

PENERAPAN *MACHINE LEARNING* MENGGUNAKAN ALGORITMA *RANDOM FOREST* UNTUK PREDIKSI HARGA MOBIL BEKAS

Belinda Eka Sarah Dewi¹, Sandy Haikal², H.S. Sulistyowati³, Rina Fitriani⁴,
Domo Pranowo K⁵

¹Universitas Bani Saleh, belindaesd@ubs.ac.id

²Universitas Bani Saleh, SANDYHAIKAL226@ubs.ac.id

³Universitas Bani Saleh, sulis@ubs.ac.id

⁴Universitas Bani Saleh, rina@ubs.ac.id

⁵Universitas Bani Saleh, domo@ubs.ac.id

ABSTRAK

Perkembangan teknologi informasi telah membawa dampak signifikan dalam berbagai sektor industri, termasuk sektor otomotif. Salah satu aspek yang semakin menarik perhatian adalah prediksi harga mobil bekas. Dalam studi ini, algoritma *Random Forest* digunakan untuk memprediksi harga mobil bekas berdasarkan data historis. Faktor-faktor seperti model mobil, tahun produksi, kapasitas mesin, jarak tempuh, transmisi, bahan bakar, dan wilayah dimasukkan ke dalam model untuk meningkatkan akurasi prediksi. Aplikasi prediksi ini dikembangkan dalam bentuk *WebApp*, memudahkan pengguna untuk memasukkan parameter mobil dan memperoleh hasil prediksi. Hasil pemodelan menunjukkan bahwa model mampu menjelaskan 90% variabilitas harga berdasarkan input pengguna. Dengan demikian, model ini dapat menjadi acuan berguna bagi penjual dan pembeli mobil bekas dalam menetapkan harga yang sesuai dengan kondisi pasar. Aplikasi ini memberikan wawasan tentang faktor-faktor yang memengaruhi harga mobil bekas dan potensinya dalam membantu pengambilan keputusan. Teknologi *Machine Learning* terbukti efektif dalam menangani kompleksitas pasar otomotif yang dinamis, memberikan prediksi harga yang lebih akurat dibandingkan metode tradisional.

Kata Kunci: *Random Forest*, *Machine Learning*, Prediksi Harga, Mobil Bekas, *E-Commerce*

ABSTRACT

The development of information technology has had a significant impact on various industrial sectors, including the automotive sector. One aspect that is increasingly attracting attention is the prediction of used car prices. This study employs the Random Forest algorithm to predict used car prices using historical data, incorporating factors such as car model, year of production, engine capacity, mileage, transmission, fuel type, and region to enhance prediction accuracy. The prediction tool is implemented as a WebApp, facilitating user input of car parameters to obtain price predictions. These metrics suggest that the model accounts for 90% of the variability in prices based on user inputs, making it a valuable resource for sellers and buyers in aligning prices with market conditions. This application offers a detailed understanding of factors impacting used car prices and their utility in informed decision-making. Machine Learning technology effectively addresses the complexities of the dynamic automotive market, delivering more accurate price predictions than traditional methods.

Keywords: *Random Forest*, *Machine Learning*, Price Prediction, Used Car, *E-Commerce*

PENDAHULUAN

Perkembangan teknologi informasi telah membawa dampak signifikan dalam berbagai sektor industri, termasuk sektor otomotif. Salah satu aspek yang semakin menarik perhatian adalah prediksi harga mobil bekas. Penentuan harga mobil bekas merupakan proses kompleks yang dipengaruhi oleh berbagai faktor, seperti merek, model, tahun pembuatan, kondisi kendaraan, jumlah kilometer yang ditempuh, hingga lokasi penjualan. Keterbatasan metode konvensional dalam menilai harga secara akurat telah mendorong adopsi teknologi berbasis *machine learning* sebagai solusi yang lebih efektif dan efisien.

