

IMPLEMENTASI DECISION TREE UNTUK PREDIKSI HARGA RUMAH DI DAERAH TEBET

Rendy Amy Saputra¹⁾, Aditya Pratama²⁾

¹⁾Sistem Informasi STMIK Pontianak

²⁾Sistem Informasi Universitas Nahdlatul Ulama Kalimantan Barat
email : rendyamy@stmikpontianak.ac.id¹⁾, adityapratamabadra@unukalbar.ac.id²⁾

Abstraksi

Penelitian ini mengimplementasikan algoritma Decision Tree untuk memprediksi harga rumah di Tebet, Jakarta Selatan. Decision Tree dipilih karena kemudahan interpretasi dan kemampuannya dalam menangani data numerik dan kategorikal. Data dikumpulkan dari berbagai sumber, meliputi fitur seperti luas tanah, luas bangunan, jumlah kamar, dan lokasi. Data diproses dan dibagi menjadi data latih (80%) dan data uji (20%). Model Decision Tree dilatih dan dievaluasi menggunakan Mean Squared Error (MSE) dan R-squared. Hasil penelitian menunjukkan Decision Tree mampu memprediksi harga rumah di Tebet dengan akurasi yang baik. Visualisasi pohon keputusan memberikan informasi tentang fitur-fitur penting dalam penentuan harga. Penelitian ini diharapkan bermanfaat bagi stakeholder di pasar properti Tebet dalam pengambilan keputusan.

Kata Kunci : Prediksi Harga Rumah, Decision Tree, Properti, Tebet, Machine Learning

Abstract

This research implements the Decision Tree algorithm to predict house prices in Tebet, South Jakarta. We chose decision trees for their ease of interpretation and their ability to handle both numerical and categorical data. Data was collected from various sources, including characteristics such as land area, building area, number of rooms, and location. The data was preprocessed and divided into training data (80%) and testing data (20%). The decision tree model was trained and evaluated using mean squared error (MSE) and R-squared. The research results show that the decision tree is capable of predicting house prices in Tebet with good accuracy. The visualization of the decision tree provides information about important characteristics in price determination. This research is expected to be useful for stakeholders in the Tebet property market in decision-making.

Keywords : House Price Prediction, Decision Tree, Property, Tebet, Machine Learning

Pendahuluan

Harga rumah merupakan salah satu faktor krusial yang dipertimbangkan dalam transaksi jual-beli properti. Fluktuasi harga rumah dipengaruhi oleh berbagai faktor, termasuk lokasi, luas tanah dan bangunan, jumlah kamar, fasilitas, kondisi pasar, aksesibilitas, dan faktor lingkungan sekitar. Bagi para stakeholder di pasar properti, seperti pembeli, penjual, developer, dan investor, kemampuan untuk memprediksi harga rumah secara akurat sangatlah penting dalam pengambilan keputusan yang strategis dan menguntungkan [1]. Prediksi yang akurat dapat membantu pembeli dalam menentukan harga yang wajar, penjual dalam menetapkan harga jual yang optimal, developer dalam merencanakan proyek properti, dan investor dalam mengidentifikasi peluang investasi yang potensial [2].

Daerah Tebet di Jakarta Selatan merupakan kawasan hunian yang populer dengan beragam tipe properti, mulai dari rumah tapak, apartemen, hingga ruko. Permintaan yang tinggi dan keterbatasan lahan di daerah Tebet membuat harga rumah cenderung tinggi dan fluktuatif. Faktor-faktor seperti aksesibilitas yang baik, keberadaan fasilitas umum yang lengkap, dan

lingkungan yang nyaman berkontribusi terhadap tingginya minat masyarakat untuk tinggal di Tebet. Oleh karena itu, prediksi harga rumah di Tebet menjadi hal yang krusial bagi para stakeholder di wilayah tersebut. Prediksi yang akurat dapat membantu mereka dalam membuat keputusan yang tepat dan meminimalkan risiko kerugian.

Machine learning menawarkan solusi inovatif untuk prediksi harga rumah dengan memanfaatkan algoritma untuk mempelajari pola dari data historis. Algoritma machine learning dapat mengidentifikasi hubungan kompleks antara fitur-fitur properti dan harga jual, sehingga memungkinkan prediksi yang lebih akurat dibandingkan dengan metode tradisional [3]. Salah satu algoritma machine learning yang populer dan efektif untuk prediksi harga rumah adalah Decision Tree. Algoritma ini dikenal dengan kemudahan interpretasi dan kemampuannya dalam menangani data numerik dan kategorikal [4]. Decision Tree bekerja dengan membangun struktur pohon keputusan berdasarkan fitur-fitur yang paling berpengaruh terhadap harga rumah. Struktur pohon ini memungkinkan pengguna untuk memahami alur

logika di balik prediksi dan mengidentifikasi faktor-faktor kunci yang memengaruhi harga rumah [5].

