

Implementasi *Machine Learning* untuk Prediksi Harga Laptop Menggunakan Algoritma Regresi Linear Berganda

Machine Learning Implementation for Laptop Price Prediction Using Multiple Linear Regression Algorithm

Noer Nilam Sari^{1*}, Tiyar Tohirotun Anisah², Risma Fitriani³

^{1,2}Program Studi Akuntansi, Universitas Singaperbangsa Karawang, Karawang, Indonesia

³Program Studi Teknik Industri, Universitas Singaperbangsa Karawang, Karawang, Indonesia

*E-mail: noernilamsarii1@gmail.com

Abstrak

Artikel ini membahas implementasi pembelajaran mesin (*machine learning*) menggunakan algoritma regresi linear berganda guna melakukan prediksi harga laptop. Tujuan utama penelitian ini, yaitu untuk merancang model prediktif yang tepat berdasarkan berbagai fitur seperti spesifikasi teknis dan merek laptop. Tahapan penelitian mencakup studi literatur, pengumpulan dataset yang relevan, pra-pemrosesan atau data cleaning, *Exploratory Data Analysis*, *Feature Engineering*, pemisahan data, pembentukan model, dan pengembangan website. Hasil penelitian menunjukkan bahwa model yang diusulkan mampu memberikan prediksi harga dengan tingkat ketepatan yang tinggi, diukur menggunakan metrik evaluasi seperti Mean Absolute Error (MAE) yang menghasilkan nilai sebesar 0.18 dan R-squared (R^2) yang memiliki nilai 0.68. Implementasi dilakukan dengan mengintegrasikan *Jupyter Notebook*, *Visual Studio Code (VS Code)* dan *Streamlit* untuk pengembangan aplikasi web interaktif. Kesimpulan dari penelitian ini adalah bahwa algoritma regresi linear berganda efektif dalam memprediksi harga laptop, serta dapat digunakan sebagai alat bantu dalam pengambilan keputusan bisnis dan strategi penetapan harga.

Kata kunci: *Prediksi Harga, Regresi Linear, Machine Learning, Laptop*

Abstract

This article discusses the implementation of machine learning using multiple linear regression algorithms to predict laptop prices. The main objective of this research is to design an appropriate predictive model based on various features such as technical specifications and laptop brands. The research stages include literature study, collection of relevant datasets, pre-processing or data cleaning, *Exploratory Data Analysis*, *Feature Engineering*, data splitting, model building, and website development. The results show that the proposed model is able to provide price predictions with a high level of accuracy, measured using evaluation metrics such as Mean Absolute Error (MAE) which produces a value of 0.18 and R-squared (R^2) which has a value of 0.68. Implementation was done by integrating *Jupyter Notebook*, *Visual Studio Code (VS Code)* and *Streamlit* for interactive web application development. The conclusion of this research is that the multiple linear regression algorithm is effective in predicting laptop prices, and can be used as a tool in making business decisions and pricing strategies.

Keywords: *Price Prediction, Linear Regression, Machine Learning, Laptop*

Naskah diterima 30 Mei 2024; direvisi 01 Jul. 2024; dipublikasikan 01 Okt. 2024.

JAMIKA is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



I. PENDAHULUAN

Di tengah gempuran revolusi industri 4.0 yang menjadikan teknologi kian berkembang pesat, masyarakat makin dibuat berkebutuhan akan penggunaan teknologi di kesehariannya. Teknologi dapat memberikan khalayak berbagai manfaat termasuk dalam mendapatkan informasi penting terkait peluang dalam penjualan produknya bagi masyarakat yang menekuni dunia berbisnis [1]. Para penjual diharuskan untuk menggunakan peluang yang ada sebaik mungkin untuk meraih keuntungan yang maksimal dengan salah satu cara yang mungkin berupa usaha dalam memprediksi apa yang akan terjadi di masa mendatang.

Peluang yang ada dalam meningkatkan keuntungan bagi penjual yang dalam kajian ini mengungkit para penjual produk elektronik yang ikut terbawa arus perkembangan industri elektronika. Dalam hal ini produk

elektronik, yaitu laptop kian terdapat variansi yang beraneka macam dengan kualitas tinggi yang ditawarkan masing-masing merek laptop. Penawaran tersebut serta didukung dengan keadaan yang mengharuskan membuat masyarakat makin minat untuk memiliki. Keadaan yang mengharuskan ini beberapa terkhususnya yang dialami oleh para pelajar ataupun mahasiswa dan para pekerja lainnya yang membutuhkan tersedianya laptop untuk membantu menyelesaikan urusannya yang berhubungan dengan proses pengetikan, analisis olah data, tugas-tugas dalam kegiatan pembelajaran dan lainnya yang dapat sangat terbantu dengan adanya laptop [2]. Analisis perlu dilakukan oleh penjual laptop untuk mengetahui berbagai merek laptop yang dijualnya, amat penting untuk dilakukan analisa terkait merek laptop yang sedang populer di khalayak ramai serta merek-merek yang paling banyak digunakan untuk kemudian dilakukan proses manajemen untuk mengendalikan totalan produk yang dijual [1].

Laptop ditawarkan dengan berbagai harga sesuai dengan spesifikasinya. Masyarakat yang ingin membeli laptop seringkali dihadapkan dengan harga yang harus menyesuaikan dengan kebutuhan. Dengan kisaran biaya yang sudah dipersiapkan untuk memperoleh sebuah laptop, tentunya membuat para pembeli harus benar-benar hilangkan kebimbangan dan memilih laptop dengan penuh keyakinan ditengah banyaknya spesifikasi laptop yang ditawarkan dari berbagai merek yang saling memiliki perbedaan dalam segi keunggulan [3]. Hal ini juga turut memberikan tuntutan bagi pembeli untuk memahami dan dapat berbagai informasi mengenai kelebihan serta kekurangan yang dimiliki laptop pilihannya. Dengan adanya permasalahan tersebut, maka dijadikan latar belakang penerapan *Machine Learning* guna prediksi harga laptop dengan penggunaan algoritma regresi linear berganda guna melihat nilai *Rsquare* dan MAE dari model.

Prediksi harga laptop merupakan proses yang dilakukan untuk memperkirakan harga laptop di masa depan berdasarkan informasi harga laptop dari masa lalu [4]. Tujuan dari Prediksi ini adalah untuk menemukan harga terdekat mungkin yang akan terjadi, tidak pasti tepat akan tetapi setidaknya dengan dilakukan prediksi, akan dapat memperkecil selisih dari harga sesungguhnya dengan hasil perkiraan [5]. Project prediksi harga laptop ini dapat bermanfaat baik bagi pembeli ataupun penjual. pembeli dapat gunakan sebagai rujukan pengambilan keputusan terkait laptop yang ingin dibeli dengan memperoleh informasi seputar target harga dan tampilan lainnya mengenai spesifikasi laptop yang dicari. Manfaat yang diraih penjual, yaitu sebagai rujukan dalam menentukan harga jual untuk laptop di masa mendatang, seperti penjualan beberapa bulan setelahnya. Prediksi ini dapat dilakukan penjual agar dapat menetapkan harga laptop dengan bijak melalui manajemen yang baik sehingga tetap memaksimalkan keuntungan dari hasil penjualan laptop tersebut.

Otomatisasi sistem komputer dalam proses perancangan prediksi akan dijalankan dengan menerapkan ilmu pengembangan algoritma *Machine Learning* yang memungkinkan teknologi untuk dapat menciptakan mesin guna memproses pembelajaran secara mandiri tanpa harus diarahkan oleh penggunanya [6]. *Machine Learning* bekerja dengan diawali tahapan input dataset terkait permasalahan yang ingin diproses untuk kemudian dipelajari melalui algoritma yang telah ditetapkan [7]. Algoritma *Machine Learning* terdiri dari beberapa jenis yang dapat diterapkan dalam pengujian, salah satunya adalah algoritma jenis *Supervised Learning* yang diterapkan dalam model prediksi ini. *Supervised Learning* merupakan salah satu bagian dari pengelompokan yang ada pada metode pembelajaran *machine learning*. *Supervised Learning* diartikan sebagai konsep ajar yang dalam pengawasan atau dalam satu jalan yang terarah. *Supervised Learning* mengenakan label atas data train untuk kemudian dijadikan acuan dalam pengamatan label data yang baru. Model dalam *Supervised Learning* diciptakan guna sebagai bahan menghasilkan output yang terperinci dengan dasar masukan sebelumnya.

Salah satu algoritma yang tergolong kedalam *Supervised learning* cakupan *Machine Learning* adalah algoritma Regresi Linear Berganda. Regresi Linear Berganda merupakan suatu regresi linear yang terdiri atas 2 variabel independen atau lebih [8]. Regresi linear digunakan untuk memperkirakan perilaku variabel terikat atas dasar nilai beberapa variabel bebas disekitarnya [6]. Model regresi yang telah dibentuk kemudian akan digunakan sebagai dasar prediksi harga laptop sehingga perlu dilakukan pengukuran keakuratan model prediksi dengan mencari nilai *R Square* dan *Mean Absolute Error*.

