



ISSN : 2339 - 1871

## BETRIK BESEMAH TEKNOLOGI INFORMASI & KOMPUTER

Editor Office : Pusat Penelitian & Pengabdian Pada Masyarakat  
(PPPM) ITPA

Phone : 0857-9716-9578

email : [betrikitpa@itpa.ac.id](mailto:betrikitpa@itpa.ac.id)

### Implementasi Algoritma Regresi Linear Untuk Memprediksi Harga Laptop

Risky Harahap<sup>1</sup>, Karpen<sup>2</sup>, Helda Yenni<sup>3</sup>, Muhamad Jamaris<sup>4</sup>

Program Studi Teknik Informatika, Universitas Sains dan Teknologi Indonesia<sup>1,2,3,4</sup>

Jalan Purwodadi No. KM. 10, Sidomulyo Barat, Tampan, Kota Pekanbaru, Riau 28294, Indonesia

Sur-el : [harahaprisky74@gmail.com](mailto:harahaprisky74@gmail.com)<sup>1</sup>, [karpenkarpen49@gmail.com](mailto:karpenkarpen49@gmail.com)<sup>2</sup>, [helayenni@sar.ac.id](mailto:helayenni@sar.ac.id)<sup>3</sup>,  
[muhamadjamaris@gmail.com](mailto:muhamadjamaris@gmail.com)<sup>4</sup>

Penulis Korespondensi: Risky Harahap, [harahaprisky74@gmail.com](mailto:harahaprisky74@gmail.com)

**Abstrak:** Perkembangan teknologi laptop mendorong kebutuhan prediksi harga yang akurat untuk membantu konsumen dalam mengambil keputusan pembelian secara tepat dan efisien. Penelitian ini mengimplementasikan algoritma Regresi Linear untuk memprediksi harga laptop berdasarkan 4 fitur utama yang termasuk *Brand*, *Processor*, *RAM*, dan *GPU*. Dataset yang digunakan terdiri dari 11.768 data yang diperoleh dari platform Kaggle yang diproses melalui tahapan *preprocessing*, transformasi fitur, serta evaluasi model dengan berbagai metrik performa. Hasil analisis menunjukkan bahwa fitur RAM memiliki pengaruh paling signifikan terhadap harga laptop, diikuti oleh Processor, Brand, dan GPU. Model Regresi Linear yang dikembangkan berhasil mencapai nilai R-squared sebesar 0,6453, yang mengindikasikan bahwa model mampu menjelaskan 64,53% variasi harga laptop berdasarkan fitur yang dianalisis. Penelitian ini memberikan kontribusi dalam pengembangan sistem prediksi harga laptop yang akurat serta menyediakan alat bantu praktis untuk mendukung pengambilan keputusan pembelian berbasis data secara efektif dan efisien.

**Kata kunci:** Evaluasi Model, Fitur Harga Laptop, Prediksi Harga Laptop, *Preprocessing*, *Regresi Linear*.

**Abstract:** The development of laptop technology has driven the need for accurate price predictions to assist consumers in making purchasing decisions appropriately and efficiently. This study implements a Linear Regression algorithm to predict laptop prices based on 4 main features including Brand, Processor, RAM, and GPU. The dataset used consists of 11,768 data obtained from the Kaggle platform which is processed through preprocessing, feature transformation, and model evaluation stages with various performance metrics. The analysis results show that the RAM feature has the most significant influence on laptop prices, followed by Processor, Brand, and GPU. The developed Linear Regression model successfully achieved an R-squared value of 0.6453, which indicates that the model is able to explain 64.53% of the variation in laptop prices based on the analyzed features. This study contributes to the development of an accurate laptop price prediction system and provides a practical tool to support data-based purchasing decisions effectively and efficiently.

**Keywords:** Model Evaluation, Laptop Price Features, Laptop Price Prediction, Preprocessing, Linear Regression.

Received: 21-07-2025 | Accepted: 28-07-2025 | Published Online: 30-08-2025

All author: Risky Harahap, Karpen, Helda Yenni, Muhammad Jamaris

## 1. PENDAHULUAN

Laptop pertama kali diperkenalkan oleh Alan Kay pada tahun 1970-an di Xerox Palo Alto Research Center melalui konsep Dynabook, yaitu komputer portabel tanpa kabel dengan berukuran kecil seperti buku tulis. IBM memperkenalkan IBM 5100 pada September 1975, yang menjadi komputer portabel pertama yang dipasarkan secara komersial. Kemudian, pada tahun 1981, Epson mengembangkan HX-20, sebuah komputer portabel yang dilengkapi layar LCD, baterai isi ulang, dan printer mini dengan berat sekitar 1,6 kg [1]. Sejak peluncuran IBM 5100, perkembangan laptop terus berlanjut hingga kini telah berevolusi menjadi perangkat canggih yang dibutuhkan oleh berbagai kalangan dalam kehidupan sehari-hari [2].

Perkembangan teknologi yang pesat telah menjadikan laptop sebagai salah satu perangkat dalam menunjang mobilitas kerja maupun aktivitas di berbagai instansi. Seiring waktu, laptop mengalami berbagai peningkatan, baik dari segi desain, ukuran, maupun spesifikasi, sehingga mampu memberikan performa yang semakin optimal. Perubahan ini membuat laptop menjadi solusi yang efektif terhadap keterbatasan komputer desktop yang bersifat statis dan kurang fleksibel. Dengan bobot yang lebih ringan, bentuk yang ringkas, serta kemampuan yang dapat disesuaikan dengan kebutuhan pengguna, laptop semakin diminati sebagai perangkat komputasi utama [2].

Laptop juga merupakan salah satu produk elektronik yang paling diminati dan dibutuhkan oleh seluruh kalangan, terutama pelajar, mahasiswa dan pekerja. Namun, konsumen sering kali mengalami kebingungan dalam menentukan pilihan laptop yang sesuai dengan kebutuhan spesifikasi maupun anggaran yang akan mereka keluarkan. Hal ini disebabkan oleh banyaknya variasi produk laptop yang tersedia dipasaran, dimana setiap merek dan tipe menawarkan spesifikasi serta keunggulan yang beragam [3].