Algoritma *Random Forest* merupakan salah satu teknik *machine learning* yang banyak digunakan dalam masalah prediksi, terutama karena kemampuannya menangani data yang bersifat non-linear dan memiliki banyak variabel. Algoritma ini bekerja dengan membangun sejumlah pohon keputusan (*decision trees*) dan menggabungkan hasil prediksinya untuk menghasilkan estimasi yang lebih akurat. Selain itu, *Random Forest* juga memiliki kemampuan untuk mengurangi risiko *overfitting*, yang sering menjadi masalah dalam algoritma prediktif lainnya.

Dalam konteks prediksi harga mobil bekas, penerapan algoritma *Random Forest* dapat membantu dealer, konsumen, maupun platform jual beli online dalam menentukan harga pasar yang lebih objektif dan transparan. Dengan adanya model prediktif yang akurat, konsumen bisa mendapatkan harga yang adil, sedangkan dealer atau penjual dapat mengoptimalkan keuntungan mereka tanpa mengurangi daya tarik produk.

Penelitian ini bertujuan untuk menerapkan algoritma *Random Forest* dalam proses prediksi harga mobil bekas dan mengevaluasi kinerja model yang dibangun berdasarkan metrik akurasi tertentu, seperti *mean absolute error* (MAE) dan *root mean square error* (RMSE). Adapun data yang digunakan dalam penelitian ini mencakup variabel-variabel penting yang relevan dengan harga mobil bekas, termasuk usia mobil, jarak tempuh, dan kondisi kendaraan.

Dengan adanya penelitian ini, diharapkan dapat memberikan kontribusi dalam bentuk model prediksi yang handal dan dapat diterapkan oleh berbagai pihak di industri otomotif, sekaligus memberikan wawasan bagi pengembangan metode prediksi harga di masa mendatang.

TINJAUAN PUSTAKA

Berikut adalah beberapa referensi yang berkaitan dengan objek pembahasan pada penelitian ini, diantaranya adalah :

- a. *Using Linear Regression For Used Car Price Predict* (MUTI & YILDIZ, 2023)
Penelitian ini menunjukkan bahwa model regresi linear efektif memprediksi harga mobil bekas dengan akurasi memadai. Faktor seperti tahun pembuatan, kilometer tempuh, dan jenis bahan bakar signifikan mempengaruhi harga. Regresi linear terbukti sebagai pendekatan sederhana namun efektif dalam prediksi ini.
- b. *A Review on Machine Learning Algorithms* (Raj, 2019)
Tinjauan ini memberikan pemahaman mendalam tentang berbagai algoritma *Machine Learning*, dengan fokus pada aplikasi mereka dalam konteks yang beragam. Tinjauan algoritma seperti *Decision Trees*, *Neural Networks*, dan *Support Vector Machines* memberikan pemahaman yang komprehensif tentang kerangka kerja dasar yang digunakan dalam pembuatan model prediktif.
- c. *Comparison of Random Forest and XGBoost Models for Predicting Crime Rates in*

West Java Province (Khairunnisa, 2023)

Penelitian ini membandingkan kinerja model *Random Forest* dan *XGBoost* dalam memprediksi tingkat kejahatan di Provinsi Jawa Barat. Penelitian ini menunjukkan bahwa integrasi pendekatan *Random Forest Regression* dan *Cosine Similarity* dapat

efektif dalam analisis properti.

d. *Vehicle Price Classification and Prediction Using Machine Learning in the IoT Smart Manufacturing Era* (Al-Turjman et al., 2022)

Penelitian ini bertujuan untuk mengklasifikasikan dan memprediksi harga kendaraan menggunakan teknik *Machine Learning* dalam era manufaktur pintar *Internet of Things* (IoT). Para peneliti menggunakan dataset yang mencakup berbagai atribut kendaraan seperti merek, model, tahun pembuatan, dan spesifikasi lainnya. Mereka menerapkan berbagai algoritma *Machine Learning* seperti *Decision Trees*, *Random Forest*, dan *Gradient Boosting* untuk melakukan klasifikasi dan prediksi harga kendaraan. Hasil eksperimen menunjukkan bahwa model *Machine Learning* yang dikembangkan mampu memberikan prediksi harga kendaraan yang akurat. Penelitian ini memberikan kontribusi penting dalam memanfaatkan teknologi IoT dan *Machine Learning* untuk meningkatkan proses manufaktur dan pengambilan keputusan di industri otomotif.