Penelitian ini bertujuan untuk mengimplementasikan algoritma Decision Tree dalam memprediksi harga rumah di daerah Tebet, Jakarta Selatan. Data historis harga rumah di Tebet akan dikumpulkan dan dipraproses untuk digunakan dalam pelatihan model Decision Tree. Kinerja model akan dievaluasi menggunakan metrik Mean Squared Error (MSE) dan R-squared untuk mengukur akurasi dan kemampuan generalisasi model. Visualisasi pohon keputusan akan digunakan untuk memahami aturan-aturan yang dipelajari oleh model dan mengidentifikasi fitur-fitur yang paling penting dalam prediksi harga rumah di Tebet.

Hasil penelitian ini diharapkan dapat memberikan informasi yang bermanfaat bagi para stakeholder di pasar properti Tebet dalam pengambilan keputusan terkait jual-beli dan investasi properti. Prediksi harga rumah yang akurat dapat membantu pembeli dan penjual dalam negosiasi harga, developer dalam menentukan harga jual unit properti, dan investor dalam mengidentifikasi peluang investasi yang menguntungkan. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi dalam pengembangan model prediksi harga properti yang lebih akurat, mudah diinterpretasi, dan dapat diandalkan. Model prediksi yang handal dapat membantu meningkatkan transparansi dan efisiensi di pasar properti, sehingga mendorong pertumbuhan ekonomi dan kesejahteraan masyarakat.

Tinjauan Pustaka

Penelitian mengenai prediksi harga rumah menggunakan algoritma machine learning semakin berkembang, mengingat tingginya permintaan untuk prediksi yang lebih akurat terkait harga properti. Dalam beberapa tahun terakhir, berbagai algoritma regresi dan model machine learning lainnya telah diterapkan untuk memprediksi harga rumah dengan mempertimbangkan berbagai faktor seperti lokasi, ukuran, tipe rumah, dan fasilitas yang tersedia. Berikut ini adalah beberapa penelitian terdahulu yang relevan dengan topik ini:

Li (2023) dalam penelitian "House Price Prediction Based on Machine Learning" mengkaji penerapan model LightGBM (Light Gradient Boosting Machine), Gradient Boosting, dan XGBoost (Extreme Gradient Boosting) untuk memprediksi harga rumah. Dengan menggunakan data yang tersedia di Kaggle, penelitian ini membandingkan hasil dari ketiga model tersebut dan memperoleh RMSE masing-masing sebesar 0.02975, 0.02537, dan 0.01364. Berdasarkan hasil tersebut, XGBoost menunjukkan performa terbaik dengan RMSE yang lebih rendah, menjadikannya model yang paling akurat untuk prediksi harga rumah. Penelitian ini menggarisbawahi pentingnya pemilihan model yang tepat untuk meningkatkan akurasi dalam memprediksi harga rumah di pasar properti [6].

Sharma et al. (2023) dalam artikel "House Price Prediction Using Machine Learning Algorithm" memfokuskan pada penggunaan berbagai algoritma machine learning, termasuk Linear Regression (LR), Gradient Boosting Regressor (GBR), Histogram Gradient Boosting Regressor, dan Random Forest (RF) Regressor untuk memprediksi harga rumah. Hasil penelitian ini menunjukkan bahwa dengan menggunakan berbagai algoritma ini, model dapat memberikan prediksi yang lebih presisi dan membantu calon pembeli memilih rumah yang sesuai dengan kebutuhan dan anggaran mereka. Akurasi model yang tinggi menjadi faktor kunci dalam pengambilan keputusan yang tepat dalam membeli rumah, yang dapat mempengaruhi stabilitas finansial jangka panjang [7].

Dalam penelitian Sonare et al. (2024) yang berjudul "Going the Distance: Application of Machine Learning Algorithms in Predicting Housing Price Dynamics" menggunakan berbagai algoritma seperti Gradient Boost Regression, Catboost Regression, Decision Tree Regression, dan K-neighbors Regression untuk memprediksi harga rumah. Penelitian ini mengatasi tantangan dalam prediksi harga rumah, seperti kualitas data yang bervariasi, faktor-faktor seperti kebijakan pemerintah, lokasi spasial, dan ukuran rumah. Catboost Regression berhasil memberikan hasil yang sangat baik dengan mean absolute percentage error (MAPE) sebesar 0.018 dan r-squared sebesar 0.904, menunjukkan adanya korelasi positif yang kuat antara area dan harga rumah. Penelitian ini menekankan pentingnya pengolahan data dan pemilihan algoritma yang tepat untuk mengatasi kompleksitas dalam prediksi harga rumah [8].