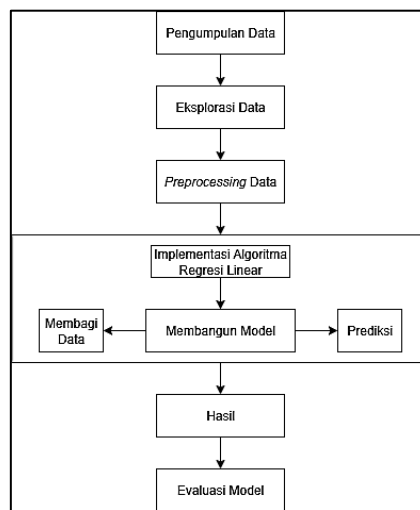
Peningkatan kebutuhan laptop di Indonesia saat ini mencerminkan pentingnya teknologi sebagai penunjang berbagai kegiatan dalam aspek, seperti sosial, pemerintahan, pendidikan, agama, dan lainnya. Ragam kegiatan ataupun aktivitas di berbagai sektor tersebut, diperlukan *work device* (laptop). Harga laptop sebagai salah satu aspek yang diperhitungkan konsumen terkadang sering mengalami perubahan yang membuat konsumen bingung untuk mempersiapkan anggaran yang akan dikeluarkan untuk membeli laptop. Masyarakat atau konsumen laptop pun akhirnya membutuhkan cara agar dapat mengantisipasi budget yang tidak sesuai dari dana yang telah diperhitungkan atau disisihkan [4].

Permasalahan di atas, kemudian dapat diatasi salah satunya dengan melakukan sebuah prediksi harga laptop. Prediksi merupakan bagian penting dalam proses perencanaan dan pengambilan keputusan dan konsumen menunjukkan sesuatu yang akan terjadi dalam keadaan tertentu [5]. Permasalahan tersebut menjadi salah satu alasan peneliti memilih algoritma Regresi Linear, karena bersifat yang sederhana, efisien secara komputasi, dan sangat cocok untuk memodelkan hubungan linear antara fitur spesifikasi laptop dan harga. Selain itu, model Regresi Linear juga mudah diimplementasikan serta dipahami oleh pengguna umum. Regresi linear merupakan salah satu metode statistik yang menghubungkan antara dua variabel atau lebih. Algoritma ini juga memiliki kelebihan dan kekurangan dalam meningkatkan akurasi yang baik dan lebih efisien [6]. Berdasarkan hal tersebut, penelitian bertujuan untuk mengimplementasikan sebuah

algoritma Regresi Linear untuk memprediksi harga laptop berdasarkan spesifikasinya. Hasil penelitian diharapkan dapat membantu masyarakat atau konsumen dalam menentukan laptop pilihan yang sesuai budget kebutuhan dan anggaran yang dimiliki.

## 2. METODOLOGI PENELITIAN

Desain penelitian merupakan gambaran umum mengenai tahapan-tahapan yang dilakukan secara sistematis dari awal hingga akhir untuk mencapai tujuan penelitian. Diagram ini menggambarkan alur kerja yang logis dan terstruktur dalam membangun model prediksi harga laptop berdasarkan fitur-fitur utama dari spesifikasi perangkat. Tahapan-tahapan tersebut dijelaskan pada Gambar 1.



Gambar 1. Metodologi Penelitian

### 2.1 Pengumpulan Data

Pengumpulan data merupakan langkah utama yang sangat penting dalam penelitian. Tujuan penelitian adalah untuk memastikan kualitas data yang sangat mempengaruhi hasil akhir analisis. Adapun tahapan dalam pengumpulan data yang digunakan pada penelitian ini adalah sebagai berikut:

#### 2.1.1. Sumber Data

Sumber data yang digunakan dalam penelitian ini dengan melalui website *kaggle*. Penulis dapat melalui akses <https://www.kaggle.com/datasets/armaanpreet123/laptop-price-dataset> dengan sebuah platform daring yang menyediakan berbagai macam sumber daya, termasuk kompetisi, ilmu data, machine learning, dan analisis statistik yang melalui secara publik.

#### 2.1.2. Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini yang mencakup berbagai jenis termasuk, teks, numerik, kategorikal, dan temporal yang terakhir diperbarui selama 4 bulan. Penulis mengunduh dataset melalui platform *kaggle* yang menggunakan kata kunci “Laptop Price Dataset”. Data ini berupa mentah yang berisi harga laptop berdasarkan spesifikasi yang terdiri 11768 baris dengan 10 fitur dan 1 label yang termasuk Brand, Processor, RAM, Storage, GPU, Screen Size, Resolution, Operating System, Battery Life, Weight, Price.

## 2.2 Eksplorasi Data

Eksplorasi data dilakukan sebagai tahap awal untuk memahami struktur, distribusi, dan karakteristik dataset sebelum dilakukan preprocessing. Tahapan ini mencakup pemeriksaan jenis data, jumlah nilai kosong, serta analisis statistik. Hasil dari eksplorasi ini menjadi dasar dalam pengambilan keputusan pada tahap selanjutnya, khususnya dalam pemilihan fitur dan strategi preprocessing data.

## 2.3 Preprocessing Data

Sebelum memulai modelling, proses *preprocessing* data yang perlu dilakukan. Tujuan *preprocessing* data adalah untuk meningkatkan kualitas data yang digunakan agar memastikan hasil analisis yang baik dan tahapan ini juga menggunakan informasi dari data yang mentah sehingga data siap dalam pengolahan berikutnya. Adapun tahapan dalam preprocessing data adalah sebagai berikut:

### 2.3.1. Pembersihan Data

Pembersihan data adalah proses untuk menghilangkan elemen yang tidak relevan, memperbaiki kesalahan format, menghapus duplikasi, serta menangani nilai ekstrem (outlier). Langkah ini dilakukan demi menjaga konsistensi dan kualitas data agar siap digunakan dalam proses pengolahan dan analisis lebih lanjut.