Nilai R^2 atau *R square* merupakan analisa yang dapat dilakukan terhadap model regresi untuk mendapatkan hasil besaran variasi yang ada di variabel terikat yang dapat diterangkan oleh variasi dalam variabel bebasnya. Terdapat rentang antar 0 hingga 1 yang mana semakin nilai koefisien determinan mendekati angka 1, maka lebih besar pula kemampuan variasi variabel bebas tersebut dalam jabarkan variasi variabel terikat dengan amat baik [9]. *R square* dapat digunakan untuk melakukan sebuah pengamatan atau prediksi dengan hasil akhir yang berupa nilai keakuratan atas seluruh bagian variabel independen yang mampu menjelaskan nilai variabel dependen.

Dalam algoritma regresi linear terdapat istilah *Mean Absolute Error* (MAE) yang didefinisikan sebagai metode guna menghitung tingkatan akurat dalam sebuah model regresi untuk perkiraan tertentu. Output dari analisa MAE akan menerangkan seberapa rata-rata dari error absolut dari hasil perkiraan dengan hasil sebenarnya [2]. Nilai MAE yang dianggap bagus adalah nilai yang makin kecil dikarenakan dengan ini berarti dapat disimpulkan model tersebut akan semakin baik atau bagus dalam tindakan prediksi yang dilakukannya.

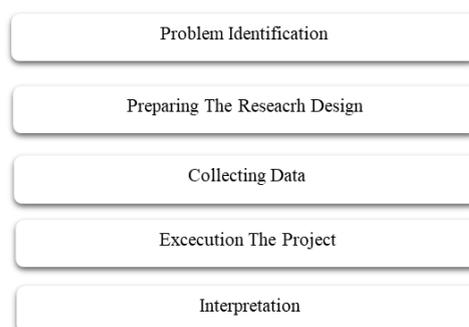
Terdapat penelitian terdahulu terkait prediksi harga laptop dengan algoritma yang sama seperti yang telah dilakukan oleh Syaiful Nur Wardani dan Nurmalitasari pada tahun 2023 dengan penelitian yang berjudul “Penerapan Regresi Linier Berganda Untuk Memprediksi Harga Laptop Dengan Menggunakan Software Python” yang memperoleh hasil keakuratan prediksi 0.347 dengan nilai MAE 26.84. Telaah ini dilakukan untuk menganalisa harga laptop yang menjadi variabel terikat dengan variabel bebas yang dapat memengaruhinya [2].

Adapun penelitian dengan algoritma yang sama lainnya dengan tingkat akurasi yang baik adalah penelitian prediksi harga mobil bekas telah dilakukan oleh Ernianti Hasibuan dan Aldian Karim pada tahun 2022 dengan data training 80% dan data testing 20% [5], penelitian analisis curah hujan telah dilakukan oleh Jesi Pebralia pada tahun 2022 dengan data training 80% dan data testing 20% [10], penelitian prediksi biaya asuransi kesehatan telah dilakukan oleh Muhammad Sholeh, dkk. pada tahun 2022 dengan data training 80% dan data testing 20% [11], penelitian prediksi harga rumah telah dilakukan oleh Andi Saiful, dkk. Pada tahun 2021 dengan data training 80% dan data testing 20% [12].

Observasi ini dilakukan dengan tujuan untuk menghasilkan *website* yang dapat digunakan untuk memprediksi harga laptop. Penelitian ini dapat membantu konsumen dalam membeli laptop pada waktu terbaik saat perkiraan menyatakan harganya sedang turun serta dapat membantu penjual untuk menetapkan strategi terkait pemberian harga bagi laptop yang dijualnya untuk dapat memaksimalkan keuntungan. Dari penelitian ini, dapat diketahui apa saja spesifikasi laptop yang dapat dikatakan sebagai faktor utama yang mampu mempengaruhi harga. Model pengembangan dalam penelitian ini juga dapat digunakan untuk model prediksi produk lainnya, tidak hanya untuk produk laptop. Adapun selain kontribusi penelitian ini yang dapat diberikan, penelitian ini juga mempunyai beberapa keterbatasan seperti akurasi yang kurang dapat dipicu dari ketidaktersediaan atau kurang lengkapnya data historis harga laptop yang terdapat dalam dataset, data tidak terbaharui atau tidak terkini sehingga prediksi yang dihasilkan mungkin saja kurang tepat.

II. METODE PENELITIAN

Tahapan penelitian terkait langkah-langkah detail prosedural yang berguna dalam proses penelitian ditunjukkan pada gambar 1 [13].



Gambar 1. Tahapan Penelitian
 Sumber: Dokumentasi Pribadi

Problem Identification

Pada tahapan ini dilakukan analisis terkait urgensi dari penelitian prediksi harga laptop menggunakan *machine learning*. Permasalahan yang menjadi rujukan dari penelitian ini, yaitu ketika seseorang ingin membeli sebuah laptop maka dibutuhkan aplikasi yang kompatibel untuk memberikan harga prediksi yang sesuai dengan spesifikasi yang diinginkan. Tujuan dari prediksi ini adalah untuk menemukan harga terdekat yang mungkin terjadi, tidak pasti tepat akan tetapi setidaknya dengan dilakukan prediksi, akan dapat memperkecil selisih dari harga sesungguhnya dengan hasil perkiraan.

Preparing The Research Design

Setelah menentukan masalah dan tujuan penelitian, selanjutnya, yaitu pemilihan metodologi dan penentuan variabel (spesifikasi laptop). Karena penelitian ini bertujuan untuk memberikan suatu prediksi maka dipilih metode yang memiliki kecocokan dengan tujuan penelitian, yaitu metode regresi linear berganda. Regresi linear adalah salah satu teknik yang paling umum diterapkan dalam analisis data untuk memahami hubungan antara variabel dependen atau Y (harga laptop) dan satu atau lebih variabel independen atau X (spesifikasi teknis laptop). Dengan mengumpulkan training data, kemudian *machine learning* akan menyelaraskan sebuah fungsi ke dalam data ini yang selanjutnya akan dipelajarilah Y sebagai fungsi dari X [14]. Rumus yang diterapkan ditunjukkan pada persamaan (1).