METODE PENELITIAN

Pada penelitian ini menggunakan metode *Random Forest* untuk melakukan prediksi harga mobil bekas. Berikut ini merupakan langkah-langkah penelitian yang dilakukan.

a. Pemilihan Dataset

Proses pemilihan *dataset* atau pengumpulan data untuk penelitian ini dilakukan melalui teknik *web scraping*. *Web scraping* adalah metode otomatis untuk mengekstrak data dari situs *web*. Dalam konteks penelitian ini, *web scraping* digunakan untuk mengumpulkan informasi mengenai mobil bekas yang dijual di platform *e-commerce* OLX.COM di Indonesia, langkah-langkah untuk melakukan *web scraping*:

1. Identifikasi Target

Mengidentifikasi platform *e-commerce* yang akan menjadi sumber data. Dalam hal ini, platform *e-commerce* OLX.COM.

2. Pemilihan alat

Pemilihan *library* yang sesuai untuk melakukan *web scraping* seperti *Python* (*Beautiful Soup*, *pandas*) dan *Selenium*.

3. Konfigurasi

Membuat konfigurasi yang sesuai untuk mengekstrak informasi yang relevan dari halaman *web* target. Informasi ini biasanya meliputi Model mobil, Tahun produksi, Bahan bakar, Jarak tempuh, Transmisi, Harga, Wilayah, Kapasitas mesin.

4. Pengumpulan Data

Konfigurasi yang telah dibuat kemudian dijalankan untuk melakukan *scraping* pada halaman *web* target. Data yang berhasil diekstrak disimpan dalam format *CSV*.

Gambar 1 merupakan hasil *dataset* yang digunakan dalam penelitian ini terdiri dari 7656 sampel data mobil bekas yang dijual di platform OLX.COM. *Dataset* ini memiliki 8 kolom atau fitur yang berhubungan dengan karakteristik mobil dan informasi penjualan, yaitu model mobil, tahun produksi, bahan bakar, jarak tempuh, transmisi, harga, wilayah, dan kapasitas mesin.

	model_mobil	tahun_produksi	bahan_bakar	jarak_tempuh	transmisi	harga	wilayah	kapasitas_mesin
0	0	2019	0	17.5	1	150000000	19	1.4
1	0	2019	0	37.5	1	149000000	19	1.4
2	0	2018	0	37.5	0	150000000	19	1.4
3	0	2023	0	37.5	0	150000000	19	1.4
4	0	2017	0	37.5	1	150000000	19	1.4
...
7650	11	2017	0	202.5	0	103000000	11	1.2
7651	11	2017	0	82.5	1	100000000	11	1.2
7652	11	2021	0	82.5	0	100000000	11	1.2
7653	11	2021	0	87.5	1	100000000	11	1.2
7654	11	2019	0	142.5	0	100000000	11	1.2

7655 rows x 8 columns

Gambar 1. Dataset

b. Pra-Pengolahan Data

Tahap pra-pemrosesan data bertujuan untuk mempersiapkan data agar siap digunakan dalam pemodelan. Langkah-langkah yang dilakukan meliputi deteksi dan penanganan *outlier*. Deteksi *outlier* dilakukan dengan menggunakan metode *Interquartile Range* (IQR) pada kolom jarak tempuh. *Outlier* didefinisikan sebagai nilai yang berada di bawah batas bawah ($Q1 - 1.5 * IQR$) atau di atas batas atas ($Q3 + 1.5 * IQR$), dimana Q1 dan Q3 adalah kuartil bawah dan atas. Nilai *outlier* kemudian dihapus untuk menghindari gangguan pada kinerja model prediksi dan meningkatkan akurasi hasil prediksi.

c. Pemodelan *Random Forest*

Model prediksi harga mobil bekas dikembangkan menggunakan algoritma *Random Forest* karena kemampuannya dalam menangani data yang kompleks, tidak memerlukan asumsi linearitas, dan memiliki interpretabilitas yang baik dengan membagi dataset menjadi data latih (80%) dan data uji (20%).