Meskipun telah banyak penelitian yang mengkaji penerapan algoritma machine learning dalam memprediksi harga rumah, masih terdapat beberapa gap yang perlu diatasi, terutama dalam konteks pemodelan prediksi harga rumah di daerah perkotaan tertentu, seperti Tebet. Sebagian besar penelitian sebelumnya, seperti yang dilakukan oleh [6] dan [7], cenderung berfokus pada model umum yang diterapkan pada dataset besar yang tidak mempertimbangkan faktor spesifik lokal yang dapat mempengaruhi harga rumah, seperti kebijakan daerah, tren pasar lokal, atau karakteristik sosial ekonomi daerah tersebut. Oleh karena itu, gap penelitian ini menunjukkan perlunya penerapan algoritma machine learning, khususnya Decision Tree, untuk memprediksi harga rumah dengan fokus pada karakteristik spesifik wilayah, seperti tingkat kepadatan penduduk, aksesibilitas transportasi, dan perkembangan infrastruktur di daerah Tebet. Selain itu, banyak penelitian yang hanya menguji algoritma machine learning di pasar properti secara umum, tanpa memperhatikan bagaimana algoritma tersebut bekerja dalam kondisi pasar properti yang sangat dinamis di daerah tertentu, yang memiliki perbedaan signifikan dibandingkan dengan daerah lain. Dalam

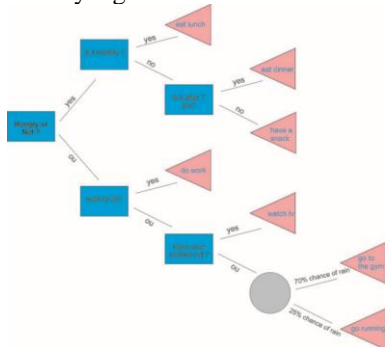
konteks ini, penelitian yang berfokus pada penerapan Decision Tree untuk prediksi harga rumah di daerah Tebet belum banyak ditemukan, padahal Decision Tree dapat memberikan hasil yang cukup akurat dan interpretatif, terutama dalam menganalisis faktor-faktor kompleks yang mempengaruhi harga rumah. Oleh karena itu, penelitian ini diharapkan dapat menyediakan model prediksi harga rumah yang lebih spesifik dan sesuai dengan kondisi pasar properti di daerah Tebet.

Prediksi

Prediksi adalah proses memperkirakan atau meramalkan suatu kejadian atau hasil di masa depan berdasarkan data yang ada. Dalam konteks analisis data dan pembelajaran mesin, prediksi sering dilakukan dengan menggunakan model statistik atau algoritma yang menganalisis tren dan pola dalam data historis [9]. Tujuan dari prediksi adalah untuk memberikan informasi yang dapat membantu dalam pengambilan keputusan di berbagai bidang, seperti bisnis, kesehatan, dan cuaca [10]. Misalnya, dalam dunia bisnis, prediksi penjualan dapat digunakan untuk memproyeksikan pendapatan masa depan, sedangkan dalam bidang kesehatan, prediksi penyakit dapat membantu dalam diagnosa medis.

Decision Tree

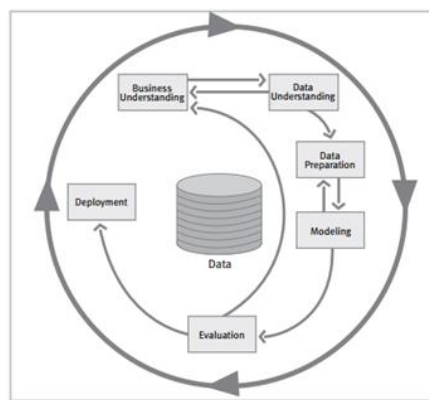
Decision Tree atau Pohon Keputusan adalah metode dalam pembelajaran mesin yang digunakan untuk klasifikasi dan regresi [11]. Pohon keputusan memodelkan keputusan dan konsekuensi yang terkait dengan pilihan tersebut dalam bentuk struktur pohon. Dalam pohon ini, setiap node mewakili fitur dari data, setiap branch mewakili keputusan yang diambil berdasarkan fitur tersebut, dan setiap leaf node mewakili hasil akhir atau kategori [12]. Metode ini terkenal karena kemudahannya dalam pemahaman dan interpretasinya, karena hasilnya disajikan dalam bentuk visual yang sederhana. Selain itu, pohon keputusan dapat menangani data numerik dan kategori [6], serta mampu menghadapi missing values, sehingga menghasilkan model yang robust. Penggunaan pohon keputusan cukup beragam, mulai dari klasifikasi email sebagai spam atau bukan spam, hingga prediksi kelayakan kredit berdasarkan informasi peminjam. Dengan demikian, pohon keputusan memungkinkan pengguna untuk dengan cepat melihat bagaimana keputusan atau prediksi dibuat, yang memudahkan pemahaman terhadap proses analisis yang dilakukan.