### 2.3.2. Transformasi Data Dan Normalisasi Data

Transformasi data merupakan proses mengubah nilai-nilai dalam kumpulan data guna memperbaiki karakteristik atau distribusinya agar lebih sesuai dengan kebutuhan analisis. Tujuan dari transformasi ini adalah untuk meningkatkan kualitas data, memperbaiki kinerja model, dan memenuhi asumsi statistik yang diperlukan dalam pemodelan. Transformasi ini memastikan bahwa data siap digunakan dalam proses machine learning dan mendukung hasil prediksi yang lebih akurat.

Normalisasi data yang bertujuan untuk memastikan bahwa variabel-variabel yang memiliki skala atau rentang nilai yang berbeda-beda sehingga dapat diukur dalam membandingkan secara akurat. Dengan demikian, normalisasi data dapat menjalankan proses data numerik sehingga menghasilkan proses analisis yang secara akurat.

## 2.4 Machine Learning (Pembelajaran Mesin)

*Machine learning* merupakan salah satu pendekatan dalam kecerdasan buatan (*Artificial Intelligence*) yang digunakan untuk meniru atau menggantikan perilaku manusia dalam menyelesaikan masalah secara otomatis. *Machine learning* memiliki karakteristik unik dalam proses pelatihan menggunakan data training, yang memungkinkan sistem belajar dari pengalaman untuk menghasilkan keputusan secara otomatis. [7]. *Machine learning* telah menjadi komponen penting dari kecerdasan buatan dalam perkembangan kemajuan teknologi, khususnya di era digital yang semakin terhubung dan didominasi oleh data. Kecerdasan buatan juga berfokus dengan pengembangan sistem komputer yang mampu menjalankan tugas-tugas dalam membutuhkan kecerdasan manusia. Selain itu, kecerdasan buatan berusaha untuk menciptakan mesin yang dapat memahami, belajar, merencanakan, dan beradaptasi dengan lingkungan sekitar [8].

## 2.5 Prediksi

Menurut Kamus Besar Bahasa Indonesia (KBBI), prediksi merupakan suatu proses yang memperkirakan nilai pada masa yang datang dalam peristiwa yang akan terjadi masa lalu. Prediksi menunjukkan dalam suatu keadaan dengan proses perencanaan dan pengambilan keputusan yang masa depan. Selain pernyataan KBBI prediksi dapat didefinisikan oleh Nasution (2003, yang dicantumkan oleh [9]). merupakan prediksi yang akan memperkirakan dari kebutuhan pada periode masa datang yang mencakup kebutuhan dalam ukuran kuantitas (jumlah), kualitas (mutu), waktu dan lokasi yang diperlukan untuk memenuhi barang atau jasa.

## 2.6 Laptop

Laptop merupakan istilah yang mengacu pada komputer portabel dapat digunakan dalam berbagai tujuan seperti bermain game, bekerja, belajar atau menyelesaikan tugas sehari-hari. Namun, ada banyak berbagai jenis dan ukuran laptop yang berbeda dalam setiap kebutuhan. Laptop ini juga memiliki banyak komponen seperti processor, memori, penyimpanan data, keyboard, touchpad, mouse pad, dan port koneksi seperti HDMI dan USB, dan lainnya. Selain itu, laptop ini memiliki berbagai sistem operasi yang termasuk Windows, macOS, dan Linux untuk menjalankan aplikasi dan mengelola berbagai tugas.[10]

## 2.7 Pemrograman Python

Python merupakan bahasa pemrograman yang pertama kali diperkenalkan oleh Guido van Rossum pada tahun 1991 di Amsterdam, Belanda. Nama Python sendiri terinspirasi dari acara komedi Inggris yang berjudul "*Monty Python's Flying Circus*". Bahasa ini dirancang agar memiliki sintaks yang sederhana, mudah dipahami serta mendukung dalam berbagai gaya pemrograman. Saat ini, Python menjadi salah satu bahasa pemrograman yang sangat populer, khususnya dalam pengembangan web, data science, pengembangan perangkat lunak, dan bidang penelitian, karena fleksibilitas dan kemudahan dalam mempelajarinya [11].

## 2.8 Algoritma Regresi Linear

Regresi Linear merupakan salah satu algoritma dalam teknik statistika yang digunakan untuk menganalisis hubungan antara variabel independen dan variabel dependen [12]. Regresi Linear juga mempunyai dua metode yang termasuk Regresi Linear Sederhana dan Regresi Linear Berganda. Regresi Linear Sederhana yang digunakan ketika hanya satu variabel independen, sedangkan Regresi Linear Berganda juga digunakan apabila terdapat lebih dari satu variabel independen. Persamaan umum yang digunakan dalam algoritma Regresi Linear adalah sebagai berikut [6]:

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

Dimana:

$Y$ : Variabel dependen (Price)

$X$ : Variabel independen (Brand, Processor, RAM dan GPU)

$a$ : Intercept atau konstanta

$b$ : Koefisien regresi dalam setiap perubahan satu unit pada variabel  $X$

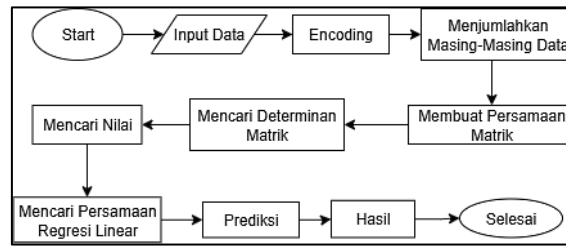
$\varepsilon$ : Error atau residual (selisih antara nilai aktual dan prediksi)

Received: 21-07-2025 | Accepted: 28-07-2025 | Published Online: 30-08-2025

All author: Risky Harahap, Karpen, Helda Yenni, Muhammad Jamaris

### 2.8.1 Prinsip Kerja

Adapun prinsip kerja dari algoritma Regresi Linear yang diimplementasikan dalam penelitian dapat dijelaskan melalui tahapan-tahapan berikut [13]:



Gambar 2. Alur Prinsip Kerja Algoritma Regresi Linear

Berdasarkan alur pada Gambar 2 diatas, tahapan prinsip kerja algoritma Regresi Linear dapat dijelaskan sebagai berikut:

