



Prediksi Harga Rumah di Boston Dengan Model Regresi Linear Menggunakan Python

Intania Dharma Hartarti¹, Intan Amelia Septiyani², Daniel Armando Gultom³, Yayan Hendrian⁴, Shynde Limar Kinanti⁵

^{1,2,3,4,5}Informatika, Teknik dan Informatika, Universitas Bina Sarana Informatika

¹intaniadharma21@gmail.com, ²intanameliaaa22@gmail.com, ³danielarmandogultom2000@gmail.com,

⁴yayan.yhn@bsi.ac.id, ⁵shynde.slk@bsi.ac.id

Abstrak

Prediksi harga rumah merupakan aspek krusial dalam sektor properti dan ekonomi karena dapat membantu proses pengambilan keputusan bagi pembeli, penjual, serta pengembang. Penelitian ini memiliki tujuan untuk membangun model prediksi harga rumah dengan pendekatan algoritma regresi linear. Dataset yang digunakan adalah Boston Housing, yang memuat berbagai variabel penentu harga rumah. Beberapa variabel penting dalam dataset tersebut mencakup jumlah kamar, usia bangunan, tingkat kriminalitas di lingkungan sekitar, dan jarak ke pusat kota. Pengolahan data dilakukan menggunakan bahasa pemrograman Python di editor Visual Studio Code (VS Code). Proses analisis mencakup eksplorasi data, analisis korelasi antar variabel, pemilihan fitur, pembangunan model, evaluasi performa model, dan visualisasi hasil. Model regresi linear kemudian dievaluasi dengan metrik seperti MAE, MSE, RMSE, dan R^2 untuk mengukur akurasi prediksi. Hasil evaluasi menunjukkan bahwa regresi linear mampu menghasilkan prediksi yang cukup baik dengan nilai R^2 yang menggambarkan korelasi yang kuat antara fitur dan harga. Berdasarkan temuan ini, regresi linear dinilai efektif sebagai pendekatan awal dalam memprediksi harga rumah berdasarkan data historis.

Kata kunci: *Prediksi Harga Rumah, Regresi Linear, Boston Housing, Python, Machine Learning*.

1. Latar Belakang

Perkembangan teknologi informasi yang pesat dalam dua dekade terakhir telah membawa perubahan besar di berbagai sektor, termasuk sektor properti. Digitalisasi data dan kemajuan komputasi memungkinkan pelaku industri untuk mengakses, mengolah, dan menganalisis informasi dalam jumlah besar secara lebih cepat dan akurat. Salah satu kebutuhan penting dalam industri properti adalah kemampuan untuk memprediksi harga rumah secara tepat. Prediksi harga rumah tidak hanya menjadi acuan bagi pembeli dan penjual dalam menentukan nilai transaksi, tetapi juga menjadi dasar penting dalam pengambilan keputusan oleh investor, pengembang, serta pembuat kebijakan dalam menyusun strategi pembangunan yang berkelanjutan [1].

Di era digital saat ini, pendekatan konvensional dalam menilai harga rumah mulai tergantikan oleh metode berbasis data dan kecerdasan buatan. Salah satu cabang dari kecerdasan buatan yang relevan dalam konteks ini adalah *machine learning*, khususnya pendekatan *supervised learning* yang memanfaatkan data historis untuk membangun model prediktif [2]. Dalam konteks prediksi harga rumah, algoritma regresi linier masih menjadi pilihan populer karena kemampuannya dalam memodelkan hubungan linier antara variabel-variabel input (fitur) dengan target output, yaitu harga rumah. Metode ini tidak hanya sederhana dalam implementasinya, tetapi juga cukup efektif dalam memberikan hasil yang dapat dijelaskan secara matematis [3].

