

# **ANALISIS DATA PERUMAHAN UNTUK MEMPREDIKSI HARGA RUMAH MENGUNAKAN *MACHINE LEARNING***

**Syafrial Fachri Pane  
Rakasona  
Rifqi Fathurrohman**



# **ANALISIS DATA PERUMAHAN UNTUK MEMPREDIKSI HARGA RUMAH MENGUNAKAN MACHINE LEARNING**

***Penulis :***

Rakasona  
Rifqi Fathurrohman  
Syafrial Fachri Pane

ISBN :

***Editor :***

Roni Andarsyah

***Penyunting :***

Roni Andarsyah

***Desain sampul dan Tata letak :***

Rifqi Fathurrohman

***Penerbit :***

Penerbit Buku Pedia

***Redaksi :***

Athena Residence Blok. E No. 1, Desa Ciwaruga,  
Kec. Parongpong, Kab. Bandung Barat 40559  
Tel. 628-775-2000-300  
Email : [penerbit@bukupedia.co.id](mailto:penerbit@bukupedia.co.id)

***Distributor :***

Informatics Research Center  
Jl. Sariasih No. 54  
Bandung 40151  
Email : [irc@ulbi.ac.id](mailto:irc@ulbi.ac.id)

Cetakan Pertama, 2023

Hak cipta dilindungi undang-undang  
Dilarang memperbanyak karya tulis ini dalam bentuk dan  
dengan cara apa pun tanpa ijin tertulis dari penerbit

# KATA PENGANTAR

Sebagai salah satu bentuk upaya untuk memecahkan masalah yang ada dalam dunia properti, kami menghadirkan buku yang berjudul "Analisis Data Perumahan untuk Memprediksi Harga Rumah Menggunakan Machine Learning". Buku ini bertujuan untuk memberikan pemahaman dan solusi terhadap masalah yang ada dalam dunia properti, khususnya memprediksi harga rumah. Dengan adanya teknologi machine learning, kami yakin bahwa buku ini akan memberikan solusi dan pemahaman yang lebih baik bagi para pembaca. Pembahasan yang diberikan akan menunjukkan bagaimana teknologi machine learning dapat digunakan untuk memprediksi harga rumah dan membantu para pemain industri properti dalam membuat keputusan yang lebih baik.

Buku ini sendiri merupakan wujud dari kebutuhan akan pentingnya meningkatkan skill dengan tantangan yang akan dihadapi oleh perusahaan dan generasi milenial di Era Industri 4.0, terutama yang berkaitan dengan komitmen kerja/bisnis dan adversity quotient (AQ). Kami berharap bahwa buku ini dapat memberikan manfaat bagi para pembaca dan dapat menjadi salah satu sumber referensi yang berguna bagi para profesional dalam bidang properti.

Buku ini disajikan dalam setiap bab yang berisi langkah-langkah untuk pembaca pemula. Setiap bab dibuka dengan definisi dan kerangka kerja. Pada bagian selanjutnya diterangkan cara membuat kode program beserta penjelasan setiap baris kode program tersebut. Pada bagian akhir disertakan kode program untuk latihan dan evaluasi pembaca. Setiap chapter atau bab yang disajikan dalam buku ini disertakan juga kode program yang bisa diakses melalui link: <https://github.com/rifqifathurrohman/PREDIKSI-HARGA-RUMAH-MENGGUNAKAN-MACHINE-LEARNING>