Model *Random Forest* dilatih dengan 200 pohon ($n_estimators=200$) menggunakan data latih dan dievaluasi kinerjanya pada data uji menggunakan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan *R-squared* (R^2).

Mean Absolute Error (MAE), yang dirumuskan sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

RMSE digunakan untuk mengukur kualitas *split*. Rumus RMSE adalah sebagai berikut:

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

Dan untuk koefisien determinasi *R-squared* (R^2), adalah sebagai berikut:

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

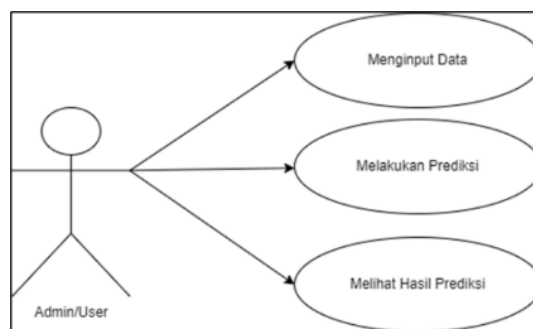
Keterangan:

- i. y_i adalah nilai aktual
- ii. \hat{y}_i adalah nilai prediksi dari model
- iii. \bar{y} adalah rata-rata dari nilai aktual
- iv. n adalah jumlah pengamatan

d. Rancangan UML

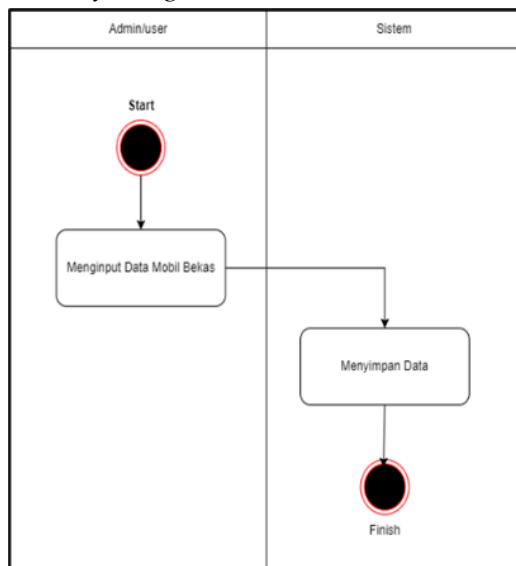
Rancangan *user interface* (UI) ini menjelaskan tentang bisnis proses yang akan dijalankan dan gambaran layout aplikasi yang akan dibuat. Adapun diagram yang digunakan yaitu *use case diagram*, *activity diagram*, dan *sequence diagram*.