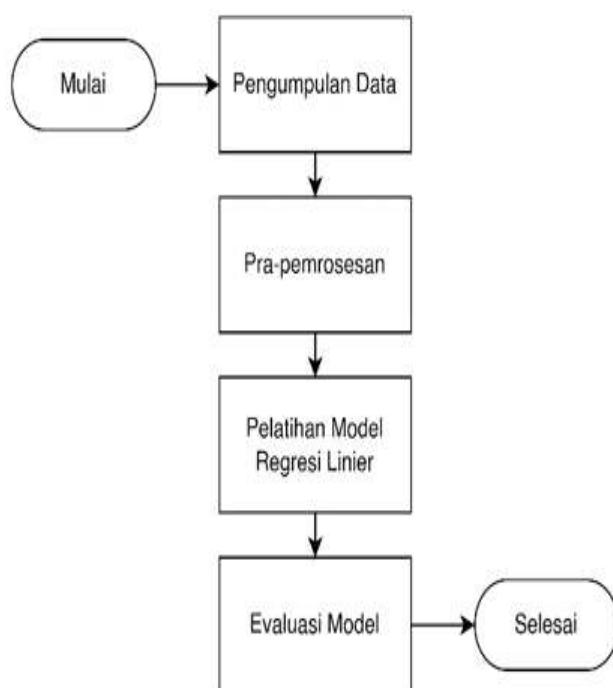
Penelitian ini bertujuan untuk membangun model prediksi harga rumah di Kota Boston menggunakan algoritma *Multiple Linear Regression*. Data yang digunakan diambil dari *Boston Housing Dataset* yang tersedia di platform Kaggle dan diolah menggunakan bahasa pemrograman Python pada lingkungan pengembangan Visual Studio Code [4]. Dataset ini memiliki kompleksitas tinggi karena mencakup berbagai variabel dari aspek fisik, sosial, ekonomi, hingga lingkungan, seperti tingkat kriminalitas, kualitas pendidikan, pajak properti, dan jarak ke pusat kota. Pemilihan dataset ini dilakukan agar model yang dibangun memiliki cakupan yang lebih luas dan realistik dibandingkan dengan penelitian terdahulu yang umumnya hanya memanfaatkan variabel terbatas, seperti penelitian oleh Muhammad Labib [5] yang hanya memfokuskan analisis pada variabel fisik dan wilayah tertentu seperti Jakarta Selatan.

Inovasi utama dari penelitian ini terletak pada dua aspek: pertama, penggunaan pendekatan multidimensi terhadap faktor-faktor yang memengaruhi harga rumah, dan kedua, penerapan metrik evaluasi model yang lebih komprehensif. Penilaian performa model dilakukan menggunakan empat metrik utama yaitu *Mean Absolute Error (MAE)*, *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)*, dan *koefisien determinasi (R²)*, untuk memberikan gambaran yang lebih lengkap mengenai akurasi dan efisiensi model [6]. Evaluasi ini bertujuan agar hasil penelitian tidak hanya memiliki nilai akademis, tetapi juga dapat diterapkan secara praktis di dunia nyata. Selain menyajikan kontribusi teknis dalam hal pemodelan dan evaluasi, penelitian ini juga memberikan manfaat strategis bagi pihak-pihak yang terlibat dalam industri properti dan perencanaan tata kota.

Dengan memanfaatkan teknologi machine learning, prediksi harga rumah dapat dilakukan secara lebih cepat, efisien, dan adaptif terhadap perubahan data pasar. Hal ini diharapkan dapat membantu pembuat kebijakan dalam menetapkan strategi pembangunan yang lebih tepat sasaran dan membantu masyarakat dalam mengambil keputusan investasi yang lebih bijak. Dengan demikian, penelitian ini menawarkan pendekatan yang lebih modern, aplikatif, dan menyeluruh dalam memprediksi harga rumah. Tidak hanya berfokus pada aspek teknis, tetapi juga menekankan pentingnya keberagaman variabel dan cakupan data sebagai dasar pengambilan keputusan berbasis data. Pendekatan ini diharapkan dapat melengkapi kekurangan dari penelitian-penelitian sebelumnya yang masih bersifat sempit dalam konteks fitur maupun wilayah analisis [7].