1. Proses ini dimulai dengan insialisasi program. Data dimasukkan oleh pengguna baik secara manual maupun melalui file CSV.
2. Setelah data dimasukkan, maka dilakukan proses encoding untuk mengubahnya menjadi nilai numerik.
3. Setelah semua variabel berada dalam format numerik, maka proses penjumlahan data sesuai dengan masing-masing kolom.
4. Berdasarkan data yang telah diproses, maka dapat dibentuk beberapa matriks, yaitu matriks utama dan matriks turunan yang sesuai dengan jumlah variabel independen.
5. Setiap seluruh matriks terbentuk dapat dihitung nilai determinan dari masing-masing matriks dengan menggunakan perangkat lunak Microsoft Excel.
6. Nilai determinan digunakan untuk menghitung koefisien regresi.
7. Nilai koefisien Regresi kemudian disusun ke dalam bentuk persamaan Regresi Linear disajikan pada 2.1
8. Pengguna dapat memasukkan nilai variabel independen untuk dilakukan prediksi harga laptop, lalu nilai data kategorikal dari input harus dimapping terlebih dahulu dengan metode yang sama seperti saat pelatihan model.
9. Nilai input telah diubah ke bentuk numerik kemudian dimasukkan ke dalam persamaan Regresi untuk menghitung nilai prediksi harga laptop.
10. Setelah hasil akhir dalam proses perhitungan, maka ditampilkan sebagai output prediksi harga laptop berdasarkan spesifikasi yang berikan. Proses ini juga berakhir setelah hasil yang ditampilkan.

### 2.8.2 Kelebihan Dan Kekurangan

Adapun kelebihan dan kekurangan dalam algoritma Regresi Linear adalah sebagai berikut [14]:

1. Salah satu kelebihan utama dalam algoritma Regresi Linear adalah kemudahan dalam implementasi serta efisiensi penggunaannya, sehingga algoritma ini memiliki kapasitas komputasi terbatas. Bentuk persamaan matematis dalam Regresi Linear mudah dipahami dan memiliki kompleksitas waktu yang rendah.
2. Regresi Linear juga memiliki kekurangan, salah satunya adalah mengasumsi bahwa setiap variabel bersifat independen. Selain itu, algoritma ini rentan mengalami underfitting karena model tidak mampu mengenali pola dalam data secara optimal. Selain itu, algoritma ini juga sangat sensitif terhadap outlier yaitu, nilai-nilai ekstrem yang menyimpang dari distribusi data.

## 2.9 Evaluasi Model

Evaluasi model juga dilakukan untuk mengukur kinerja algoritma Regresi Linear dalam memprediksi harga laptop. Proses evaluasi dilakukan menggunakan data uji (testing set) setelah model diuji maka, juga diperlukan data latih (training set). Evaluasi ini juga bertujuan untuk mengetahui performa model, serta melihat seberapa besar error yang dihasilkan model dalam memprediksi harga laptop. Model ini dibangun dalam menggunakan empat metrik adalah sebagai berikut:

1. *Mean Absolute Error* (MAE)

Mengukur rata-rata kesalahan absolut antara nilai actual dan prediksi. Semakin rendah nilai MAE, semakin baik dalam memprediksi harga laptop.

$$MAE = \frac{1}{n} \sum_{i=0}^n |y_i - \hat{y}_i| \quad (2)$$

2. *Mean Square Error* (MSE)

Menghitung kesalahan kuadrat rata-rata antara nilai actual dan prediksi. Nilai RMSE yang rendah menunjukkan model memiliki error yang kecil dalam prediksi harga laptop.

$$MSE = \sqrt{\frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2} \quad (3)$$

3. *Root Mean Square Error* (RMSE)

RMSE juga merupakan akar dari MSE yang memberikan gambaran seberapa rata-rata kesalahan yang sama dengan target harga laptop

$$RMSE = \sqrt{MSE} \quad (4)$$

4. *R-squared* ( $R^2$ )

Mengukur seberapa baik model dapat menjelaskan variabilitas data actual. Nilai R-squared yang mendekati 1 menunjukkan bahwa memiliki kemampuan yang baik dalam prediksi harga laptop.

$$R^2 = 1 - \frac{\sum_{i=0}^n (y_i - \hat{y}_i)^2}{\sum_{i=0}^n (y_i - \bar{y})^2} \quad (5)$$

## 3. HASIL DAN PEMBAHASAN

### 3.1. Dataset

Dataset yang digunakan penelitian ini berasal dari *kaggle dataset* yang digunakan dalam penelitian

ini berasal dari situs *kaggle* yang menggunakan kata kunci “*Laptop Price Dataset*”. Dataset ini juga memiliki 11768 baris terdiri dari 10 fitur dan 1 label. *Dataset* yang disajikan pada gambar 3.

	Brand	Processor	RAM	Storage	GPU	Screen Size	Resolution	Battery Life	Weight	Operating System	Price
0	Apple	AMD Ryzen 3	64	512GB SSD	Nvidia GTX 1650	17.3	2560x1440	8.9	1.42	FreeDOS	3997.07
1	Razer	AMD Ryzen 7	4	1TB SSD	Nvidia RTX 3080	14.0	1366x768	9.4	2.57	Linux	1355.78
2	Asus	Intel i5	32	2TB SSD	Nvidia RTX 3060	13.3	3840x2160	8.5	1.74	FreeDOS	2673.07
3	Lenovo	Intel i5	4	256GB SSD	Nvidia RTX 3080	13.3	1366x768	10.5	3.10	Windows	751.17
4	Razer	Intel i3	4	256GB SSD	AMD Radeon RX 6600	16.0	3840x2160	5.7	3.38	Linux	2059.83
...	...	...	...	...	...	...	...	...	...	...	...
11763	Acer	Intel i3	4	2TB SSD	Nvidia RTX 2060	17.3	1366x768	11.5	1.58	macOS	704.82
11764	Asus	Intel i3	4	2TB SSD	AMD Radeon RX 6800	16.0	1366x768	9.5	2.14	Linux	775.59
11765	Razer	AMD Ryzen 9	4	2TB SSD	AMD Radeon RX 6600	15.6	2560x1440	8.2	2.05	Linux	2789.46
11766	Samsung	AMD Ryzen 7	16	512GB SSD	Integrated	13.3	1920x1080	7.5	1.48	macOS	1067.13
11767	Samsung	Intel i7	8	256GB SSD	Nvidia RTX 3080	17.3	2560x1440	6.4	2.45	FreeDOS	1579.55