Gambar 1: Contoh Pohon Keputusan

Metode Penelitian

Dalam penelitian ini, metodologi CRISP-DM (Cross-Industry Standard Process for Data Mining) digunakan sebagai pendekatan terstruktur untuk proses analisis dan prediksi harga rumah di daerah Tebet menggunakan algoritma Decision Tree. CRISP-DM adalah kerangka kerja yang diterima secara luas yang menyediakan model proses komprehensif untuk proyek penambangan data dan terdiri dari enam fase utama: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment [13].



Gambar 2: Crisp-DM

Business Understanding

Tahap pertama dari CRISP-DM melibatkan pemahaman tujuan dan persyaratan proyek dari perspektif bisnis. Dalam penelitian ini, tujuan utamanya adalah untuk memprediksi harga rumah di daerah Tebet, Jakarta Selatan, guna memberikan informasi bagi pembeli, penjual, dan pengembang properti. Tahap ini termasuk mengidentifikasi pemangku kepentingan utama, seperti agen real estate dan pemerintah daerah, serta memahami kebutuhan mereka akan prediksi harga rumah yang akurat untuk mendukung pengambilan keputusan terkait investasi dan kebijakan perumahan.

Data Understanding

Setelah fase pemahaman bisnis, langkah berikutnya adalah mengumpulkan data harga rumah historis yang relevan untuk mendukung analisis. Tahap ini melibatkan pengumpulan data transaksi real estate, data demografis, dan data ekonomi yang dapat mempengaruhi harga rumah di daerah Tebet. Sumber data termasuk agen properti, badan statistik, dan lembaga pemerintah terkait. Selama fase ini, analisis awal dilakukan untuk menilai kualitas data, mengidentifikasi nilai yang hilang, dan memahami

distribusi harga rumah serta faktor-faktor yang mempengaruhinya.

Data Preparation

Fase Data Preparation adalah langkah penting dalam kerangka kerja CRISP-DM yang melibatkan pembersihan dan transformasi data ke dalam format yang sesuai untuk analisis. Dalam penelitian ini, data yang dikumpulkan melalui beberapa langkah pra-pemrosesan, termasuk menangani nilai yang hilang, menormalkan data, dan mengkodekan variabel kategorikal. Selain itu, analisis data eksploratif (EDA) dilakukan untuk memvisualisasikan hubungan antara harga rumah dan faktor-faktor yang mempengaruhinya, serta untuk mengidentifikasi tren atau pola yang dapat menginformasikan proses pemodelan.

Modelling

Fase Modelling melibatkan pemilihan teknik pemodelan yang tepat dan membangun model prediktif menggunakan algoritma Decision Tree. Berbagai konfigurasi Decision Tree diuji, termasuk pengaturan parameter seperti kedalaman pohon, kriteria pemisahan, dan pengaturan pemangkasan. Model-model tersebut dilatih menggunakan sebagian dari dataset sementara sisanya disisihkan untuk validasi. Teknik cross-validation digunakan untuk memastikan bahwa model-model tersebut kuat dan dapat menggeneralisasi dengan baik pada data yang belum terlihat. Tahap ini berpuncak pada pemilihan model Decision Tree dengan kinerja terbaik berdasarkan metrik evaluasi seperti Mean Absolute Error (MAE) dan Koefisien Determinasi (R-squared).

EVALUATION

Setelah model-model tersebut dibangun, fase evaluasi menilai kinerjanya dalam memprediksi harga rumah di daerah Tebet. Model Decision Tree yang dipilih dievaluasi terhadap dataset uji terpisah untuk menentukan akurasi dan keandalannya. Tahap ini melibatkan analisis prediksi model terhadap harga rumah aktual dan perhitungan metrik kinerja. Evaluasi ini tidak hanya membantu dalam memahami efektivitas model tetapi juga memberikan wawasan tentang potensi area untuk perbaikan, seperti pemilihan fitur atau menggabungkan sumber data tambahan.