2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif yang berfokus pada pengumpulan dan analisis data numerik untuk menguji hipotesis secara objektif. Pendekatan ini berlandaskan pada filsafat positivisme, yang menekankan pada observasi empiris dan logika deduktif dalam memahami fenomena [8]. Dalam konteks ini, data berupa angka dianalisis menggunakan regresi linier. Metode ini dipilih karena kemampuannya dalam mengidentifikasi dan mengukur hubungan linier antara variabel independen dan dependen. Sumber data adalah Boston Housing Dataset yang memuat informasi numerik seperti jumlah kamar, tingkat kejahatan, dan jarak ke pusat kota. Seluruh variabel ini digunakan sebagai prediktor terhadap variabel target yaitu MEDV (*Median Value of Owner-Occupied Homes*).



Gambar 1. Alur Penelitian

2.1. Pengumpulan Data

Penelitian ini menggunakan Boston Housing Dataset versi terbaru yang diunduh dari platform Kaggle, yang diunggah oleh Motahare Shokr pada tahun 2024. Dataset ini merupakan pembaruan dari versi asli yang dikembangkan oleh Harrison dan Rubinfeld pada tahun 1978, dengan penyempurnaan format serta peningkatan kualitas data agar lebih relevan untuk analisis modern. Dataset ini terdiri dari 506 observasi dengan 14 variabel yang merepresentasikan berbagai aspek penting dalam penentuan harga rumah, seperti tingkat kejahatan di lingkungan sekitar, kualitas pendidikan, akses terhadap fasilitas transportasi, serta kondisi sosial ekonomi lainnya.

di wilayah Boston [9]. Data yang digunakan dalam penelitian ini merupakan data sekunder, yaitu data yang telah tersedia sebelumnya dan diperoleh bukan secara langsung dari lapangan, melainkan dari sumber yang sudah dipublikasikan resmi, seperti website instansi, artikel ilmiah, jurnal ataupun dataset digital yang tersedia secara daring [10]. Dalam penelitian ini, data diimpor file CSV.

Tabel 1. Dataset Boston Housing

No	Nama	Deskripsi
1	Crim	Tingkat kriminalitas per kapita per kota
2	Zn	Proporsi lahan dengan zonasi untuk rumah mewah
3	Indus	Proporsi area bisnis non-retail
4	Chas	Dummy variabel: 1 jika berbatasan dengan sungai Charles
5	Nox	Konsentrasi oksida nitrogen (parts per 10 juta)
6	Rm	Rata-rata jumlah kamar per rumah
7	Age	Proporsi unit yang dibangun sebelum 1940
8	Dis	Jarak ke pusat bisnis Boston
9	Rad	Aksesibilitas ke jalan raya utama
10	Tax	Pajak properti per \$10.000
11	Ptratio	Rasio murid-guru per kota
12	B	Populasi orang kulit hitam
13	Lstat	Persentase populasi berstatus sosial rendah
14	Medv	Median harga rumah (dalam ribuan USD) - target

2.2. Pra-pemrosesan

Pra-pemrosesan data adalah langkah awal yang penting dalam membangun model machine learning, dimulai dari Pembersihan data dilakukan secara sistematis, dimulai dengan mengubah semua nama kolom menjadi huruf kecil (df.columns.str.lower()) untuk konsistensi. Lima data pertama ditampilkan menggunakan head(), diikuti pengecekan jumlah baris dan kolom (shape) serta tipe data tiap kolom (dtypes). pandas digunakan untuk manipulasi dan analisis data dalam bentuk tabel (DataFrame), yang sangat membantu dalam pengorganisasian dan pembersihan data [11], nilai kosong diperiksa dengan isnull().sum(), sementara outlier dideteksi menggunakan metode Interquartile Range (IQR) dan divisualisasikan melalui boxplot.