Gambar 3. Tampilan Awal *Dataset*

Pada Gambar 3 menampilkan dataset berisi informasi harga laptop yang terdiri dari fitur, yaitu *Brand, Processor, RAM, Storage, GPU, Screen Size, Resolution, Operating System, Battery Life, Weight, Price*. Penjelasan masing-masing fitur adalah:

1. Brand : Nama merek atau produsen laptop.
2. Processor : Komponen utama memproses data dan menjalankan perintah..
3. RAM : Kapasitas memory sementara yang mendukung kinerja aplikasi.
4. Storage : Media penyimpanan data, seperti HDD/SDD.
5. GPU : Komponen untuk pengolahan grafis dan tampilan visual.
6. Screen Size : Ukuran layar laptop secara diagonal dalam inci.
7. Resolution : Jumlah piksel layar yang menentukan tingkat kerjenihan.
8. Battery Life : Durasi daya tahan baterai saat laptop digunakan.
9. Weight : Jumlah berat laptop keseluruhan.
10. System Operation: Sistem yang mengatur dan mengelola perangkat keras serta perangkat lunak.
11. Price : Harga jual laptop dipasaran.

```

Informasi Dataset Awal:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11768 entries, 0 to 11767
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Brand                  11768 non-null object
1   Processor              11768 non-null object
2   RAM                    11768 non-null int64
3   Storage                11768 non-null object
4   GPU                    11768 non-null object
5   Screen Size            11768 non-null float64
6   Resolution             11768 non-null object
7   Battery Life           11768 non-null float64
8   Weight                 11768 non-null float64
9   Operating System       11768 non-null object
10  Price                  11768 non-null float64
dtypes: float64(4), int64(1), object(6)
memory usage: 1011.4+ KB
    
```

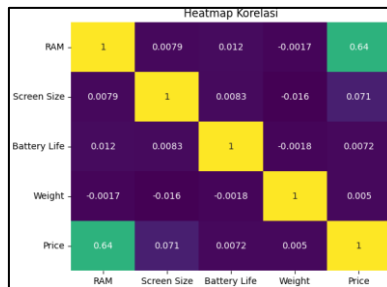
Gambar 4. Informasi Dataset Awal

Gambar 4 menampilkan struktur *dataset* harga laptop yang diperoleh menggunakan fungsi `df.info()` dari pustaka Pandas. Dataset ini terdiri dari 11 kolom dan 11.768 baris, dengan berbagai tipe data seperti

objek, integer, dan float. Informasi ini juga memudahkan pemahaman mengenai tipe data dan kelengkapan fitur sebelum tahap *preprocessing* dan pemodelan.

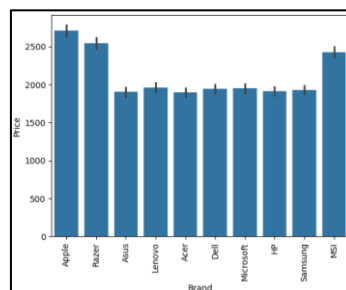
### 3.2. Eksplorasi Data

Sebelum memasuki tahap *preprocessing*, dilakukan tahap eksplorasi data (Exploratory Data Analysis/EDA) untuk memahami pola, hubungan antar fitur, dan karakteristik distribusi data dalam dataset harga laptop. Hasil eksplorasi ini menjadi dasar dalam pengambilan keputusan pada tahap *preprocessing* dan pemodelan.



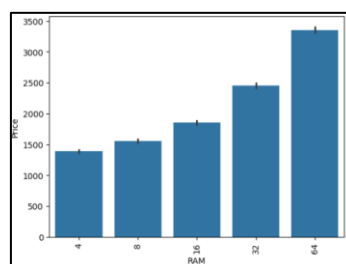
Gambar 5. Heatmap Korelasi Fitur

Pada Gambar 5 ditampilkan hasil analisis menunjukkan bahwa fitur RAM memiliki korelasi positif yang cukup kuat terhadap harga laptop dengan nilai 0,636. Fitur seperti, *Screen Size* (0,071), *Battery Life* (0,007), dan *Weight* (0,005) memiliki nilai korelasi yang sangat rendah. Informasi ini sangat penting dalam proses analisis dan pemodelan, karena membantu menentukan fitur-fitur yang memberikan kontribusi signifikan terhadap variabel target.



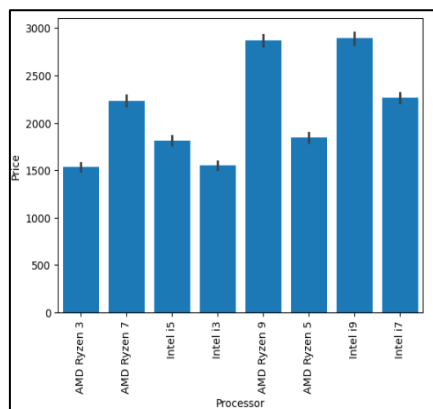
Gambar 6. Visualisasi Rata-Rata Harga Laptop Berdasarkan Brand

Pada Gambar 6 ditampilkan hasil analisis menunjukkan bahwa Apple memiliki harga rata-rata tertinggi, mencerminkan posisi sebagai produsen perangkat premium. Sementara itu, HP, Asus, dan Acer cenderung memiliki harga rata-rata lebih rendah, sesuai dengan segmentasi pasar menengah ke bawah. Temuan ini menunjukkan bahwa merek laptop memiliki pengaruh signifikan terhadap harga, karena setiap brand mengusung positioning dan target konsumen yang berbeda dalam pasar.