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (1)$$

Collecting Data

Tahapan awal yang akan dilakukan dalam melakukan observasi terkait prediksi adalah menetapkan apa yang akan menjadi bahasan utama dalam prediksi tersebut, topik prediksi dapat dipastikan di tahapan ini lalu langkah selanjutnya yang dapat dilakukan adalah pengambilan dataset yang sesuai dengan topik prediksi, dataset dapat diperoleh dari website penyedia berbagai dataset yang diketahui sebagai Kaggle.com. Bahan pembelajaran yang dipakai berupa dataset *laptop_gaming* milik Muhammad Fatur Febrianto di Kaggle.com [15]. yang kemudian akan berbentuk file excel yang akan diformat menjadi csv agar lebih mudah untuk terbaca saat ditarik kedalam bahasa pemrograman python melalui *Jupyter Notebook* [14]. Dataset yang berasal dari platform Kaggle.com berisi informasi tentang harga laptop beserta fitur-fitur seperti brand, tipe penyimpanan, ukuran layar dalam inch, warna, VGA, RAM, berat laptop dan lain-lain. Dataset *laptop_gaming* ditampilkan pada gambar 2.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Nama, Berat, Dimensi, Garansi, VGA, Tipe Penyimpanan, Processor, Penyimpanan, Ukuran Layar, Warna, RAM, Keyboard, Microsoft Office, Sistem Operasi, Brand, Tipe Layar, Harga																			
2	Razer Blade Advance 18	INTEL I9 13950HX	64GB DDR5	2TB SSD	RTX4090-16GB	18.0 QHD+	240HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4090, SSD,	Intel Core					
3	Razer Blade Advance 16	INTEL I9 13950HX	32GB DDR5	2TB SSD	RTX4090-16GB	16.0 QHD+	240HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4090, SSD,	Intel Core					
4	Razer Blade Advance 16	INTEL I9 13950HX	32GB DDR5	2TB SSD	RTX4090-16GB	16.0 QHD+	240HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4090, SSD,	Intel Core					
5	Razer Blade Advance 16	INTEL I9 13950HX	32GB DDR5	1TB SSD	RTX4070-8GB	16.0 QHD+	240HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4070, SSD,	Intel Core					
6	Razer Blade Advance 16	INTEL I9 13950HX	16GB DDR5	1TB SSD	RTX4070-8GB	16.0 QHD+	240HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4070, SSD,	Intel Core					
7	Razer Blade Advance 15	INTEL I7 13800H	32GB DDR5	1TB SSD	RTX3080TI-16GB	15.6 QHD	240HZ	100DCIP3	PKRGB	WIN11HOME, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3080TI, SSD,	Intel Core					
8	Razer Blade Advance 15	INTEL I7 12800H	32GB DDR5	1TB SSD	RTX3080TI-16GB	15.6 QHD	240HZ	RGB	WIN11HOME, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3080TI, SSD,	Intel Core						
9	Razer Blade Advance 15	INTEL I7 12800H	16GB DDR5	1TB SSD	RTX3070TI-8GB	15.6 QHD	240HZ	RGB	WIN11HOME, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3070TI, SSD,	Intel Core						
10	Razer Blade 14	AMD RYZEN 9 6900HX	16GB DDR5	1TB SSD	RTX3080TI-16GB	14.0 QHD	165HZ	RGB	WIN11HOME	QUARTZ, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3080TI, SSD,	AMD Ryzen 9,					
11	Razer Blade 14	AMD RYZEN 9 6900HX	16GB DDR5	1TB SSD	RTX3070TI-8GB	14.0 QHD	165HZ	RGB	WIN11HOME	QUARTZ, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3070TI, SSD,	AMD Ryzen 9,					
12	Razer Blade 14	AMD RYZEN 9 6900HX	16GB DDR5	1TB SSD	RTX3070TI-8GB	14.0 QHD	165HZ	RGB	WIN11HOME	MERCURY, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3070TI, SSD,	AMD Ryzen 9,					
13	Razer Blade 14	AMD RYZEN 9 6900HX	16GB DDR5	1TB SSD	RTX3070TI-8GB	14.0 QHD	165HZ	RGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	RAZER 1 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3070TI, SSD,	AMD Ryzen 9,					
14	MSI VECTOR GP68HX 13VH	INTEL I9 12900HX	32GB DDR5	2TB SSD	RTX4080-12GB	16.0 QHD+	IPS 240HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4080, SSD,	Intel Core i9, 2T					
15	MSI VECTOR GP68HX 12VH	INTEL I9 12900HX	16GB DDR5	1TB SSD	RTX4080-12GB	16.0 QHD+	IPS 240HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4080, SSD,	Intel Core i9, 1T					
16	MSI VECTOR GP66 12UGS	INTEL I7 12700H	16GB DDR5	1TB SSD	RTX3070-8GB	15.6 QHD	IPS 165HZ	PKRGB	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX3070, SSD,	Intel Core i7, 1TB, 16					
17	MSI TITAN GT77HX 13V1	INTEL I9 13980HX	64GB DDR5	4TB SSD	RTX4090-16GB	17.3 UHD	144HZ	PKRGB	MINILED	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4090, SSD,	Intel Core i9,				
18	MSI TITAN GT77HX 13V1	INTEL I9 13980HX	64GB DDR5	4TB SSD	RTX4080-12GB	17.3 UHD	144HZ	PKRGB	MINILED	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4080, SSD,	Intel Core i9,				
19	MSI TITAN GT77 13V1	INTEL I9 13980HX	64GB DDR5	4TB SSD	RTX4080-12GB	17.3 3&e	UHD	144HZ	PKRGB	MINILED	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4080, SSD,	Intel Core i9,			
20	MSI STEALTH 17M STUDIO A13VH	INTEL I9 13900H	64GB DDR5	4TB SSD	RTX4090-16GB	17.3 UHD	144HZ	PKRGB	MINILED	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4090, SSD,	Intel Core i9,				
21	MSI STEALTH 17 STUDIO A13VH	INTEL I9 13900H	64GB DDR5	4TB SSD	RTX4080-12GB	17.3 UHD	144HZ	PKRGB	MINILED	WIN11HOME	BLACK, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4080, SSD,	Intel Core i9,				
22	MSI STEALTH 16 A13VG	INTEL I7 13700H	32GB DDR5	2TB SSD	RTX4070-8GB	16.0 QHD+	240HZ	100SRGB	PKRGB	WIN11HOME	STAR BLUE, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4070, SSD,	Intel Core i7,				
23	MSI STEALTH 16 A13VG	INTEL I7 13700H	32GB DDR5	2TB SSD	RTX4070-8GB	16.0 QHD+	240HZ	100SRGB	PKRGB	WIN11HOME	STAR BLUE, 5 kg, 40 A	20 A	10 cm,	MSI 2 YEAR (SERVICE & SPAREPART),	NVIDIA RTX4070, SSD,	Intel Core i7,				

Gambar 2. Dataset yang Ditampilkan dalam Excel
Sumber: Kaggle.com

Execution The Project

Execution The Project merupakan proses penerapan metodologi penelitian yang telah dirancang sebelumnya. Terdapat beberapa proses yang dilakukan, yaitu:

1. Data cleaning

Library python yang digunakan untuk observasi ini adalah *pandas*, *seaborn*, *matplotlib* dan *numpy* [16].

Dataset yang sudah berbentuk file csv dimunculkan kedalam program python dengan *library* *pandas* untuk dideskripsikan serta dihitung total data dengan bantuan bahasa pemrogramman untuk keseluruhan laptop yang terinput dalam dataset. Gambar 3 menunjukkan bahwa dalam dataset *laptop gaming* untuk observasi ini terdapat total 646 data dan 17 kolom serta informasi mengenai tipe data yang ditunjukkan dalam gambar 4.