1. Use Case Diagram

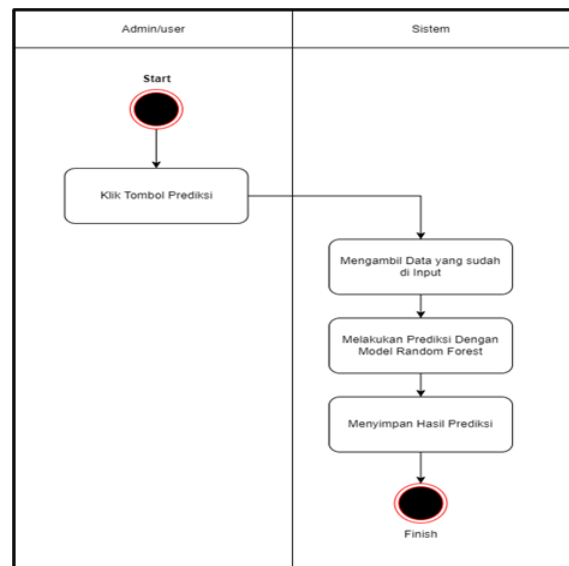


Gambar 2. Use Case Diagram Admin

2. Activity Diagram

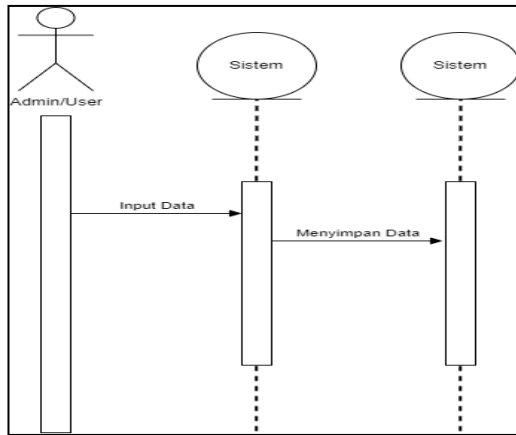


Gambar 3. Activity Diagram Input Data

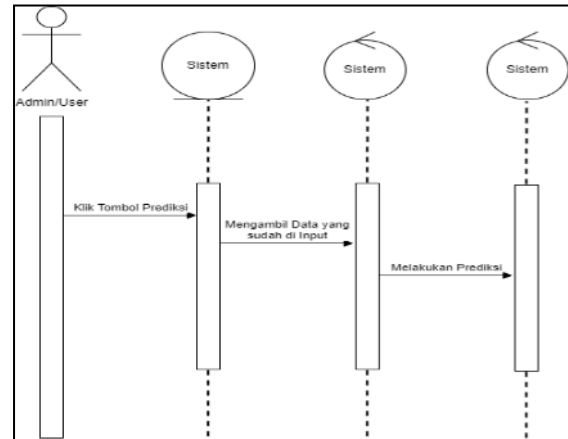


Gambar 4. Activity Diagram Proses Prediksi

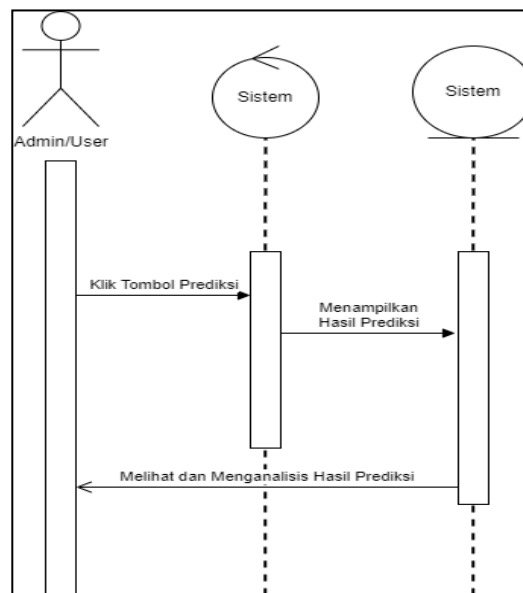
3. Sequence Diagram



Gambar 5. *Sequence Diagram Input Data*



Gambar 6. *Sequence Diagram Proses Prediksi*



Gambar 7. *Sequence Diagram Hasil Prediksi*

HASIL DAN PEMBAHASAN

Encoding Fitur Kategorikal

Dataset ini memiliki beberapa variabel kategorikal seperti model_mobil, bahan_bakar, transmisi, dan wilayah. *Model machine learning*, seperti *Random Forest*, hanya dapat bekerja dengan data *numerik*, sehingga variabel-variabel ini perlu diubah menjadi bentuk *numerik*. Hasil *encoding* fitur kategorikal ditunjukkan pada Tabel 1, Tabel 2, Tabel 3, dan Tabel 4.