Bandung, 12 Februari 2023

Penulis

# DAFTAR ISI

<b>KATA PENGANTAR .....</b>	<b>i</b>
<b>DAFTAR ISI .....</b>	<b>ii</b>
<b>DAFTAR GAMBAR.....</b>	<b>v</b>
<b>DAFTAR TABEL .....</b>	<b>vii</b>
<b>BAB 1 PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang.....	1
<b>BAB 2 TUJUAN INTRUKSIONAL DAN CAPAIAN PEMBELAJARAN .....</b>	<b>4</b>
2.1 Tujuan Intruksional .....	4
2.2 Capaian Pembelajaran .....	4
<b>BAB 3 URAIAN MATERI .....</b>	<b>5</b>
3.1 <i>State of the art</i> .....	5
3.2 Teori - teori yang digunakan .....	6
3.2.1 Data.....	6
3.2.2 Prediksi .....	7
3.2.3 Harga Rumah .....	7
3.2.4 <i>Machine Learning</i> .....	7
3.2.5 <i>Random Forest</i> .....	9
3.2.6 <i>Outlier</i> .....	11
3.2.7 Metode Evaluasi .....	13
3.2.8 CRISP-DM .....	17
3.2.9 UML ( <i>Unified Modeling Language</i> ).....	18
3.3 Alat dan aplikasi yang digunakan .....	19
3.3.1 HTML.....	19
3.3.2 CSS .....	20
3.3.3 XAMPP .....	20
3.3.4 Python.....	21

3.3.5	<i>Google Colab</i> .....	22
3.3.6	<i>VScode</i> .....	23
3.3.7	<i>Tableau</i> .....	24
3.3.8	<i>Django</i> .....	25
<b>BAB 4 HASIL DAN PEMBAHASAN .....</b>		<b>27</b>
4.1	Analisis .....	27
4.2	Business Understanding.....	27
4.3	Data Understanding .....	27
4.4	Preprocessing Data .....	29
4.4.1	Membaca Data (Pandas).....	29
4.4.2	Input Nilai Pada Kumpulan Data ( <i>Simple Imputer</i> ).....	30
4.4.3	Melakukan Encoding ( <i>Ocean Proximity</i> ).....	32
4.4.4	Mendapatkan Nilai <i>Outlier</i> .....	33
4.4.5	Standarisasi Data .....	37
4.4.6	Train Test Split .....	39
4.4.7	Pemodelan.....	39
4.5	Evaluasi Model .....	41
4.6	Analisis Sistem.....	42
4.6.1	<i>Flowmap</i> Prediksi Harga Rumah.....	42
4.7	Perancangan Sistem (UML) .....	43
4.7.1	Use Case Diagram .....	43
4.7.2	Class Diagram.....	51
4.7.3	Activity Diagram.....	52
4.7.4	Statechart Diagram.....	58
4.7.5	ERD (Entity Relationship Diagram) .....	63
4.7.6	Struktur Menu .....	64
<b>BAB 5 IMPLEMENTASI APLIKASI .....</b>		<b>66</b>
5.1	Implementasi Kode Aplikasi.....	66
5.1.1	Implementasi Package Yang Digunakan .....	66

5.1.2	Implementasi Proses <i>Login</i> .....	67
5.1.3	Implementasi Proses Register .....	68
5.1.4	Implementasi Proses Prediksi Harga Rumah .....	68
5.1.5	Implementasi Proses CRUD Data Rumah .....	70
5.1.6	Implementasi Proses Visualisasi .....	73
5.2	Implementasi Halaman Antar Muka.....	74
5.2.1	Implementasi Halaman Depan .....	74
5.2.2	Implementasi Halaman Login .....	74
5.2.3	Implementasi Halaman Register.....	75
5.2.4	Implementasi Halaman Prediksi Rumah.....	76
5.2.5	Implementasi Halaman Hasil Prediksi Rumah .....	77
5.2.6	Implementasi Halaman Menambahkan Data Rumah Baru .....	77
5.2.7	Implementasi Halaman Daftar data rumah baru .....	78
5.2.8	Implementasi Halaman Mengedit Data Rumah Baru .....	78
5.2.9	Implementasi Halaman Visualisasi .....	79
<b>BAB 6</b>	<b>LATIHAN.....</b>	<b>81</b>
6.1	LATIHAN .....	81
<b>BAB 7</b>	<b>RANGKUMAN MATERI .....</b>	<b>80</b>
7.1	RANGKUMAN MATERI.....	80
6.2	Saran .....	80
<b>DAFTAR PUSTAKA</b>	<b>.....</b>	<b>81</b>
<b>GLOSARIUM</b>	<b>.....</b>	<b>88</b>
<b>TENTANG PENULIS</b>	<b>.....</b>	<b>98</b>