646 rows x 17 columns

Gambar 3. Tampilan Dataset
Sumber: Dokumentasi Pribadi

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 646 entries, 0 to 645
Data columns (total 17 columns):
#   Column                               Non-Null Count  Dtype
---  ---                               -
0   Nama                                 646 non-null    object
1   Berat                               646 non-null    object
2   Dimensi                             646 non-null    object
3   Garansi                             642 non-null    object
4   VGA                                  646 non-null    object
5   Tipe Penyimpanan                   646 non-null    object
6   Processor                           646 non-null    object
7   Penyimpanan                         646 non-null    object
8   Ukuran Layar                       646 non-null    object
9   Warna                               646 non-null    object
10  RAM                                  646 non-null    object
11  Keyboard                            227 non-null    object
12  Microsoft Office                    382 non-null    object
13  Sistem Operasi                      646 non-null    object
14  Brand                               646 non-null    object
15  Tipe Layar                          646 non-null    object
16  Harga                               646 non-null    object
dtypes: object(17)
memory usage: 85.9+ KB
```

Gambar 4. Informasi Awal Tipe Data
Sumber: Dokumentasi Pribadi

Dari informasi yang terlihat pada gambar, tahap selanjutnya yang dilakukan dalam *Jupyter Notebook* adalah pembersihan data. Dalam tahapan ini akan dilakukan penghapusan data yang terdapat duplikat dan penghapusan kolom yang dirasa kurang diperlukan [17]. Selengkapnya akan dilakukan perincian yang akan menerangkan beberapa perubahan setelah proses data cleaning dalam tabel 1.

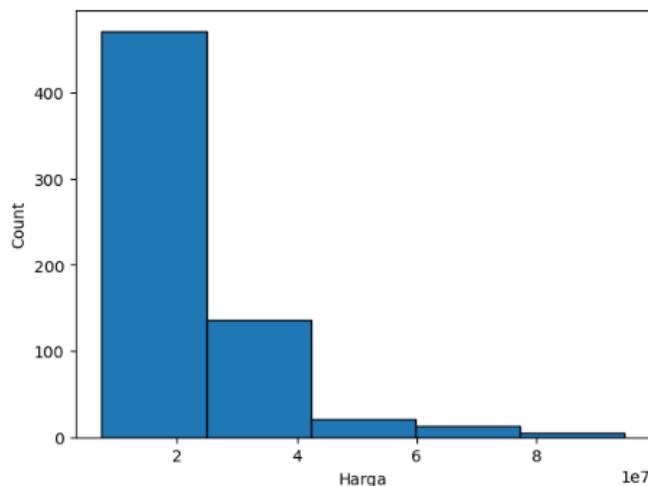
TABEL 1
PROSES PERUBAHAN DALAM DATA CLEANING

Kolom	Setelah Data Cleaning
Berat	Kolom telah dihapus setelah pengungkapan pada <i>Exploratory Data Analysis (EDA)</i>
Dimensi	Kolom telah dihapus
Garansi	Kolom telah dihapus
VGA	Kolom telah dihapus setelah pembentukan kolom baru pada proses Feature Engineering

Kolom	Setelah Data Cleaning
Tipe Penyimpanan	Kolom telah dihapus setelah pembentukan kolom baru pada proses Feature Engineering
Penyimpanan	Kolom telah dihapus setelah pembentukan kolom baru pada proses Feature Engineering
Ukuran Layar	'Inch' pada data ukuran layar dihapus dan tipe data telah diubah dari object menjadi float64
Ram	'GB' pada data ram telah dihapus dan tipe data telah diubah dari object menjadi int32
Keyboard	Kolom telah dihapus
Microsoft Office	Kolom telah dihapus
Sistem Operasi	Kolom telah dihapus setelah pembentukan kolom baru pada proses Feature Engineering
Tipe Layar	Kolom telah dihapus setelah pembentukan kolom baru pada proses Feature Engineering
Harga	'Rp' dan '.' pada data harga telah dihapus dan tipe data telah diubah dari object menjadi float64

2. Exploratory Data Analysis

Exploratory Data Analysis (EDA) merupakan langkah awal yang dapat dilakukan guna memahami pengungkapan pada data yang tersaji. Data yang telah disajikan terlebih dahulu melewati dasar analisis yang dilakukan untuk mengetahui karakteristik ataupun komponen yang ada dalam dataset agar dapat dibentuk model prediksi yang baik setelah dilakukan analisis statistik. Pengungkapan yang dapat dilakukan atas perilaku pembeli jika dilihat dari harga laptop ditunjukkan dalam gambar 5.

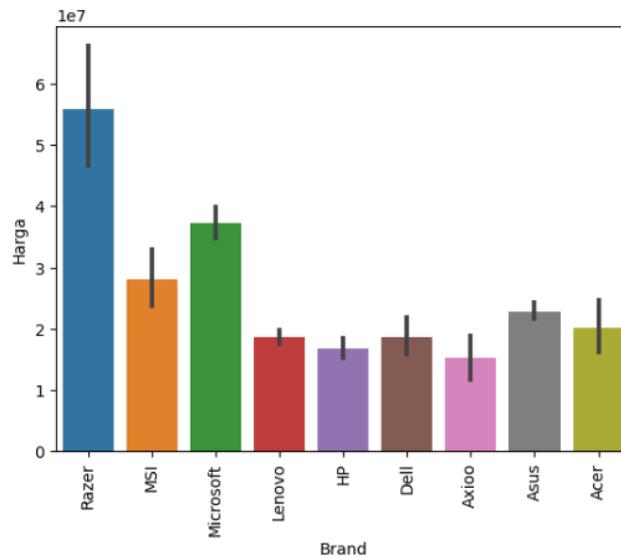


Gambar 5. Histogram Harga Laptop
Sumber: Dokumentasi Pribadi

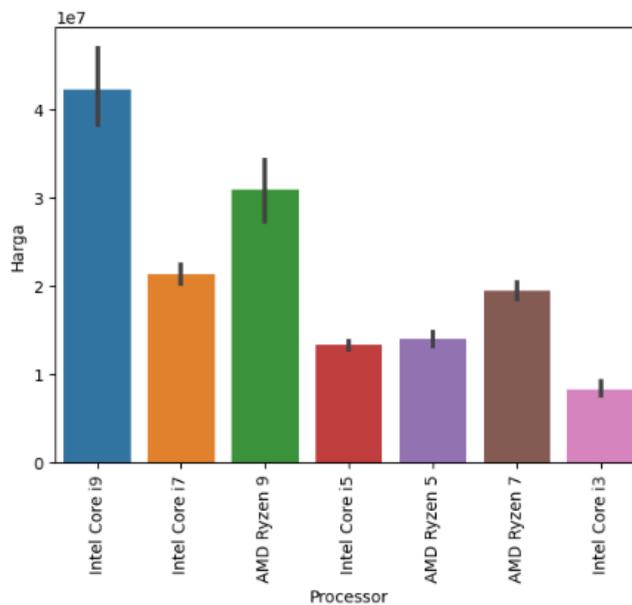
Berdasarkan hasil yang telah ditampilkan, terdapat laptop dengan harga yang lebih rendah atau murah dan beberapa laptop dengan harga yang sangat tinggi. Penggunaan istilah count yang terdapat dalam histogram harga laptop untuk mewakili jumlah pembeli dapat menunjukkan bahwa pembeli laptop dengan harga rendah atau murah lebih banyak jika dibandingkan pembeli laptop dengan harga tinggi. Dilakukan pengungkapan atas rata-rata harga jika didasarkan pada merek laptop ditunjukkan dalam gambar 6.

Hasil analisis berbagai merek laptop dengan harga yang beragam di Indonesia memperlihatkan harga laptop Razer melambung tinggi berbeda dari merek laptop lain. Merek Microsoft merupakan merek laptop dengan harga tertinggi kedua, lalu merek laptop dengan harga yang sangat terjangkau ada di merek Axioo. Dilakukan pengungkapan atas rata-rata harga jika didasarkan pada processor ditunjukkan dalam gambar 7.

Harga rata-rata dari berbagai processor laptop didasari oleh pengelompokkan processor laptop yang memberikan performa yang berbeda-beda. seperti *Intel Core i9* yang termasuk dalam kategori processor kelas atas dengan performa paling baik yang dimiliki intel untuk para pengguna yang memerlukannya. *AMD Ryzen 9* yang turut menawarkan performa terbaik untuk pengguna yang memerlukan untuk mengerjakan tugas-tugas berat.

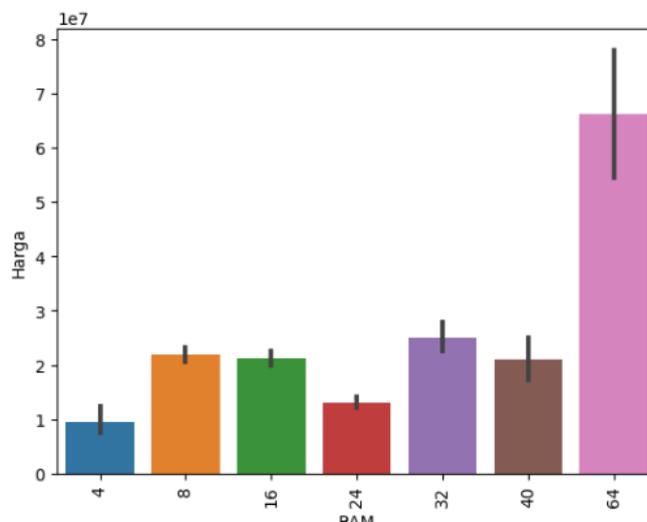


Gambar 6. Grafik Rata-rata Harga Laptop di Indonesia Berdasar pada Merek
Sumber: Dokumentasi Pribadi



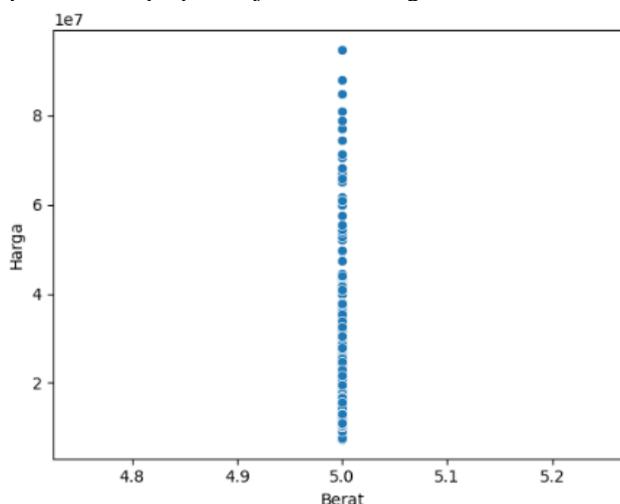
Gambar 7. Grafik Rata-rata Harga Berdasar pada Processor Laptop
Sumber: Dokumentasi Pribadi

Selanjutnya ada *Intel Core i7* yang berada dibawah *Intel Core i9* tetap mampu memberikan performa terbaik yang cukup bagi pengguna, *AMD Ryzen 7* yang berada dikelas menengah hingga atas yang dimiliki *AMD* menawarkan performa yang sebanding dengan harga, *AMD Ryzen 5* menawarkan performa yang cukup baik untuk mengerjakan tugas ringan dengan harga yang lebih terjangkau dari *AMD Ryzen 9* dan *AMD Ryzen 7*, *Intel Core i5* juga turut memberikan performa yang sebanding dengan harga yang dikenakan dan yang terakhir ada *Intel Core i3* yang menjadi prosesor yang ramah bagi pemula keluaran intel yang menawarkan performa dasar dengan harga paling rendah yang ada dalam dataset ini. Dilakukan pengungkapan atas rata-rata harga jika didasarkan pada ram laptop ditunjukkan dalam gambar 8.



Gambar 8. Grafik Rata-rata Harga Berdasar pada Ram Laptop
Sumber: Dokumentasi Pribadi

Berdasarkan gambar 8, didapatkan hasil analisis harga yang mengalami kenaikan dan penurunan dari 4GB hingga 64GB. Rerata harga laptop tertinggi hingga terendah jika diurutkan, yaitu laptop dengan ram 64GB, 32GB, 8GB, 16GB, 40GB, 24 GB dan 4GB. Dengan ini dapat ditarik simpulan bahwa semakin bertambahnya Ram tidak menjamin kenaikan harga laptop. Dapat diartikan bahwa terdapat korelasi yang rendah antara ram dan harga laptop. Terdapat faktor lainnya yang dapat berpengaruh terhadap harga seperti laptop dengan processor yang lebih cepat serta kapasitas penyimpanan yang besar dengan SSD, desain yang lebih premium atau faktor lainnya yang dapat memengaruhi harga. Dilakukan pengungkapan atas rata-rata harga jika didasarkan pada berat laptop ditunjukkan dalam gambar 9.



Gambar 9. Scatterplot Rata-rata Harga Berdasar pada Berat Laptop
Sumber: Dokumentasi Pribadi

Jika dilihat dari Scatterplot yang ditunjukkan dalam gambar 9, analisis yang didapatkan adalah dalam mencari hubungan linear antar variabel berat laptop terhadap harga laptop, diperoleh hasil variabel berat laptop tidak memiliki hubungan linear terhadap harga atau dapat dikatakan berat laptop tidak memberikan pengaruh terhadap harga laptop. Hal ini dikarenakan seluruh laptop yang terdapat dalam dataset memiliki nilai berat yang sama, yaitu 5 kg. Berat yang sama antar laptop yang terdata dapat memvalidasi bahwa variabel berat tidak memberikan pengaruh terhadap harga laptop, melainkan terdapat faktor lainnya yang dapat menjadi faktor terpengaruhnya harga laptop.

3. Feature Engineering

Dalam tahapan ini memungkinkan untuk membentuk lebih banyak fitur dari dataset sebelumnya [18]. Contohnya dalam kolom Tipe Penyimpanan dalam tahap ini akan dipisahkan ke beberapa kolom baru yang berupa HDD dan SSD. Beberapa perubahan *feature engineering* yang terjadi terlihat pada tabel 2.

TABEL 2
PROSES PERUBAHAN DALAM FEATURE ENGINEERING

Kolom	Setelah Data Cleaning
Nama	Pembentukan kolom baru dengan informasi IPS
VGA	Pembentukan kolom baru dengan informasi VGA Brand
Tipe Penyimpanan	Pembentukan kolom baru dengan informasi HDD dan SSD
Penyimpanan	Dipilah berdasarkan jenisnya lalu membentuk kolom baru dengan informasi HDD dan SSD
Sistem Operasi	Pembentukan kolom baru dengan informasi OS_Windows 11 Pro, OS_Windows 11 Home, OS_Windows 10 Pro dan OS_Windows 10 Home
Tipe Layar	Pembentukan kolom baru dengan informasi Touchscreen

4. Pemisahan data

Pre-processing dilakukan terhadap dataset untuk menyiapkan model yang baik, dalam tahapan tersebut telah dilakukan pembersihan data dengan menghapus kolom yang dirasa kurang diperlukan dan data yang banyak terdapat duplikat lalu dilakukan pembentukan lebih banyak fitur yang dikembangkan dari dataset sebelumnya. Dataset yang telah melewati proses pembersihan dan penambahan fitur dianggap telah siap untuk selanjutnya dilakukan pembentukan model. Dataset setelah tahap *pre-processing* ditunjukkan pada gambar 10.

	Nama	Processor	Ukuran Layar	Warna	RAM	Sistem Operasi	Brand	Harga	Touchscreen	IPS	SSD	HDD	VGA Brand
0	RAZER BLADE ADVANCE 18 INTEL I9 13950HX 64GB D...	Intel Core i9	18.0	Black	16	Windows 11 Home	Razer	94799000.0	0	0	2000.0	0.0	NVIDIA
1	RAZER BLADE ADVANCE 18 INTEL I9 13950HX 32GB D...	Intel Core i9	18.0	Black	16	Windows 11 Home	Razer	84899000.0	0	0	2000.0	0.0	NVIDIA
2	RAZER BLADE ADVANCE 16 INTEL I9 13950HX 32GB D...	Intel Core i9	16.0	Black	16	Windows 11 Home	Razer	66999000.0	0	0	2000.0	0.0	NVIDIA
3	RAZER BLADE ADVANCE 16 INTEL I9 13950HX 32GB D...	Intel Core i9	16.0	Black	16	Windows 11 Home	Razer	66999000.0	0	0	1000.0	0.0	NVIDIA
4	RAZER BLADE ADVANCE 16 INTEL I9 13950HX 16GB D...	Intel Core i9	16.0	Black	8	Windows 11 Home	Razer	51999000.0	0	0	1000.0	0.0	NVIDIA
...
641	ACER ASPIRE 5 SLIM A514 55G INTEL I7 1255U 8GB...	Intel Core i7	14.0	Blue	8	Windows 11 Home	Acer	13199000.0	0	1	512.0	0.0	NVIDIA
642	ACER ASPIRE 5 SLIM A514 55G INTEL I5 1235U 8GB...	Intel Core i5	14.0	Red	8	Windows 11 Home	Acer	10999000.0	0	1	512.0	0.0	NVIDIA
643	ACER ASPIRE 5 SLIM A514 55G INTEL I5 1235U 8GB...	Intel Core i5	14.0	Grey	8	Windows 11 Home	Acer	10999000.0	0	1	512.0	0.0	NVIDIA
644	ACER ASPIRE 5 SLIM A514 55G INTEL I5 1235U 8GB...	Intel Core i5	14.0	Gold	8	Windows 11 Home	Acer	10999000.0	0	1	512.0	0.0	NVIDIA
645	ACER ASPIRE 5 SLIM A514 55G INTEL I5 1235U 8GB...	Intel Core i5	14.0	Blue	8	Windows 11 Home	Acer	10999000.0	0	1	512.0	0.0	NVIDIA

646 rows x 13 columns

Gambar 10. Tampilan Dataset Setelah Proses Data Cleaning dan Feature Engineering
Sumber: Dokumentasi Pribadi

Pembentukan model dapat dilakukan setelah data melalui berbagai proses sebelumnya guna mempersiapkan data. Setelah data telah siap, kemudian akan dilakukan pemisahan data menjadi dua bagian dengan perbandingan 85% : 15%, yaitu 85% menjadi data training atau data pelatihan dan 15% menjadi data test atau data pengujian seperti yang terlihat dalam gambar 11 [19].