Eksplorasi data dilakukan dengan membuat histogram tiap fitur menggunakan seaborn.histplot() untuk memahami distribusi data. Statistik deskriptif ditampilkan dengan describe(), dan korelasi antar fitur dengan target medv dihitung menggunakan corr(), lalu divisualisasikan dalam bentuk heatmap menggunakan seaborn.heatmap(). Dalam proses ini, NumPy juga berperan penting dalam komputasi numerik dan pengolahan array data [12]. Untuk mendukung visualisasi, digunakan pustaka Matplotlib dan Seaborn memungkinkan pembuatan grafik seperti histogram, scatter plot, heatmap, dan residual plot guna mempermudah pemahaman distribusi data dan evaluasi model [13]. Pemilihan fitur dan target dilakukan melalui input pengguna menggunakan fungsi input(), kemudian divalidasi untuk memastikan bahwa variabel yang dipilih tersedia dalam dataset.

2.3. Pelatihan Model Regresi Linear

Menurut [14], analisis regresi mengukur kekuatan dan arah hubungan antara variabel dependen dan independen. Regresi linier berganda menjelaskan hubungan linear antara beberapa variabel independen (X) dengan variabel dependen (Y). Dalam prediksi harga rumah, model ini mencari garis lurus terbaik yang memetakan fitur ke target, dirumuskan sebagai:

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon \quad (1)$$

Di mana y adalah variabel dependen, x adalah variabel independen, α merupakan konstanta (intercept), β sebagai koefisien regresi (slope), dan ϵ error atau residual.

2.4. Evaluasi Model

Data dibagi menjadi data latih (80%) dan data uji (20%) menggunakan fungsi train_test_split() dari scikit-learn untuk melatih model dan menguji kinerjanya pada data baru. Model regresi linier dibuat dengan LinearRegression() dari sklearn.linear_model dan dilatih menggunakan data latih dengan metode fit(). Menurut [15], scikit-learn memudahkan implementasi regresi linier serta menyediakan metrik evaluasi seperti MAE sebagai Rata-rata kesalahan absolut dari hasil prediksi, MSE sebagai Rata-rata kuadrat dari selisih antara nilai aktual dan prediksi, RMSE sebagai Akar dari MSE yang semakin kecil nilainya, semakin baik performa model dan R^2 sebagai pengaruh variabel independent untuk menunjukkan bahwa model dapat menjelaskan berapa persen variasi dari harga rumah berdasarkan fitur-fitur input. Evaluasi model dilakukan dengan menghitung:

Mean Absolute Error

$$MAE = \left(\frac{1}{n} \right) \sum |y_i - \hat{y}_i| \quad (2)$$

Mean Squared Error

$$MSE = \left(\frac{1}{n} \right) \sum (y_i - \hat{y}_i)^2 \quad (3)$$

Root Mean Squared Error

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

Koefisien Determinasi

$$R^2 = 1 - \left[\frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \right] \quad (5)$$

Metrik-metrik ini digunakan untuk mengukur seberapa baik model memprediksi data dan menjelaskan variasi variabel target.

3. Hasil dan Diskusi

3.1. Analisis Korelasi Pearson



Gambar 2. Heatmap Korelasi

Gambar heatmap tersebut menggambarkan hubungan korelasi Pearson antar fitur dalam dataset Boston Housing, termasuk variabel target medv yang merepresentasikan harga rumah dalam ribuan dolar. Warna merah menunjukkan korelasi positif, sedangkan warna biru menunjukkan korelasi negatif. Fitur dengan korelasi positif terkuat adalah rm (jumlah rata-rata kamar) sebesar +0.70, artinya semakin banyak kamar, harga rumah cenderung meningkat. Fitur zn (+0.36), chas (+0.18), dan b (+0.33) juga berkorelasi positif, meskipun tidak sekuat rm.

Sebaliknya, lstat (persentase penduduk berpenghasilan rendah) memiliki korelasi negatif paling kuat terhadap medv, yaitu -0.74, menunjukkan bahwa area dengan tingkat kemiskinan tinggi cenderung memiliki harga rumah lebih rendah. Sebuah penelitian korelasi ini sejalan oleh peneliti sebelumnya yang menemukan bahwa fitur sosial-ekonomi seperti status sosial penduduk memengaruhi nilai properti [6].