# DAFTAR GAMBAR

Gambar 3.1 <i>Data Processing</i> .....	6
Gambar 3.2 Grafik Regresi dan Klasifikasi .....	8
Gambar 3.3 <i>Outlier box plot</i> .....	11
Gambar 3.4 Residual RMSE.....	14
Gambar 3.5 CRISP-DM .....	17
Gambar 3.6 HTML.....	19
Gambar 3.7 CSS.....	20
Gambar 3.8 XAMPP.....	20
Gambar 3.9 Python .....	21
Gambar 3.10 <i>Google Colab</i> .....	22
Gambar 3.11 VScode .....	23
Gambar 3.12 Tableau .....	24
Gambar 3.13 Django .....	25
Gambar 4.1 Sesudah <i>Encoding</i> .....	32
Gambar 4.2 Sebelum <i>Encoding</i> .....	32
Gambar 4.3 Plot sebelum <i>cleaning data</i> .....	34
Gambar 4.4 Plot setelah <i>cleaning data</i> .....	36
Gambar 4.5 Normalisasi Data .....	37
Gambar 4.6 <i>Heatmap</i> Normalisasi.....	38
Gambar 4.7 <i>Plot scatter</i> nilai prediksi dan aktual.....	40
Gambar 4.8 Visualisasi Model .....	41
Gambar 4.9 <i>Flowmap</i> Prediksi Harga Rumah .....	42
Gambar 4.10 <i>Use Case Diagram</i> .....	44
Gambar 4.11 <i>Class Diagram</i> .....	51
Gambar 4.12 <i>Activity Login</i> .....	52
Gambar 4.13 <i>Activity Registrasi</i> .....	53
Gambar 4.14 <i>Activity Pengolahan Data</i> .....	54
Gambar 4.15 <i>Activity</i> Prediksi Harga Rumah .....	55
Gambar 4.16 <i>Activity Diagram</i> Visualisasi .....	56
Gambar 4.17 <i>Activity Diagram Logout</i> .....	57
Gambar 4.18 <i>Statechart Diagram Login Admin</i> .....	58

Gambar 4.19 <i>Statechart</i> Register .....	59
Gambar 4.20 <i>Statechart</i> Prediksi harga Rumah .....	60
Gambar 4.21 <i>Statechart</i> pengolahan data .....	61
Gambar 4.22 <i>Statechart</i> Visualisasi Diagram.....	62
Gambar 4.23 <i>Entity Relationship Diagram</i> .....	64
Gambar 4.24 Struktur Menu.....	64
Gambar 5.1 Tampilan UI Halaman depan.....	74
Gambar 5.2 Tampilan UI Depan.....	74
Gambar 5.3 Tampilan UI Register .....	75
Gambar 5.4 Halaman UI Input nilai prediksi.....	76
Gambar 5.5 Halaman Hasil Prediksi Rumah .....	77
Gambar 5.6 Halaman Tambah data rumah baru .....	77
Gambar 5.7 Halaman Daftar data rumah baru .....	78
Gambar 5.8 Halaman Edit Data rumah baru .....	78
Gambar 5.9 Halaman Visualisasi dari Tableau.....	79



## DAFTAR TABEL

Tabel 3.1 Residual RMSE.....	15
Tabel 4.1 Penjelasan Dataset.....	28
Tabel 4.2 Table Housing.head.....	29
Tabel 4.3 Table <i>Missing Value</i> .....	31
Tabel 4.4 Imputasi median .....	31
Tabel 4.5 Setelah Imputasi.....	32
Tabel 4.6 parameter pelatihan dan pengujian .....	39
Tabel 4.7 Hasil harga aktual dan prediksi .....	40
Tabel 4.8 Table Definisi Admin .....	44
Tabel 4.9 Table Definisi <i>Use Case</i> .....	46
Tabel 4.10 Skenario <i>Use Case</i> Pada <i>Machine learning</i> .....	47
Tabel 4.11 Skenario <i>Use Case</i> Melakukan Prediksi.....	47
Tabel 4.12 Skenario <i>Use Case</i> Melakukan Visualisasi .....	48
Tabel 4.13 Skenario <i>Use Case</i> Melakukan Pengolahan Data .....	49
Tabel 4.14 Skenario <i>Use Case</i> Melakukan <i>Logout</i> .....	50