```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test = train_test_split(X,y,test_size=0.15,random_state=2)
```

Gambar 11. Pemisahan Data Test dan Data Training
Sumber: Dokumentasi Pribadi

Terdapat total 549 data pelatihan yang akan digunakan dalam membentuk model dari total keseluruhan data 646 seperti yang tercantum dalam gambar 12. Selisih dengan total 97 merupakan data pengujian.

	Nama	Processor	Ukuran Layar	Warna	RAM	Sistem Operasi	Brand	Harga	Touchscreen	IPS	SSD	HDD	VGA Brand
505	ASUS ROG ZEPHYRUS G14 GA402RK R9X7B7GO AMD RYZ...	AMD Ryzen 9	14.0	White	8	Windows 10 Home	Asus	29499000.0	0	1	1000.0	0.0	NVIDIA
13	MSI VECTOR GP68HX 12VH INTEL I9 1200HX 16GB 1...	Intel Core i9	16.0	Black	16	Windows 11 Home	MSI	36499000.0	0	1	1000.0	0.0	NVIDIA
221	LENOVO IDEAPAD GAMING 3i 15 INTEL I5 12500H 8G...	Intel Core i5	16.0	Grey	8	Windows 11 Home	Lenovo	14299000.0	0	0	1000.0	0.0	NVIDIA
453	ASUS TUF DASH F15 FX517ZC I53588TO INTEL I5 12...	Intel Core i5	16.0	Black	16	Windows 11 Home	Asus	15699000.0	0	1	1000.0	0.0	NVIDIA
573	ASUS ROG G15 G513RM AMD RYZEN 7 6800H 16GB DOR...	AMD Ryzen 7	16.0	Grey	16	Windows 11 Home	Asus	21299000.0	0	1	1000.0	0.0	NVIDIA
...
534	ASUS ROG STRIX G17 G713RC R73587GO AMD RYZEN 7...	AMD Ryzen 7	17.0	Grey	16	Windows 11 Home	Asus	18599000.0	0	1	512.0	0.0	NVIDIA
584	ASUS ROG FLOW Z13 GZ301ZC INTEL I7 12700H 19GB...	Intel Core i7	13.0	Black	16	Windows 11 Home	Asus	21499000.0	1	1	1000.0	0.0	NVIDIA
493	ASUS ROG ZEPHYRUS G16 GU803VV I946C6GO INTEL I...	Intel Core i9	16.0	Grey	16	Windows 11 Home	Asus	31999000.0	0	0	1000.0	0.0	NVIDIA
527	ASUS ROG STRIX SCAR G17 G733PZ R948C8TO AMD RY...	AMD Ryzen 9	17.0	Black	32	Windows 11 Home	Asus	60899000.0	0	1	2000.0	0.0	NVIDIA
168	LENOVO LEGION 5i PRO 15 INTEL I7 12700H 64GB D...	Intel Core i7	16.0	Grey	8	Windows 11 Home	Lenovo	25699000.0	0	0	2000.0	0.0	NVIDIA

549 rows x 13 columns

Gambar 12. Data Training
Sumber: Dokumentasi Pribadi

5. Pengujian Model

Setelah dilakukan pemisahan data training, langkah selanjutnya adalah menerapkan algoritma Regresi Linear dalam permodelan menggunakan data training yang telah diperoleh lalu dilanjutkan dengan uji akurasi model prediksi menggunakan R^2 dan *Mean Absolute Error* seperti yang terlihat pada gambar 13.

