R. (2004) Research Methodology: Methods and Techniques. 2nd Edition, New Age International Publishers, New Delhi.
- [4] Bryman, A. (2016). Social Research Methods (5th ed.). Oxford, 373-374.
- [5] Kerlinger, F.N. (1986) Foundations of Behavioral Research. 3rd Edition, Holt, Rinehart and Winston, New York.
- [6] Joseph F. Hair, William C. Black, Barry J. Babin, Rolph E. Anderson (2013), Multivariate Data Analysis, Pearson Education Limited, London.
- [7] Guido Van Rossum (2015), The Modern Era of Python, IEEE
- [8] Edward R. Tufte (1997), Visual Explanations: Images and Quantities, Evidence and Narrative, Graphics Press
- [9] Muhammad Sholeh (2022, Penerapan Regresi Linear Ganda Untuk Memprediksi Hasil Nilai Kuesioner Mahasiswa Dengan Menggunakan Python, Jurnal Dinamika Informatika.
- [10] Yayan Adrianova Eka Tuah(2020), Implementasi Linear Sederhana Untuk Prediksi Gaji Berdasarkan Pengalaman Lama Bekerja, Jutech .