Fitur lain seperti ptratio (-0.51), nox (-0.43), indus (-0.48), tax (-0.47), dan crim (-0.39) juga berkorelasi negatif, menandakan pengaruh negatif dari kualitas lingkungan dan sosial terhadap harga. Sementara itu, dis (jarak ke pusat kerja) memiliki korelasi positif rendah (+0.25), yang bisa mengindikasikan preferensi terhadap lokasi yang lebih tenang. Kesimpulannya, fitur rm dan lstat menjadi variabel paling signifikan dalam model regresi linear, sedangkan fitur lain perlu dianalisis lebih lanjut untuk menghindari multikolinearitas.

3.2. Evaluasi Model Prediksi Harga Rumah di Boston

Model regresi linear dibangun dengan seluruh fitur sebagai input kecuali variabel target medv. Setelah pelatihan model dengan data latih, diperoleh hasil evaluasi nilai MAE sebesar 3.189 adalah rata-rata selisih absolut antara nilai aktual dan prediksi adalah sekitar 3.189. Nilai ini cukup baik karena menunjukkan bahwa model cenderung memiliki kesalahan prediksi sekitar ± 3.2 satuan harga rumah. Nilai MSE sebesar 24.291 adalah rata-rata kesalahan untuk kuadrat antara nilai aktual dan prediksi, dan nilai MSE yang lebih kecil menunjukkan bahwa model memiliki prediksi yang stabil dan tidak terlalu banyak kesalahan ekstrim. RMSE adalah akar MSE dan menyatakan kesalahan rata-rata dalam satuan yang sama dengan target (harga rumah). Dengan nilai RMSE 4.929, model memiliki deviasi rata-rata ± 4.93 dari nilai sebenarnya. Nilai ini juga menunjukkan model memiliki kemampuan untuk membuat prediksi yang cukup akurat. Nilai R^2 sebesar 0,669, atau 66,9%, menunjukkan model dapat menjelaskan sekitar 66,9 persen variasi harga yang disebabkan oleh fitur-fitur input. Meskipun nilai ini belum mencapai 100%, nilai ini menunjukkan bahwa model cukup menjelaskan hubungan antara fitur dan harga rumah.

Tabel 2. Evaluasi Model MAE, MSE, RMSE, R^2

No	Metrik	Nilai
1	MAE	3.189
2	MSE	24.291
3	RMSE	4.929
4	R^2	0.669

3.3. Interpretasi Koefisien

Hasil interpretasi koefisien dari model regresi linear dalam memprediksi harga rumah di Boston berdasarkan berbagai fitur. Model menghasilkan intercept sebesar 30.246, yang berarti jika seluruh variabel bernilai nol, harga rumah diperkirakan sebesar \$30.246 atau sekitar 30 ribu dolar. Di antara seluruh fitur, jumlah rata-rata kamar per rumah (rm) memiliki pengaruh paling besar dengan koefisien +4.4388. Artinya, setiap penambahan satu kamar akan meningkatkan harga rumah sekitar \$4.440. Sebaliknya, fitur nox (tingkat polusi udara) memiliki pengaruh negatif terbesar dengan koefisien -17.2026, yang menunjukkan bahwa peningkatan satu unit polusi akan menurunkan harga rumah sebesar \$17.200.