```
# Tes Akurasi dengan RSquare dan MAE (Mean Absolute Error)
print('R2 score',r2_score(y_test,y_pred))
print('MAE',mean_absolute_error(y_test,y_pred))
```

Gambar 13. Pengujian Model dengan R^2 dan MAE
Sumber: Dokumentasi Pribadi

Pengujian model yang dilakukan memperoleh nilai akurasi R^2 sebesar 0.68 dengan MAE senilai 0.18 atas evaluasi linear yang telah dilakukan. Hasil dari pengujian terdapat pada gambar 14.

R2 score 0.680324168620956
MAE 0.18495815550051786

Gambar 14. Nilai R^2 dan MAE
Sumber: Dokumentasi Pribadi

Perolehan nilai tersebut mengindikasikan bahwa model memiliki akurasi yang baik dengan nilai R^2 sebesar 0.68 dan MAE sebesar 0.18. Ini berarti bahwa model mampu menjelaskan 68% variabilitas dalam harga laptop dan rata-rata kesalahan prediksi sebesar 18%.

6. Pembentukan model

Data yang dipakai dalam proses ini adalah data training karena berfungsi untuk melatih algoritma. Data lain, yaitu data testing akan digunakan untuk menguji algoritma terlatih ketika terindikasi pembaruan data yang sebelumnya belum tampak. Proses pembentukan model dengan algoritma regresi linear ditunjukkan pada gambar 15.

```
# Linear Regression
step1 = ColumnTransformer(transformers=[
    ('col_tnf', OneHotEncoder(sparse_output=False, drop='first'), [0,1,7,10,11])
], remainder='passthrough')

step2 = LinearRegression()

pipe = Pipeline([
    ('step1', step1),
    ('step2', step2)
])

pipe.fit(X_train, y_train)

y_pred = pipe.predict(X_test)
```

Gambar 15. Proses Pembentukan Model Regresi Linear
 Sumber: Dokumentasi Pribadi

Dalam penelitian ini, digunakan berbagai algoritma regresi linear yang sering digunakan dalam pengembangan prediksi di bidang machine learning. Beberapa algoritma yang dipakai termasuk *KNeighborsRegressor*, *RandomForestRegressor*, dan *GradientBoostingRegressor*. Proses pembentukan model dengan algoritma regresi linear *KNeighborsRegressor* ditunjukkan pada gambar 16.

```
step2 = KNeighborsRegressor(n_neighbors=3)

pipe = Pipeline([
    ('step1', step1),
    ('step2', step2)
])

pipe.fit(X_train, y_train)

y_pred = pipe.predict(X_test)
```

Gambar 16. Proses Pembuatan Model KneighborsRegressor
 Sumber: Dokumentasi Pribadi

KNeighborsRegressor merupakan implementasi khusus dari algoritma *K-Nearest Neighbors* yang digunakan untuk regresi [20]. *KNeighborsRegressor* menghitung rata-rata nilai dari sejumlah titik data terdekat, yang ditentukan oleh parameter *k* untuk menghasilkan prediksi. Jika parameter *k* diatur menjadi 3 (*n_neighbors=3*), maka model akan mempertimbangkan tiga tetangga terdekat dari titik data yang ingin diprediksi untuk menghitung nilai rata-rata mereka dan menghasilkan prediksi akhir. Pendekatan ini memungkinkan model untuk mempertimbangkan kedekatan dan kemiripan antara titik data dalam menentukan hasil prediksi, sehingga bisa menangkap pola dan tren yang ada dalam data dengan lebih baik.

```
step2 = RandomForestRegressor(n_estimators=100,
                              random_state=3,
                              max_samples=0.5,
                              max_features=0.75,
                              max_depth=15)

pipe = Pipeline([
    ('step1', step1),
    ('step2', step2)
])

pipe.fit(X_train, y_train)

y_pred = pipe.predict(X_test)
```

Gambar 17. Proses Pembuatan Model RandomForestRegressor
 Sumber: Dokumentasi Pribadi

Gambar 17 menunjukkan proses pembentukan model dengan algoritma regresi linear *RandomForestRegressor*. *Random Forest* merupakan salah satu bagian dari *machine learning* dengan konsep *supervised* dalam pembentukan kelas [21]. Pendekatan ini terdiri dari sejumlah besar pohon keputusan (decision trees) dan akan mempelajari mengenai pola dari dataset [22]. Setiap pohon keputusan secara acak

belajar akan mengenali pola dari dataset yang diberikan, sehingga menghasilkan model yang kuat dan mampu menangani data yang kompleks serta beragam.

GradientBoostingRegressor merupakan modifikasi dari algoritma *ensemble learning* yang dikenal sebagai *Gradient Boosting Machine (GBM)*. GBM sendiri, yaitu teknik pembelajaran mesin yang ditujukan untuk isu-isu klasifikasi dan menghasilkan prediksi dalam format *decision tree* [23]. Dalam *Gradient Boosting Regression*, model bekerja dengan membangun serangkaian pohon keputusan yang lemah secara bertahap di mana setiap pohon memperbaiki kesalahan prediksi dari pohon sebelumnya [24]. Proses pembentukan model dengan algoritma regresi linear *GradientBoostingRegressor* ditunjukkan pada gambar 18.

```
step2 = GradientBoostingRegressor(n_estimators=500)

pipe = Pipeline([
    ('step1', step1),
    ('step2', step2)
])

pipe.fit(X_train, y_train)

y_pred = pipe.predict(X_test)
```

Gambar 18. Proses Pembuatan Model *GradientBoostingRegressor*
Sumber: Dokumentasi Pribadi

7. Pengembangan website

Setelah proses pembentukan dan pelatihan model yang dilakukan menggunakan *Jupyter Notebook* selesai dan model telah siap untuk dijadikan sebagai dasar prediksi, langkah selanjutnya adalah proses pengembangan *website* menggunakan *Streamlit* dengan bantuan *Visual Studio Code*. Dalam mengembangkan aplikasi web Untuk memulai pengembangan aplikasi di *Visual Studio Code (VS Code)*, pastikan semua file yang terkait dengan proyek berada dalam satu folder yang terorganisir. Buka *VS Code* dan pilih opsi "Open Folder" untuk memuat seluruh folder proyek ke dalam editor. Setelah folder proyek terbuka, langkah berikutnya adalah membuat file Python baru yang akan menjadi inti dari aplikasi web. Buatlah file dengan nama "app.py" di dalam folder proyek. File ini akan berisi seluruh kode yang diperlukan untuk mengembangkan antarmuka pengguna serta mengintegrasikan model prediksi yang telah dilatih sebelumnya. Dalam "app.py", termuat kode *Streamlit* yang memungkinkan pengguna untuk berinteraksi dengan model prediksi melalui antarmuka web yang intuitif. Pengembangan aplikasi dengan bantuan *VS Code* diperlihatkan dalam gambar 19.