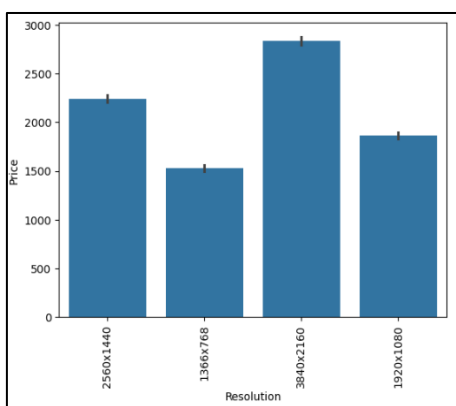
Gambar 7. Visualisasi Rata-Rata Harga Laptop Berdasarkan RAM

Pada gambar 7 ditampilkan hasil analisis menunjukkan bahwa laptop dengan RAM 64 GB memiliki harga rata-rata tertinggi, menandakan segmentasi untuk pasar premium dengan kebutuhan komputasi intensif. Sebaliknya, laptop dengan kapasitas RAM 8 GB dan 4 GB berada pada kisaran harga yang lebih rendah, mencerminkan segmentasi untuk pengguna umum atau kelas entry-level. Dari analisis ini dapat disimpulkan bahwa kapasitas RAM memiliki kontribusi signifikan dalam menentukan harga laptop, dan menjadi salah satu fitur penting dalam model prediksi harga.



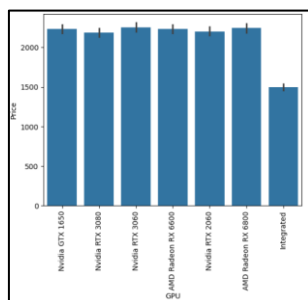
Gambar 8. Visualisasi Rata-Rata Harga Laptop Berdasarkan Processor

Pada Gambar 8 ditampilkan hasil analisis menunjukkan bahwa prosesor kelas atas seperti Intel Core i9 dan AMD Ryzen 9 berada pada posisi teratas dengan harga rata-rata tertinggi, yang mencerminkan segmentasi produk untuk pasar premium dengan kebutuhan performa tinggi. Sementara itu, prosesor Intel i5, i3, Ryzen 5, dan Ryzen 3 cenderung memiliki harga lebih rendah, menggambarkan laptop untuk kebutuhan komputasi standar atau entry-level. Dari hasil ini dapat disimpulkan bahwa jenis prosesor merupakan salah satu faktor utama yang memengaruhi harga laptop, seiring dengan perbedaan kapabilitas dan target pasar dari masing-masing kelas prosesor.



Gambar 9. Visualisasi Rata-Rata Harga Laptop Berdasarkan Resolution

Pada Gambar 9 ditampilkan hasil analisis menunjukkan bahwa laptop dengan resolusi tinggi seperti 3840x2160 (4K) memiliki harga tertinggi yang juga berada pada kisaran harga premium, sementara resolusi lebih rendah seperti 1366x768 (HD) memiliki harga paling rendah. Fenomena ini mengindikasikan bahwa resolusi layar merupakan faktor yang berpengaruh terhadap harga jual laptop, karena semakin tinggi kualitas dan kerapatan piksel layar, semakin tinggi pula harga yang ditawarkan.



Gambar 10. Visualisasi Rata-Rata Harga Laptop Berdasarkan GPU

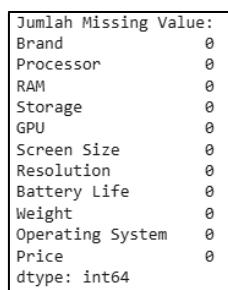
Pada Gambar 10 ditampilkan hasil analisis menunjukkan bahwa laptop yang menggunakan GPU kelas menengah ke atas seperti Nvidia RTX 3060, AMD Radeon RX 6800, dan RX 6600 memiliki rata-rata harga yang tinggi. Sebaliknya, laptop yang menggunakan GPU terintegrasi (Integrated GPU) cenderung memiliki harga lebih rendah, karena ditujukan untuk penggunaan ringan seperti pekerjaan kantor atau keperluan sehari-hari.

### 3.3. Preprocessing Data

Sebelum melakukan proses pelatihan model, diperlukan tahapan preprocessing bertujuan untuk membersihkan, mengubah dan mempersiapkan data agar siap digunakan oleh model *machine learning*. Secara umum, preprocessing dalam penelitian ini dilakukan melalui beberapa tahap sebagai berikut:

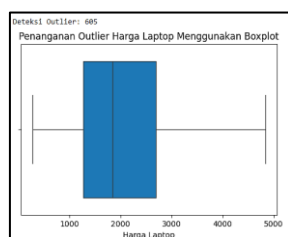
#### 3.3.1 Pembersihan Data

Proses pembersihan data diawali dengan pemeriksaan *missing value* dan deteksi outlier, untuk memastikan bahwa dataset yang digunakan bersih, konsisten, dan tidak mengandung anomali yang dapat memengaruhi kinerja model.



Gambar 11. Missing Value

Gambar 11 menampilkan hasil bahwa seluruh kolom dalam *dataset* tidak memiliki nilai kosong atau *missing value*. Hal ini menunjukkan bahwa data yang bersih, maka dataset siap digunakan untuk proses analisis dan pemodelan.



Gambar 12. Visualisasi Penanganan Outlier Terhadap Harga laptop

Pada Gambar 12 menampilkan hasil visualisasi ditemukan sebanyak 605 data yang tergolong sebagai *outlier*. Keberadaan outlier ini berpotensi memengaruhi hasil analisis dan akurasi

model prediksi dalam meningkatkan kualitas data, serta membantu model Regresi Linear menghasilkan prediksi harga laptop yang lebih akurat, dan andal.

Dalam pembangunan model prediksi harga laptop, dilakukan seleksi fitur untuk memilih variabel yang tidak digunakan karena pengaruhnya rendah terhadap harga dalam pengolahan data. Rincian fitur-fitur tersebut beserta alasan penghapusannya disajikan pada Tabel 1.