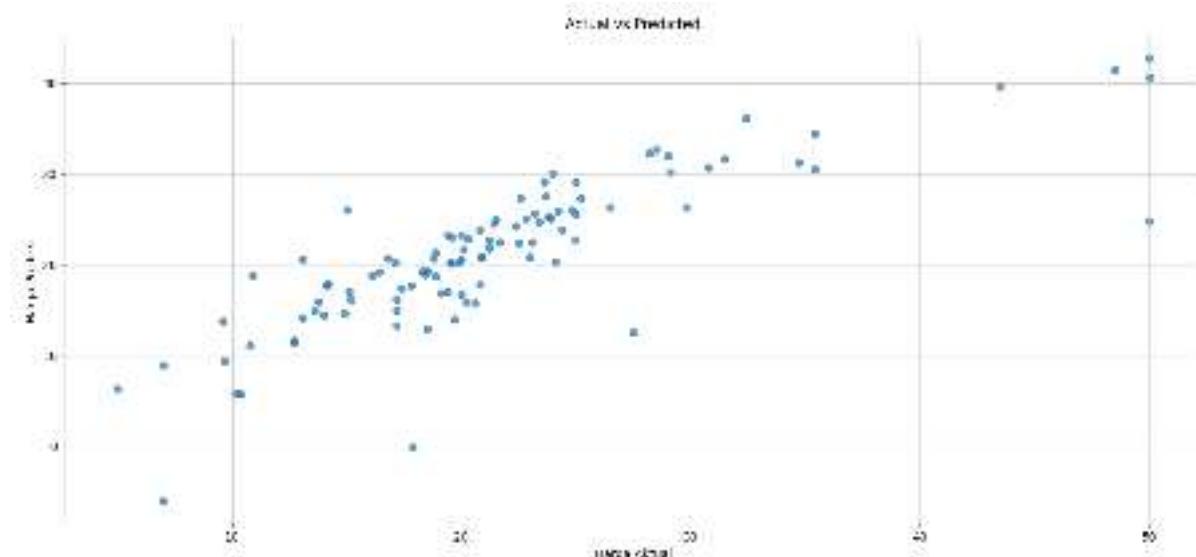
Selain itu, keberadaan rumah di dekat Sungai Charles (chas) menaikkan harga sekitar \$2.780, sedangkan meningkatnya jarak ke pusat kerja (dis) justru menurunkan harga sekitar \$1.450. Rasio murid terhadap guru (ptratio) yang lebih tinggi, yang menceerminkan kualitas pendidikan yang menurun, akan menurunkan harga rumah sebesar \$910 per satuan kenaikan. Faktor sosial ekonomi seperti lstat, yaitu persentase penduduk berpenghasilan rendah, juga memiliki dampak negatif yang signifikan; setiap kenaikan 1% pada lstat menurunkan harga rumah sebesar \$510. Fitur lain seperti tingkat kejahatan (crim) dan pajak properti (tax) juga menurunkan harga rumah, meskipun dengan pengaruh yang lebih kecil, masing-masing sekitar \$110 dan \$10 per satuan kenaikan.

Sementara itu, fitur zn (zona perumahan), indus (proporsi area industri), dan b (proporsi populasi kulit hitam) memiliki pengaruh yang sangat kecil terhadap harga rumah, hanya berkisar antara \$10 hingga \$40. Umur bangunan (age) juga hampir tidak berpengaruh signifikan, hanya menurunkan harga sekitar \$6 per tahun. Dari interpretasi ini dapat disimpulkan bahwa variabel yang paling berperan dalam memengaruhi harga rumah di Boston adalah kualitas lingkungan, sosial, dan karakteristik fisik rumah itu sendiri, dengan fitur rm, nox, dan lstat menjadi faktor kunci dalam menentukan nilai properti.

Tabel 3. Dataset Boston Housing

No	Nama	Koefisien	Analisa Fitur
1	Crim	-0.113056	
2	Zn	0.030110	
3	Indus	0.040381	
4	Chas	2.784438	
5	Nox	-17.202633	Fitur paling sedikit berpengaruh, setiap kenaikan 1 unit pada Rm, prediksi harga berubah 4.44
6	Rm	4.438835	Fitur paling berpengaruh, setiap kenaikan 1 unit pada Rm, prediksi harga berubah 4.44
7	Age	-0.006296	
8	Dis	-1.447865	
9	Rad	0.262430	
10	Tax	-0.010647	
11	Ptratio	-0.915456	
12	B	0.012351	
13	Lstat	-0.508571	
Intercept		30.24675099392388	