```
app.py > ...
1 import streamlit as st
2 import pickle
3 import numpy as np
4
5 st.write("<h1><b>Welcome!</b></h1>", unsafe_allow_html=True)
6 st.write("Kindly Select the Specifications for the Laptop You Desire")
7
8 # import the model
9 pipe = pickle.load(open('pipe.pkl', 'rb'))
10 df = pickle.load(open('df.pkl', 'rb'))
11
12 st.title("Laptop Price Predictor")
```

Gambar 19. Pengembangan Aplikasi di *Visual Studio Code*
Sumber: Dokumentasi Pribadi

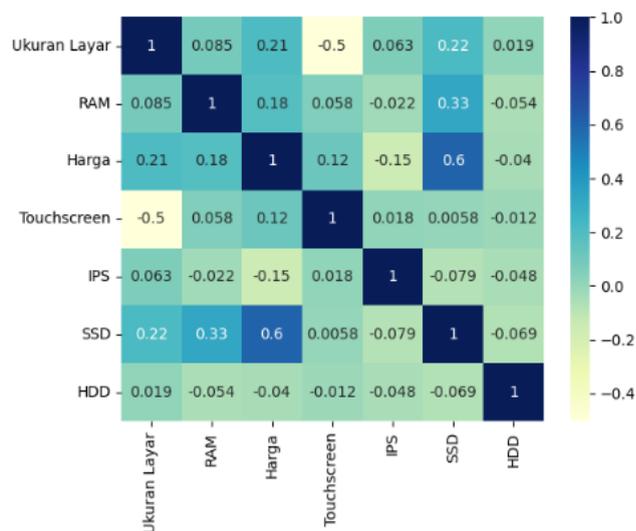
Interpretation

Interpretation merupakan proses analisis hasil dari model prediksi yang telah dibangun setelah melewati berbagai tahapan sebelumnya. Hasil prediksi tidak pasti tepat dengan harga yang sesungguhnya terjadi, akan tetapi prediksi dapat memberikan gambaran harga sedekat mungkin dan memperkecil selisih yang ada. Proses *interpretation* dapat dilanjutkan dengan analisis hasil dari model prediksi dengan harga aktual untuk dibandingkan agar dapat disimpulkan seberapa akurat aplikasi tersebut dalam memprediksi harga laptop.

III. HASIL DAN PEMBAHASAN

Uji Korelasi

Dalam observasi ini, terdapat variabel-variabel independen yang memungkinkan untuk dapat memberikan pengaruh yang signifikan terhadap variabel dependen berupa harga laptop. Analisis korelasi menggunakan heatmap untuk dapat memperlihatkan dengan jelas setiap korelasi yang dimiliki variabel independen terhadap variabel dependen. Analisis korelasi menggunakan heatmap ditunjukkan pada gambar 20.



Gambar 20. Heatmap Korelasi Variabel

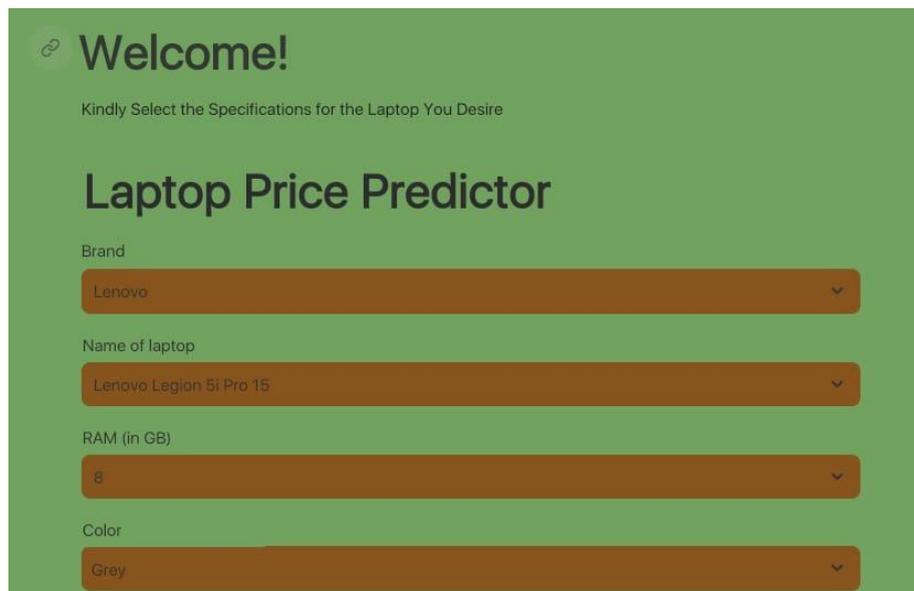
Sumber: Dokumentasi Pribadi

Hasil dari Heatmap digunakan untuk melihat korelasi variabel-variabel yang ada dalam observasi ini terhadap variabel terikatnya, yaitu harga laptop. Variabel Ukuran Layar memiliki nilai 0.21 yang artinya terdapat korelasi yang lemah terhadap harga laptop. Variabel Ram memiliki nilai 0.18 yang artinya terdapat korelasi yang sangat lemah terhadap harga laptop. Variabel Touchscreen memiliki nilai 0.12 yang artinya terdapat korelasi negatif yang sangat lemah terhadap harga laptop. Variabel Ips memiliki nilai -0.15 yang artinya lemah atau tidak ada korelasi terhadap harga laptop. Variabel Ssd memiliki nilai 0.6 yang artinya terdapat korelasi yang kuat terhadap harga laptop dan yang terakhir variabel HDD memiliki nilai -0.04 yang artinya sangat lemah atau hampir tidak ada korelasi negatif terhadap harga laptop.

Tampilan User Interface

Setelah proses pengembangan aplikasi selesai, langkah selanjutnya adalah menjalankan aplikasi *Streamlit* dari terminal di *Visual Studio Code* dengan menggunakan perintah “*streamlit run app.py*”. Dengan melakukan hal ini, aplikasi akan mulai berjalan secara lokal di komputer pengguna dan secara otomatis membuka jendela browser web dengan aplikasi yang telah dikembangkan. Pengguna dapat langsung mulai menggunakan aplikasi dengan memasukkan spesifikasi laptop yang mereka cari melalui antarmuka yang telah disediakan. Tampilan user interface terdapat pada gambar 21.

Streamlit merupakan *platform* yang sangat berguna untuk mengembangkan antarmuka pengguna yang interaktif. Dengan menggunakan *Streamlit*, pengguna dapat dengan mudah memasukkan spesifikasi laptop yang mereka cari, seperti merk, warna, tipe prosesor, RAM, penyimpanan, ukuran layar, dan parameter lainnya melalui widget input yang disediakan. Input dari pengguna kemudian diproses secara otomatis dan digunakan sebagai masukan untuk model prediksi yang telah disiapkan sebelumnya. Model ini akan melakukan perhitungan berdasarkan input yang diberikan dan menghasilkan prediksi harga laptop. Hasil prediksi ini kemudian ditampilkan kembali kepada pengguna melalui antarmuka *Streamlit*, yang memungkinkan pengguna untuk melihat perkiraan harga laptop berdasarkan spesifikasi yang mereka inginkan dengan cepat dan mudah.



Gambar 21. User Interface
Sumber: Dokumentasi Pribadi

Setelah website dapat diakses, langkah berikutnya adalah melakukan pengujian untuk memastikan bahwa semua aspek aplikasi berfungsi dengan baik. Pengujian situs web ini dilakukan dalam dua tahap, yaitu uji *interface* dan uji prediksi.

Pengujian Interface

Pengujian interface dilakukan dengan menggunakan metode *black box testing*, yaitu suatu pendekatan di mana aplikasi atau sistem diuji dari luar tanpa memperhatikan struktur internal atau implementasi kode yang ada di dalamnya dan hanya berfokus pada input yang diberikan dan output yang dihasilkan. Pendekatan ini fokus pada fungsi atau perilaku yang diharapkan dari sistem, tanpa memperhatikan bagaimana sistem mencapai hasil tersebut. Pengujian interface dilakukan oleh pihak yang tidak terlibat langsung dalam proses pembuatan model machine learning dengan tujuan untuk memvalidasi fungsionalitas website dari sudut pandang pengguna. Pengujian dilakukan oleh mahasiswa kelas 4D, di mana masing-masing mahasiswa diberikan arahan yang sama untuk menjalankan skenario pengujian *user interface*. Hasil dari pengujian tersebut disajikan dalam tabel 3.

TABEL 3
SKENARIO PENGUJIAN INTERFACE

Skenario Pengujian	Hasil
Menguji fungsionalitas <i>select box</i> brand, nama laptop, RAM, warna, touchscreen, IPS, screen resolution, CPU, HDD, SSD, VGA, dan OS.	Ketika pengguna mengklik atau menyentuh select box, daftar opsi akan muncul dan dapat dipilih.
Menguji fungsionalitas <i>number input</i> untuk opsi ukuran layar (dalam inch).	Jika input berupa bilangan positif, maka sistem akan menampilkan hasil prediksi. Sedangkan jika input berupa bilangan negatif atau 0 maka sistem akan menampilkan sebuah pesan eror.
Menguji fungsionalitas tombol prediksi harga.	Tombol prediksi harga responsif terhadap interaksi pengguna.

Pengujian Prediksi

Sementara itu, uji prediksi bertujuan untuk memastikan bahwa hasil prediksi yang diberikan oleh model sesuai dengan nilai input yang diberikan oleh pengguna. Dalam tahap ini, pengujian dilakukan untuk memverifikasi bahwa prediksi yang dihasilkan konsisten dengan tingkat akurasi yang ditemukan selama uji model regresi. Hasil dari pengujian prediksi diperlihatkan dalam gambar 22.



Gambar 22. Hasil Prediksi Harga
Sumber: Dokumentasi Pribadi

Kinerja Aplikasi

Salah satu cara untuk memverifikasi hasil prediksi adalah dengan membandingkan hasil prediksi dari aplikasi dengan nilai aktual atau data yang sudah diketahui hasilnya. Pengujian ini juga melibatkan penggunaan data input realistis untuk memastikan bahwa model dapat memberikan prediksi yang masuk akal. Perbandingan antara harga asli laptop dengan harga prediksi yang dihasilkan diuraikan dalam tabel 4.

TABEL 4
PERBANDINGAN HARGA ASLI DAN HARGA PREDIKSI

Merek	Tipe	RAM	Harga Asli	Harga Prediksi
Razer	Blade 16	32	Rp 66.999.000	Rp 61.497.200
Lenovo	Legion Slim 5	16	Rp25.299.000	Rp 28.550.529
HP	Victus 15	32	Rp 15.099.000	Rp 14.889.065
Asus	Vivobook Pro 15	16	Rp 25.399.000	Rp 29.199.404
Acer	Nitro 5	16	Rp13.899.000	Rp 12.153.265