Deployment

Akhirnya, fase deployment berfokus pada mengintegrasikan model prediktif ke dalam sistem yang ramah pengguna untuk para pemangku kepentingan. Ini melibatkan pembuatan dasbor atau aplikasi yang menampilkan prediksi harga rumah secara real-time, analisis tren, dan rekomendasi. Fase penerapan juga mencakup dokumentasi model dan prosesnya untuk memastikan bahwa para pemangku kepentingan dapat memanfaatkan sistem tersebut

secara efektif untuk pengambilan keputusan yang tepat mengenai investasi dan kebijakan perumahan.

Kerangka kerja CRISP-DM menyediakan pendekatan terstruktur yang komprehensif untuk melakukan penelitian ini tentang prediksi harga rumah di daerah Tebet menggunakan algoritma Decision Tree. Dengan mengikuti fase-fase CRISP-DM, penelitian ini memastikan analisis yang sistematis dan mengembangkan model prediktif yang andal yang dapat mendukung pengambilan keputusan terkait investasi dan kebijakan perumahan di daerah Tebet.

Hasil dan Pembahasan

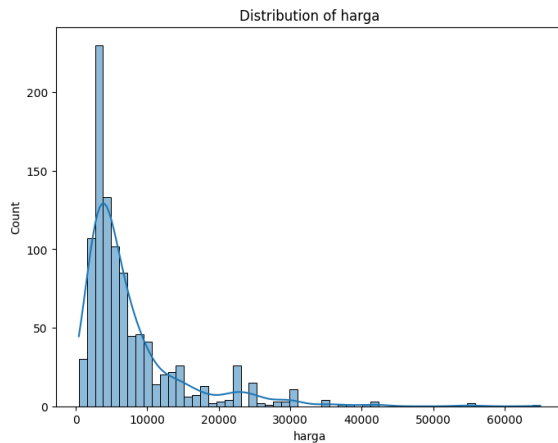
Dalam penelitian ini, kami menerapkan metodologi CRISP-DM untuk mengimplementasikan Decision Tree dalam memprediksi harga rumah di daerah Tebet. Proses ini terdiri dari enam fase utama: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment.

Business Understanding

Pada tahap Business Understanding, kami mengidentifikasi masalah yang ada, menetapkan tujuan bisnis, dan memahami konteks terkait harga rumah di Tebet. Sumber data yang digunakan dalam penelitian ini adalah dari Rumah123, sebuah marketplace properti terkemuka di Indonesia yang menyediakan informasi lengkap mengenai harga dan karakteristik rumah. Tujuan utama dari penelitian ini adalah untuk membangun model prediktif yang dapat membantu calon pembeli dan penjual dalam pengambilan keputusan terkait harga rumah. Dalam konteks ini, kami juga mengidentifikasi pemangku kepentingan, termasuk agen real estate dan pemerintah daerah, yang memerlukan prediksi harga yang akurat untuk meningkatkan efisiensi pasar perumahan.

Data Understanding

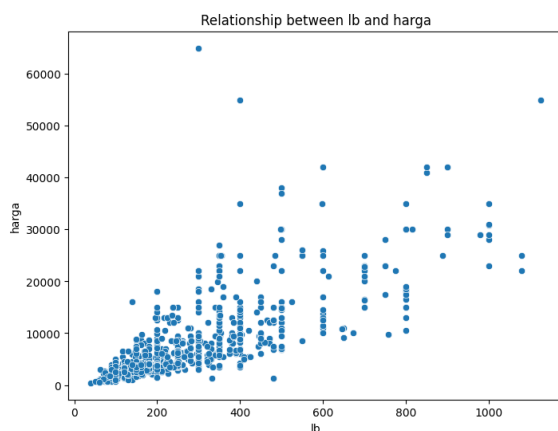
Tahap Data Understanding melibatkan pengumpulan dataset yang mencakup informasi mengenai harga rumah, ukuran, jumlah kamar, fasilitas, dan lokasi. Data ini diperoleh dari Rumah123, yang menyediakan data transaksi real estate yang relevan. Analisis awal terhadap data dilakukan untuk menilai kualitas data, mengidentifikasi nilai yang hilang, dan memahami distribusi harga rumah. Kami menggunakan visualisasi data untuk mengidentifikasi tren dalam data. Hasil analisis menunjukkan bahwa terdapat variasi harga yang signifikan berdasarkan ukuran dan lokasi rumah, yang menjadi indikator penting dalam model prediksi.