Tabel 1. Fitur Yang Tidak Relevan

Fitur	Alasan
<i>Storage</i>	Memiliki variasi tipe dan kapasitas yang sangat menyulitkan proses <i>preprocessing</i> .
<i>Screen Size</i>	Menunjukkan korelasi rendah terhadap harga.
<i>Resolution</i>	Format data yang tidak konsisten juga menyulitkan proses <i>preprocessing</i> .
<i>Battery Life</i>	Memiliki korelasi yang sangat rendah terhadap harga laptop.
<i>Operating System</i>	Didominasi oleh sistem operasi umum, sehingga variasinya rendah dan tidak memiliki pengaruh signifikan terhadap harga.
<i>Weight</i>	Variasi bobot antar laptop tidak cukup besar untuk memengaruhi harga secara signifikan.

### 3.3.2 Transformasi Data

Beberapa fitur dalam dataset dapat digunakan dalam model Regresi Linear yang dilakukan transformasi label berdasarkan segmen pasar dan performa produk. Tujuan dari pendekatan ini adalah untuk menyisipkan makna ordinal atau hirarki performa ke dalam representasi numerik.

#### 1. Brand (Merek Laptop)

Merek laptop dikelompokkan berdasarkan segmentasi pasar, mulai dari brand *entry-level* hingga brand premium. Hasil transformasi fitur Brand ditampilkan pada Tabel 2.

Tabel 2. Hasil Transformasi Fitur *Processor*

Brand	Nilai Numerik	Keterangan
Acer	0	<i>Entry-level</i>
Asus	1	<i>Mid-range</i>
Lenovo	2	<i>Bisnis</i>
HP	3	<i>Mid-range/bisnis</i>
Samsung	4	<i>Umum</i>
Dell	5	<i>Bisnis</i>
MSI	6	<i>Gaming</i>
Microsoft	7	<i>Bisnis</i>
Razer	8	<i>Gaming premium</i>
Apple	9	<i>Premium</i>

#### 2. Processor

*Prosesor* diklasifikasikan berdasarkan kelas performa *prosesor*, dari *entry-level* hingga kelas premium. Hasil transformasi fitur *Processor* yang disajikan pada Tabel 3.

Tabel 1. Hasil Transformasi Fitur Processor

Processor	Nilai Numerik	Keterangan
Intel i3	0	<i>Entry-level</i>
AMD Ryzen 3	1	<i>Entry-level</i>
Intel i5	2	<i>Mid-range</i>
AMD Ryzen 5	3	<i>Mid-range</i>

Processor	Nilai Numerik	Keterangan
Intel i7	4	High performance
AMD Ryzen 7	5	High performance
Intel i9	6	Premium
AMD Ryzen 9	7	Premium

### 3. GPU

GPU dikategorikan berdasarkan kemampuan grafis, dari GPU bawaan (integrated) hingga GPU kelas atas untuk gaming berat. Hasil transformasi fitur GPU yang disajikan pada Tabel 4.

Tabel 4. Hasil Transformasi Fitur GPU

GPU	Nilai Numerik	Keterangan
Integrated	0	Entry-level
Nvidia GTX 1650	1	Mid-range
Nvidia RTX 2060	2	Mid-range
AMD Radeon RX 6600	3	Mid-range
Nvidia RTX 3060	4	High-performance
AMD Radeon RX 6800	5	High-performance
Nvidia RTX 3080	6	Premium

### 3.4. Hasil Implementasi Regresi Linear

Proses implementasi algoritma Regresi Linear dilakukan untuk memprediksi harga laptop berdasarkan beberapa fitur utama, yaitu merek (*Brand*), Processor, RAM, dan GPU. Model dibangun menggunakan 11.768 baris data, dengan pemrograman Python melalui platform Google Colab. Data dibagi menjadi 80% sebagai data latih dan 20% sebagai data uji guna melatih dan mengevaluasi kinerja model secara terpisah. Hasil dari pelatihan model ini menghasilkan persamaan Regresi Linear yang ditampilkan pada Gambar 13.

```
Koefisien Regresi Linear:
Intercept : 12.9211
Brand : 82.4584
Processor : 206.5498
RAM : 32.5502
GPU : 70.5528
```

Gambar 13. Hasil Persamaan Regresi Linear

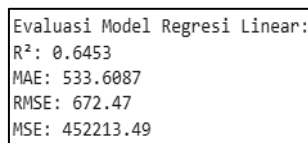
```
HASIL PREDIKSI HARGA
-----
Brand : Apple
Processor : Intel i3
GPU : Nvidia GTX 1650
RAM : 64 GB
-----
Harga USD : $2,913.22
Harga IDR : Rp 46,011,451
Kurs : 1 USD = Rp 16,000
-----
```

Gambar 14. Prediksi Regresi linear

Berdasarkan Gambar 14, ditampilkan hasil ini menunjukkan bahwa spesifikasi dengan menggunakan prosesor kelas menengah tetap menghasilkan prediksi harga yang tinggi. Hal ini mendukung temuan sebelumnya memiliki kontribusi besar terhadap harga laptop, dan model dapat memperkirakan harga secara logis berdasarkan kombinasi fitur yang diberikan.

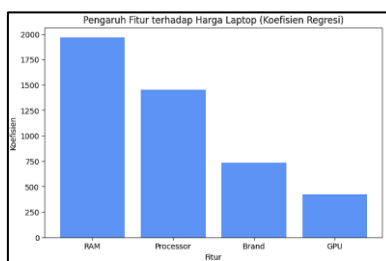
Model Regresi Linear yang digunakan dalam penelitian ini untuk mengukur kinerja model dalam memprediksi harga, digunakan sejumlah metrik evaluasi yaitu: Koefisien Determinasi ( $R^2$ ), Mean *Absolute*

Error (MAE), Root Mean Squared Error (RMSE), dan Mean Squared Error (MSE). Seluruh hasil evaluasi ditampilkan pada Gambar 15 sebagai berikut:



Gambar 15. Evaluasi Regresi Linear

1. R<sup>2</sup> = 0,6453 Menunjukkan bahwa model mampu menjelaskan sekitar 64,5% variasi dari data harga laptop.
2. Mean Absolut Error (MAE) = 533,61 Menunjukkan bahwa rata-rata selisih antara nilai prediksi dan nilai asli adalah sekitar 533 USD.
3. Root Mean Squared Error (RMSE) = 672,47 Menunjukkan hasil sedikit lebih besar dari MAE karena menghitung akar dari kesalahan kuadrat.
5. Mean Absolut Error (MSE) = 452,213.49 Menunjukkan hasil kuadrat dari perbedaan antara nilai aktual dan nilai prediksi.