3.4. Perbandingan Harga Prediksi dan Harga Aktual



Gambar 3. Harga Prediksi vs Harga Aktual

Gambar scatter plot tersebut menunjukkan hubungan antara harga rumah aktual dan harga rumah yang diprediksi oleh model regresi linear pada dataset Boston Housing. Setiap titik merepresentasikan satu rumah, dengan sumbu horizontal menunjukkan harga aktual dan sumbu vertikal menunjukkan harga prediksi, keduanya dalam satuan ribuan dolar, dengan garis ideal $y = x$. Sebagian besar titik berada dekat dengan garis diagonal imajiner dari kiri bawah ke kanan atas, yang mengindikasikan bahwa model memberikan prediksi yang cukup akurat, terutama untuk rumah dengan harga di kisaran \$15.000 hingga \$35.000. Misalnya, rumah dengan harga aktual \$25.000 umumnya diprediksi mendekati nilai tersebut. Namun, terdapat beberapa penyimpangan yang cukup signifikan. Pada harga rumah yang tinggi, terutama sekitar \$50.000, prediksi model justru bervariasi secara luas, mulai dari sekitar \$25.000 hingga \$43.000, yang menunjukkan bahwa model kesulitan menangkap harga tertinggi dengan tepat. Demikian pula, pada harga yang sangat rendah (di bawah \$10.000), terdapat beberapa prediksi yang meleset jauh bahkan mendekati nilai negatif.

4. Kesimpulan

Secara keseluruhan, hasil evaluasi dan visualisasi menunjukkan bahwa regresi linear mampu menjelaskan sebagian besar hubungan antara fitur-fitur dengan harga rumah, meskipun masih terdapat ruang untuk perbaikan, terutama dalam menangani non-linearitas dan outlier ekstrem. Berdasarkan dataset Boston Housing, model regresi linear yang dibangun menghasilkan nilai R^2 sebesar 66,9%, MAE sebesar 3,18, dan RMSE sebesar 4,93, yang menunjukkan performa prediksi yang cukup baik. Sebagian besar variabel sosial, ekonomi, dan lingkungan berhasil dijelaskan oleh model ini. Faktor paling berpengaruh terhadap harga rumah adalah jumlah rata-rata kamar (rm), di mana setiap penambahan satu kamar diperkirakan meningkatkan harga rumah sebesar \$4.440. Sebaliknya, tingkat polusi udara (nox) memiliki dampak negatif yang signifikan, dengan setiap kenaikan satu unit nox diperkirakan menurunkan harga rumah sebesar \$17.200. Hal ini menegaskan pentingnya kualitas lingkungan dan lokasi properti terhadap nilai jual rumah. Selain itu, hasil korelasi juga menunjukkan bahwa komponen lingkungan dan sosial, seperti Istat (persentase penduduk berpenghasilan rendah), turut memberikan pengaruh negatif yang kuat terhadap harga. Visualisasi antara harga aktual dan harga prediksi menunjukkan pola sebaran yang baik pada rentang \$15.000 hingga \$35.000, meskipun akurasi menurun pada harga yang sangat rendah atau sangat tinggi. Model juga menghasilkan intercept sebesar 30.246, yang berarti jika seluruh variabel bernilai nol, maka harga rumah diperkirakan sebesar \$30.246 atau sekitar 30 ribu dolar. Temuan ini menunjukkan bahwa regresi linear merupakan metode awal yang layak dan cukup efektif untuk digunakan dalam pemodelan prediktif harga rumah, meskipun masih bisa ditingkatkan dengan pendekatan model yang lebih kompleks.

Referensi

- [1] Andhika Pradifta Wicaksana, Bernadus Very Christioko, Ilham Faiq Musyaffa, Rio Eko Saputro, and Rio Eko Saputro, "Analisis Regresi Linear untuk Prediksi Harga Rumah Berdasarkan Luas Area Tanah," Universitas Semarang, 2024.