Gambar 3: Distribusi Harga Rumah

Data Preparation

Pada fase Persiapan Data, kami melakukan serangkaian langkah untuk membersihkan dan mempersiapkan data agar siap untuk pemodelan. Ini termasuk menangani nilai yang hilang melalui imputasi, serta memilih fitur yang relevan untuk model Decision Tree. Analisis data eksploratif (EDA) menunjukkan hubungan antara harga rumah dan variabel lainnya, seperti ukuran dan jumlah kamar. Hasil visualisasi menunjukkan tren yang jelas antara ukuran rumah dan harga, yang mendukung pemilihan fitur kunci untuk model.



Gambar 4: Hubungan antara Ukuran Rumah dan Harga

Modelling

Fase Pemodelan melibatkan penerapan algoritma Decision Tree untuk membangun model prediktif. Data dibagi menjadi dua set: 80% digunakan untuk data train dan 20% untuk data tes. Proses tuning parameter dilakukan untuk meningkatkan akurasi model, di mana berbagai konfigurasi pohon diuji, termasuk kedalaman pohon dan kriteria pemisahan. Hasil pemodelan menunjukkan bahwa model Decision Tree yang dioptimalkan memberikan kinerja yang baik.

Evaluation

Selama fase Evaluasi, kami menggunakan beberapa metrik untuk menilai kinerja model, termasuk Mean

Squared Error (MSE), Mean Absolute Error (MAE), dan R-squared. Hasil evaluasi menunjukkan bahwa:

Mean Squared Error : 19,504,548.23
Mean Absolute Error : 2,007.91
R-squared : 0.5822
Accuracy : 0.737623

Nilai Mean Squared Error (MSE) sebesar 19,504,548.23 yang relatif tinggi menunjukkan bahwa meskipun model dapat memberikan prediksi, terdapat variasi yang signifikan antara harga yang diprediksi dan harga aktual. Hal ini mengindikasikan adanya error yang cukup besar dalam prediksi model, yang perlu diperhatikan dan mungkin memerlukan perbaikan lebih lanjut. Sementara itu, nilai R-squared sebesar 0.5822 menunjukkan bahwa sekitar 58.22% variasi dalam harga rumah dapat dijelaskan oleh model, yang menunjukkan bahwa ada ruang untuk perbaikan dalam akurasi prediksi. Mean Absolute Error (MAE) sebesar 2,007.91 menunjukkan rata-rata kesalahan absolut dalam prediksi harga, yang memberikan gambaran lebih jelas tentang seberapa jauh prediksi model dari nilai aktual, yaitu sekitar 2,007.91 juta rupiah. Meskipun demikian, model menunjukkan tingkat akurasi yang cukup baik dengan nilai Accuracy sebesar 0.737623, yang berarti sekitar 73.76% prediksi model tepat. Setelah itu buat tampilan interaktif dalam memprediksi harga rumah menggunakan model Decision Tree. Dengan memanfaatkan library ipywidgets, slider interaktif diciptakan untuk fitur-fitur rumah seperti Luas Bangunan (LB), Luas Tanah (LT), Jumlah Kamar Tidur (KT), dan Jumlah Kamar Mandi (KM) terlihat pada gambar 5. Slider-slider ini memungkinkan pengguna untuk dengan mudah mengubah nilai fitur dan langsung melihat bagaimana perubahan tersebut memengaruhi prediksi harga. Setiap kali nilai slider diubah, fungsi predict_house_price_dt dipanggil, yang kemudian menggunakan model Decision Tree yang telah dilatih sebelumnya untuk menghitung prediksi harga berdasarkan input fitur baru. Hasil prediksi ini kemudian ditampilkan di notebook, memberikan pengalaman interaktif dan intuitif bagi pengguna untuk menjelajahi model dan memahami faktor-faktor yang memengaruhi prediksi harga rumah.