Gambar 16. Visualisasi Pengaruh Fitur Terhadap Harga Laptop

Pada Gambar 16 menunjukkan hasil visualisasi pengaruh fitur-fitur utama terhadap harga laptop berdasarkan koefisien regresi linear. Visualisasi ini memperlihatkan besarnya kontribusi masing-masing fitur dalam memengaruhi prediksi harga laptop. Penyajian visual yang ringkas ini juga menunjukkan kemampuan model dalam menginterpretasi data, sehingga mempermudah analisis dan pengambilan keputusan dalam proses prediksi harga laptop.

#### 4. KESIMPULAN

*Dataset* yang digunakan dalam penelitian ini memiliki jumlah data yang cukup besar, sehingga mempercepat proses analisis dan meningkatkan akurasi pemodelan. Hasil analisis korelasi menunjukkan bahwa fitur RAM yang paling mempengaruhi harga laptop. Proses preprocessing yang dilakukan meliputi encoding, normalisasi, dan penanganan outlier untuk meningkatkan performa model Regresi Linear. Model ini menghasilkan nilai R<sup>2</sup> sebesar 0,6453, yang menunjukkan kemampuan prediksi yang cukup baik. Model berhasil diintegrasikan dalam aplikasi berbasis Streamlit untuk prediksi harga secara langsung dan visualisasi kontribusi fitur. Untuk pengembangan lebih lanjut, disarankan penggunaan model non-linear seperti XGBoost, LightGBM, dan SVR, penambahan fitur seperti jenis penyimpanan, refresh rate layar, dan tahun rilis, lokasi atau segmentasi regional, serta validasi hasil prediksi dengan data aktual. Pengembangan aplikasi juga dapat diarahkan ke tampilan yang lebih interaktif, seperti visualisasi harga dalam bentuk grafik,

atau integrasi API dalam sistem rekomendasi harga jual beli laptop untuk membantu menetapkan harga jual optimal.

## DAFTAR RUJUKAN

- [1] N. F. Budiarto, "Preferensi Merek Terhadap Ekuitas Merek Laptop Lenovo Universitas Pendidikan Indonesia," 2016.
- [2] V. Setiady And V. Meliana, "Pengaruh Brand Credibility Dan Brand Attitude Terhadap Minat Beli Laptop Huawei," *Kalbisiana J. Sains, Bisnis Dan Teknol.*, Vol. 8, No. 2, Pp. 2247–2259, 2022.
- [3] S. N. Wardani And N. Nurmalitasari, "Penerapan Regresi Linier Berganda Untuk Memprediksi Harga Laptop Dengan Menggunakan Software Python," In *Prosiding Seminar Nasional Teknologi Informasi Dan Bisnis*, 2023, Pp. 377–382.
- [4] E. N. Waroi, A. Arief, And K. Khusnawi, "Prediksi Harga Laptop Menggunakan Algoritma Gru Dan Bilstm," *J. Sos. Teknol.*, Vol. 4, No. 7, Pp. 408–424, 2024.
- [5] M. Kafil, "Penerapan Metode K-Nearest Neighbors Untuk Prediksi Penjualan Berbasis Web Pada Boutiq Dealove Bondowoso," *Jati (Jurnal Mhs. Tek. Inform.)*, Vol. 3, No. 2, Pp. 59–66, 2019.
- [6] P. B. Utomo, E. Utami, And S. Raharjo, "Implementasi Metode K-Nearest Neighbor Dan Regresi Linear Dalam Prediksi Harga Emas," *J. Inf. Interaktif*, Vol. 4, No. 3, Pp. 155–159, 2019.
- [7] A. Ahmad Hania, "Mengenal Artificial Intelligence, Machine Learning, & Deep Learning," *J. Teknol. Indones.*, Vol. 1, No. June, Pp. 1–6, 2017.
- [8] R. F. Putra *Et Al.*, *Algoritma Pembelajaran Mesin (Dasar, Teknik, Dan Aplikasi)*, Vol. 19, No. 5. 2016.
- [9] C. P. Ananda, "Machine Learning Untuk Prediksi Gaya Hidup Berdasarkan Socioeconomic Status Ses Menggunakan Algoritma Catboost Studi Kasus: Mahasiswa Uin Jakarta," Fakultas Sains Dan Teknologi Uin Syarif Hidayatullah Jakarta, 2023.
- [10] P. C. Putro, "Penggunaan Laptop Bagi Pemenuhan Kebutuhan Mahasiswa Ditinjau Dari Teori Konsumsi Islam (Studi Mahasiswa Ekonomi Syariah Angkatan 2019)," Iain Metro, 2024.
- [11] A. Ritonga And Y. Yahfizham, "Studi Literatur Perbandingan Bahasa Pemrograman C++ Dan Bahasa Pemrograman Python Pada Algoritma Pemrograman," *J. Tek. Inform. Dan Teknol. Inf.*, Vol. 3, No. 3, Pp. 56–63, 2023.
- [12] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, And C. Lauw, "Penerapan Aplikasi Rapidminer Untuk Prediksi Nilai Tukar Rupiah Terhadap Us Dollar Dengan Metode Linear Regression," *J. Nas. Teknol. Dan Sist. Inf.*, Vol. 7, No. 1, Pp. 8–17, 2021.
- [13] E. Triyanto, H. Sismoro, And A. D. Laksito, "Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul," *Rabit J. Teknol. Dan Sist. Inf. Univrab*, Vol. 4, No. 2, Pp. 66–75, 2019.
- [14] M. A. Iqbal, "Application Of Regression Techniques With Their Advantages And Disadvantages," *Elektron Mag.*, No. September, Pp. 11–17, 2021.