Dari tabel diatas dapat dilihat bahwa hasil prediksi cenderung mengikuti pola garis linier, dengan perbedaan antara harga asli dan harga prediksi yang relatif kecil. Fenomena ini mengindikasikan bahwa model prediksi yang digunakan mampu menangkap pola-pola yang terkandung dalam data dengan baik, sehingga memberikan prediksi yang mendekati garis linier. Dengan demikian, hasil ini memberikan bukti yang kuat bahwa model prediksi yang digunakan mampu memberikan prediksi dengan tingkat akurasi yang memuaskan.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi harga laptop yang akurat menggunakan metode regresi linier. Implementasi aplikasi berbasis web dengan Streamlit menunjukkan bahwa aplikasi ini *user-friendly*, efisien, dan dapat diandalkan untuk memberikan prediksi harga laptop secara *real-time*. Hasil evaluasi memperlihatkan bahwa model mempunyai tingkat ketepatan atau akurasi yang tinggi dengan nilai R^2 sebesar 0.68 dan *Mean Absolute Error* (MAE) sebesar 0.18. Validasi yang dilakukan memastikan bahwa model memiliki performa yang konsisten dan mampu menangani data baru dengan baik. Penelitian ini mengindikasikan bahwa integrasi antara machine learning dan teknologi web dapat menghasilkan alat yang praktis dan berguna bagi pengguna dalam berbagai konteks, khususnya dalam prediksi harga laptop. Untuk pengembangan lebih lanjut terkait topik ini disarankan agar memperluas validasi model dengan menggunakan dataset eksternal yang berbeda untuk memastikan generalisasi yang lebih baik dari model serta melakukan iterasi lebih lanjut pada antarmuka pengguna aplikasi web untuk meningkatkan pengalaman pengguna, termasuk visualisasi tambahan, petunjuk penggunaan, dan fitur interaktif lainnya.

DAFTAR PUSTAKA

- [1] M. R. A. Annas, "PREDIKSI HARGA LAPTOP DENGAN METODE SINGLE EXPONENTIAL SMOOTHING BERBASIS WEB," eprints.unisla.ac.id, Feb. 02, 2021. <http://eprints.unisla.ac.id/53/> (accessed May 26, 2024).
- [2] S. N. Wardani and N. Nurmalitasari, "Penerapan Regresi Linier Berganda Untuk Memprediksi Harga Laptop Dengan Menggunakan Software Python," Prosiding Seminar Nasional Teknologi Informasi dan Bisnis, pp. 377–382, Jul. 2023, Accessed: May 26, 2024. [Online]. Available: <http://ojs.uib.ac.id/index.php/Senatib/article/view/3211>
- [3] S. R. Harahap, "Faktor-faktor yang memengaruhi keputusan konsumen dalam pembelian laptop Merek Asus : studi kasus pada Mahasiswa Fakultas Ekonomi dan Bisnis Islam IAIN Padangsidempuan.," etd.uinsyahada.ac.id, Feb. 14, 2020. <https://etd.uinsyahada.ac.id/3537/> (accessed May 29, 2024).

- [4] T. Adiando, Y. R. Nasution, and A. H. Lubis, "Implementasi Data Mining untuk Prediksi Penjualan Produk Mayora Menggunakan Metode FP-Growth," repository.uinsu.ac.id, 2023. <http://repository.uinsu.ac.id/20362> (accessed May 29, 2024).
- [5] E. Hasibuan and A. Karim, "Implementasi Machine Learning untuk Prediksi Harga Mobil Bekas dengan Algoritma Regresi Linear Berbasis Web," *Jurnal Ilmiah KOMPUTASI*, vol. XXI(4), pp. 595-602, 2022.
- [6] Ibnu Daqiqil Id, *MACHINE LEARNING : Teori, Studi Kasus dan Implementasi Menggunakan Python*. Unri Press.
- [7] R. P. Sari, "Apa itu Machine Learning? Pengertian dan Contohnya," *Cloud Computing Indonesia*. <http://cloudcomputing.id/pengetahuan-dasar/apa-itu-machine-learning> (accessed May 26, 2024).
- [8] W. Bhirawa, "Proses Pengolahan Data Dari Model Persamaan Regresi Dengan Menggunakan Statistical Product and Service Solution (SPSS)." Available: <https://journal.universitassuryadarma.ac.id/index.php/jmm/article/viewFile/528/494>
- [9] D. A. Rhamadhani and E. E. D. Saputra, "Analisa Model Machine Learning dalam Memprediksi Laju Produksi Sumur Migas 15/9-F-14H," *Journal of Sustainable Energy Development*, vol. 1, no. 1, pp. 48–55, May 2023, Accessed: May 30, 2024. [Online]. Available: <https://journal.unej.ac.id/JSED/article/view/307>
- [10] Jesi Pebralia, "Analisis Curah Hujan Menggunakan Machine Learning Metode Regresi Linier Berganda Berbasis Python dan Jupyter Notebook," *JIFP (Jurnal Ilmu Fisika dan Pembelajarannya)/Jurnal Ilmu Fisika dan Pembelajarannya (JIFP)*, vol. 6, no. 2, pp. 23–30, Dec. 2022, doi: <https://doi.org/10.19109/jifp.v6i2.13958>.
- [11] Muhammad Sholeh, Suraya Suraya, and D. Andayati, "Machine Linear untuk Analisis Regresi Linier Biaya Asuransi Kesehatan dengan Menggunakan Python Jupyter Notebook," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. 8, no. 1, pp. 20–20, Apr. 2022, doi: <https://doi.org/10.26418/jp.v8i1.48822>.
- [12] A. Saiful, "Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 1, pp. 41–50, Mar. 2021, doi: <https://doi.org/10.35957/jatisi.v8i1.701>.
- [13] V. T. Sapanji, S. Lestari, M. and R. Samihardjo, "Prediksi Indeks Bursa Efek Indonesia 2023 Pendekatan ARIMA, Machine Learning dengan R Programming," *Jurnal Manajemen Informatika (JAMIKA)*, vol. XIII, no. 2, pp. 163-177, 2023.
- [14] W. Budiharto, *Machine Learning dan Computational Intelligence*, Yogyakarta: Penerbit ANDI, 2016.
- [15] M. F. Febrianto, "Gaming Laptop Dataset" Kaggle, 2023. Available: <https://www.kaggle.com/datasets/faturfebr/gaming-laptop-dataset> [Accessed June 25, 2024]
- [16] M. L. Mu'tashim, S. A. Damayanti, H. N. Zaki, T. Muhayat and R. Wirawan, "Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression," *JURNAL INFORMATIK*, vol. XVII(3), pp. 238-245, 2021.
- [17] V. Advani, "34 Open-Source Python Libraries You Should Know About," *GreatLearning*, Sep. 11, 2020. <https://www.mygreatlearning.com/blog/open-source-python-libraries/>
- [18] "Data Cleaning in Machine Learning: Steps & Process [2022]," [www.v7labs.com](https://www.v7labs.com/blog/data-cleaning-guide). <https://www.v7labs.com/blog/data-cleaning-guide>
- [19] M. Uddin, J. Lee, S. Rizvi, and S. Hamada, "Proposing Enhanced Feature Engineering and a Selection Model for Machine Learning Processes," *Applied Sciences*, vol. VIII, no. 4, p. 646, Apr. 2018, doi: <https://doi.org/10.3390/app8040646>.
- [20] Abu Tholib, *Implementasi Algoritma Machine Learning Berbasis Web dengan Framework Streamlit*. Pustaka Nurja.
- [21] R. B. Widodo, *Machine Learning Metode k-Nearest Neighbors Klasifikasi Angka Bahasa Isyarat*, Malang: Media Nusa Creative, 2022.
- [22] A. D. Siburian et al., "Laptop Price Prediction with Machine Learning Using Regression Algorithm," *Jurnal Sistem Informasi dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, vol. VI, no. 1, pp. 87–91, Sep. 2022, doi: <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v6i1.2850>.
- [23] S. Saadah and H. Salsabila, "Prediksi Harga Bitcoin Menggunakan," *Jurnal Politeknik Caltex Riau*, vol. VII, no. 1, pp. 24-32, 2021.
- [24] V. Atlantic, E. Sulistianingsih and H. Perdana, "Gradient Boosting Machine Pada Klasifikasi Kelulusan Mahasiswa," *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, vol. 12, no. 2, pp. 165-174, 2024.