- [2] Andreas Lindholm, Niklas Wahlström, Fredrik Lindsten, and Thomas B Schön, *MACHINE LEARNING A First Course for Engineers and Scientists*. Cambridge University Press, 2022. doi: <https://doi.org/10.1017/9781108919371>.
- [3] Rafif Nauval Tuah Siregar, Vijay Sitorus, and Willy Pramudia Ananta, “Analisis Prediksi Harga Rumah di Bandung Menggunakan Regresi Linear Berganda,” Universitas Negeri Medan, 2023. Accessed: Apr. 15, 2025. [Online]. Available: <https://ejurnal.politeknikpratama.ac.id/index.php/jscr/article/download/3038/2873/8515>
- [4] Mohammad Reza Faisal and Erick Kurniawan, “ASP.NET Core Web API 6 dengan Visual Studio,” Mar. 2023, Accessed: May 31, 2025. [Online]. Available: https://www.researchgate.net/publication/369143955_ASPNET_Core_Web_API_6_dengan_Visual_Studio
- [5] Muhammad Labib Mu’tashim, Sekar Ayu Damayanti, Hanan Nadia Zaki, Toni Muhayat, and Rio Wirawan, “Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression,” UPN Veteran Jakarta, 2021. Accessed: Apr. 15, 2025. [Online]. Available: <https://ejurnal.upnvj.ac.id/informatik/article/download/3635/1498/10600>
- [6] Rosalia Roja Hallan and Ika Nur Fajri, “Prediksi Harga Rumah Menggunakan Machine Learning Algoritma Regresi Linier,” Universitas Dharma Andalas, 2025. Accessed: Apr. 15, 2025. [Online]. Available: <https://jurnal.unidha.ac.id/index.php/jtekstis/article/download/1732/953/>
- [7] Ashary Vermaysha and Nurmalaitsari, “Prediksi Harga Rumah di Kabupaten Karanganyar Menggunakan Metode Regresi Linear,” Duta Bangsa University, 2023. Accessed: Apr. 15, 2025. [Online]. Available: <https://ojs.ubd.ac.id/index.php/Senatib/article/view/2997/2178>
- [8] S. HI. , M. A. , Ciq. Dr. Karimuddin Abdullah *et al.*, *METODOLOGI PENELITIAN KUANTITATIF*. Yayasan Penerbit Muhammad Zaini, 2022. Accessed: May 12, 2025. [Online]. Available: <https://repository.ar-raniry.ac.id/id/eprint/28559/1/Buku%20Metodologi%20Penelitian%20Kuantitatif.pdf>
- [9] Sayali58, “Boston_Housing.csv,” 2025, Accessed: Apr. 15, 2025. [Online]. Available: <https://www.kaggle.com/datasets/sayali6666/boston-housing-csv>
- [10] S. S. , M. Kes. Slamet Widodo *et al.*, *BUKU AJAR METODE PENELITIAN*, 1st ed. CV SCIENCE TECHNO DIRECT, 2023. Accessed: May 12, 2025. [Online]. Available: https://repository.binawan.ac.id/3303/1/Buku%20Ajar%20Metode%20Penelitian%20Full_compressed%20Highlighted.pdf
- [11] Dr. Poornima G. Naik, *Unleashing the Power of Python Libraries for Machine Learning Excellence (Covers Tools and Techniques for Building Smarter Models)*. Evincepup Publishing, 2025. [Online]. Available: <https://www.researchgate.net/publication/387933576>
- [12] S. T. , M. T. Ir. Jarot Budiasto, S. Kom. , M. Cs. Ir. Agustan Latif, S. S. M. C. I. ASEAN. Eng. Dr. Ir. Heru Ismanto, and S. Kom. , M. T. Ir. Susanto, *PYTHON UNTUK DATA SCIENCE: Panduan Praktis Menguasai Analisis Data*. Literasi Langsung Terbit, 2025. [Online]. Available: <https://www.researchgate.net/publication/390579101>
- [13] Abdellatif M. Sadeq, *Modeling and Simulation of Combustion in Python*. Zenodo, 2024. doi: 10.5281/zenodo.14577494.
- [14] S. M. S. Pardomuan Robinson Sihombing, S. M. M. Ade Marsinta Arsani, S. M. S. E. Usep Nugraha, and S. M. A. Akhmad Mun’im, *Analisis Regresi Linier Berganda Data Panel dalam Berbagai Software*. Minhaj Pustaka, 2024. [Online]. Available: <https://www.researchgate.net/publication/387159472>
- [15] Partha Majumdar, *Neural Networks for Engineers*. Amazon Kindle Direct Publishing, 2024. [Online]. Available: <https://www.researchgate.net/publication/383646169>