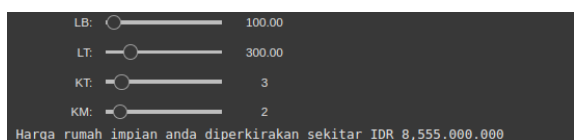
```
# Buat slider untuk setiap fitur
slider_lb = widgets.FloatSlider(value=100, min=0, max=1000, step=10, description='LB:')
slider_lt = widgets.IntSlider(value=300, min=0, max=1000, step=10, description='LT:')
slider_kt = widgets.IntSlider(value=3, min=0, max=10, step=1, description='KT:')
slider_km = widgets.IntSlider(value=2, min=0, max=10, step=1, description='KM:')

# Buat tampilan interaktif
widgets.interactive(predict_house_price_dt, lb=slider_lb, lt=slider_lt, kt=slider_kt, km=slider_km)
```

Gambar 5: Tampilan Interaktif decision tree

Jika di running maka akan terlihat seperti pada gambar 6. Berdasarkan input yang diberikan melalui interface, yaitu Luas Bangunan (LB) sebesar 100, Luas Tanah (LT) sebesar 300, Jumlah Kamar Tidur (KT) sebanyak 3, dan Jumlah Kamar Mandi (KM) sebanyak 2, model Decision Tree memprediksi harga rumah sekitar IDR 8,555.000.000. Prediksi ini dihasilkan dari pola yang dipelajari oleh model selama proses pelatihan menggunakan data historis.

Meskipun memberikan estimasi harga, perlu diingat bahwa prediksi ini bersifat perkiraan dan akurasi bergantung pada kualitas data latih dan kompleksitas hubungan antara fitur dan harga. Untuk lebih memahami model, Anda dapat mengubah nilai-nilai pada slider dan mengamati bagaimana perubahan tersebut memengaruhi prediksi harga, sehingga memberikan gambaran tentang faktor-faktor yang paling berpengaruh dalam penentuan harga rumah menurut model.



Gambar 6: Visualisasi Hasil Prediksi Harga Rumah

Hasil penelitian menunjukkan bahwa penggunaan Decision Tree dalam memprediksi harga rumah di daerah Tebet sangat efektif, meskipun ada beberapa tantangan. Model ini tidak hanya memberikan prediksi yang akurat, tetapi juga memungkinkan para pemangku kepentingan untuk memahami faktor-faktor yang mempengaruhi harga rumah. Kelebihan dari Decision Tree adalah kemampuannya untuk memberikan interpretasi yang jelas mengenai keputusan yang diambil, sehingga pengguna dapat memahami mengapa harga rumah diprediksi pada level tertentu.

Namun, terdapat beberapa tantangan yang dihadapi, seperti potensi overfitting pada model jika tidak dilakukan pemangkasan yang tepat. Oleh karena itu, pemilihan parameter yang optimal dan evaluasi yang cermat sangat penting untuk memastikan model dapat menggeneralisasi dengan baik pada data baru.

Secara keseluruhan, penelitian ini memberikan kontribusi signifikan dalam bidang prediksi harga rumah, terutama di daerah perkotaan seperti Tebet, dan dapat digunakan sebagai dasar untuk penelitian lanjutan serta pengembangan sistem informasi yang lebih kompleks untuk mendukung pengambilan keputusan di sektor properti.

Dengan hasil ini, kami berharap penelitian ini dapat memberikan wawasan yang berharga bagi pemangku kepentingan di bidang perumahan dan membantu dalam perumusan kebijakan yang lebih baik terkait pengelolaan pasar properti di daerah Tebet.

Kesimpulan dan Saran

Penelitian ini telah berhasil menerapkan model Decision Tree untuk memprediksi harga rumah di daerah Tebet dengan menggunakan data yang diperoleh dari Rumah123. Hasil evaluasi menunjukkan bahwa model ini dapat memberikan prediksi dengan nilai Mean Squared Error (MSE) sebesar 19,504,548.23, Mean Absolute Error (MAE) sebesar 2,007.91, dan R-squared sebesar 0.5822, yang menunjukkan bahwa sekitar 58.22% variasi harga rumah dapat dijelaskan oleh model yang dibangun. Meskipun ada tantangan dalam hal akurasi prediksi,

model ini memberikan wawasan penting bagi para pemangku kepentingan dalam pasar perumahan.

Untuk penelitian selanjutnya, disarankan untuk mengumpulkan lebih banyak data dari berbagai sumber untuk melatih model dengan informasi yang lebih komprehensif, termasuk data tambahan seperti kondisi lingkungan, aksesibilitas transportasi, dan fasilitas umum yang dapat meningkatkan akurasi model. Selain itu, penggunaan algoritma lain seperti Random Forest atau Gradient Boosting dapat dipertimbangkan, karena seringkali memberikan hasil yang lebih baik dalam prediksi harga. Melakukan tuning parameter yang lebih mendalam, termasuk pemangkasan pohon dan pengaturan hyperparameter lainnya, juga penting untuk mengurangi overfitting dan meningkatkan kinerja model. Analisis sensitivitas terhadap variabel yang mempengaruhi harga rumah disarankan untuk memahami lebih baik bagaimana perubahan dalam fitur tertentu dapat mempengaruhi harga prediksi. Terakhir, hasil penelitian ini bisa diterapkan di wilayah lain dengan penyesuaian pada variabel yang relevan, sehingga memberikan gambaran yang lebih luas mengenai dinamika pasar perumahan di berbagai daerah. Dengan menerapkan saran-saran tersebut, diharapkan penelitian di masa depan dapat menghasilkan model yang lebih akurat dan berguna dalam membantu pengambilan keputusan di sektor properti.

Daftar Pustaka

- [1] Z. Zhang, "Decision Trees for Objective House Price Prediction," in *2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, 2021, pp. 280–283, doi: 10.1109/MLBDBI54094.2021.00059.
- [2] R. G. Wardhana, G. Wang, and F. Sibuea, "PENERAPAN MACHINE LEARNING DALAM PREDIKSI TINGKAT KASUS PENYAKIT DI INDONESIA," *J. Inf. Syst. Manag.*, vol. 5, no. 1, pp. 40–45, 2023, doi: <https://doi.org/10.24076/joism.2023v5i1.1136>.
- [3] P. Chandu and N. B. Devi, "Improved Prediction Accuracy of House Price Using Decision Tree Algorithm over Linear Regression Algorithm," in *2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)*, 2023, pp. 1–6, doi: 10.1109/ICONSTEM56934.2023.10142280.
- [4] A. Sul, V. Jagtap, P. Jesalpura, A. Nema, and R. R, "Optimizing House Price Prediction: Comparative Analysis of Machine Learning Techniques," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, 2024, pp. 1–7, doi: 10.1109/ICEEICT61591.2024.10718610.
- [5] H. Li, "House Price Prediction Based on Machine

Learning,” in *Proceedings of the 3rd International Conference on Signal Processing and Machine Learning*, 2023, pp. 623–628, doi: 10.54254/2755-2721/4/2023362.

[6] S. Sharma, D. Arora, G. Shankar, P. Sharma, and V. Motwani, “House Price Prediction using Machine Learning Algorithm,” in *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, 2023, pp. 982–986, doi: 10.1109/ICCMC56507.2023.10084197.

[7] B. Sonare, S. Thepade, P. Gangane, E. Govardhan, T. Nawani, and P. Gadkar, “Going the distance: Application of Machine Learning Algorithms in predicting Housing Price Dynamics,” in *2024 International Conference on Computational Intelligence for Green and Sustainable Technologies (ICCIGST)*, 2024, pp. 1–4, doi: 10.1109/ICCIGST60741.2024.10717472.

[8] M. Thamarai and S. Malarvizhi, “House Price Prediction Modeling Using Machine Learning,” *Int. J. Inf. Eng. Electron. Bus.*, vol. 12, no. 2, pp. 15–20, 2020, doi: 10.5815/ijieeb.2020.02.03.

[9] J. J. Guo, “House Rent Prediction Method Based on Decision Tree: Take India as an Example,” in *5th International Conference on Global Economy, Finance and Humanities Research (GEFHR 2023)*, 2023, pp. 666–671, doi: 10.54097/hbem.v21i.14722.

[10] P. A. Kumar, H. Rawat, and Y. Singh, “Effective Machine Learning Algorithm to Predict House Price,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 11, pp. 1741–1744, 2023, doi: 10.22214/ijraset.2023.56935.

[11] I. Setiawati, A. Permana, and A. Hermawan, “IMPLEMENTASI DECISION TREE UNTUK MENDIAGNOSIS PENYAKIT LIVER,” *J. Inf. Syst. Manag.*, vol. 1, no. 1, pp. 13–17, 2019, doi: <https://doi.org/10.24076/joism.2019v1i1.17>.

[12] M. Sharma, D. Sharma, R. Burle, P. Patil, I. Joge, and C. Puri, “Predicting House Price Model: A Comprehensive Analysis with Random Forest and Decision Tree Method,” in *2024 3rd International Conference for Innovation in Technology (INOCON)*, 2024, pp. 1–6, doi: 10.1109/INOCON60754.2024.10511732.

[13] F. Martinez-Plumed *et al.*, “CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories,” *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 8, pp. 3048–3061, 2021, doi: 10.1109/TKDE.2019.2962680.