Tabel 1. Model Mobil Numerik

Model Mobil	Numerik
daihatsu ayla	0
daihatsu sigra	1
daihatsu xenia	2
honda brio	3
honda brio satya	4
honda mobilio	5
suzuki ertiga	6
suzuki ignis	7
toyota agya	8
toyota avanza	9
toyota calya	10
wuling cortez	11

Tabel 2. Bahan Bakar Numerik

Bahan Bakar	Numerik
Bensin	0
Diesel	1

Tabel 3. Transmisi Numerik

Transmisi	Numerik
Automatic	0
Manual	1

Tabel 4. Wilayah Numerik

Wilayah	Numerik	Wilayah	Numerik
Bandung Kab.	0	Karawang Kota	11
Bandung Kota	1	Malang Kab.	12
Bekasi Kab.	2	Malang Kota	13
Bekasi Kota	3	Medan Kota	14
Bogor Kab.	4	Semarang Kab.	15
Bogor Kota	5	Semarang Kota	16
Depok Kota	6	Sukabumi Kab.	17
Jakarta Barat	7	Surabaya Kota	18
Jakarta Pusat	8	Tangerang Kab.	19
Jakarta Selatan	9	Tangerang Kota	20
Jakarta Timur	10	Yogyakarta Kota	21

Explanatory Data Analysis (EDA)

EDA dilakukan untuk mendapatkan gambaran awal tentang distribusi data, mengidentifikasi *outlier*, dan menemukan pola-pola menarik. Tabel 5 menggambarkan hasil EDA dengan data awal.

Tabel 5. Hasil *Explanatory Data Analysis*

	Model Mobil	Tahun Produksi	Bahan Bakar	Jarak Tempuh	Transmisi	Wilayah	Kapasitas Mesin	Harga
Mean	5.619	2019.731	0.084	88.709	0.545	9.220	1.265	129554091.57
Median	6.000	2019.000	0.000	92.500	1.000	8.000	1.250	133000000.00
Modus	8.000	2019.000	0.000	122.500	1.000	4.000	1.200	145000000.00
Standar Deviasi	3.363	1.791	0.278	55.763	0.498	6.416	0.094	14620893.99
Min	0.000	2017.000	0.000	0.000	0.000	0.000	1.200	100000000.00
Max	11.000	2024.000	1.000	202.500	1.000	22.000	1.500	150000000.00

Data Preprocessing

Tahap ini bertujuan untuk mempersiapkan data agar siap digunakan dalam pemodelan. Beberapa langkah pra-pemrosesan yang dilakukan meliputi:

1. Deteksi *Outlier*

Mengidentifikasi *outlier* yang dapat mengganggu kinerja model. Untuk mendeteksi *Outlier* menggunakan rumus:

$$IQR = Q3 - Q1$$

$$Batas\ Bawah = Q1 - 1.5 \times IQR$$

$$Batas\ Atas = Q3 + 1.5 \times IQR$$

Deteksi *outlier* untuk kolom jarak tempuh

$$IQR = 122.5 - 37.5 = 85$$

$$Batas\ Bawah = 37.5 - 1.5 \times 85 = -90$$

$$Batas\ Atas = 122.5 + 1.5 \times 85 = 250$$

Jika nilai dari kolom jarak tempuh berada dibawah -90 dan diatas 250 maka nilai tersebut akan dinyatakan sebagai *outlier*.

2. Penanganan *Outlier*

Menghapus *Outlier* yang dapat mengganggu kinerja model prediksi. Sehingga menjadikan *dataset* yang akan digunakan terbebas dari *outlier* yang dapat mengganggu akurasi dari hasil prediksi. *Dataset* yang sudah tidak memiliki *outlier* berjumlah 6910 data.

Pengembangan Model Random Forest

Algoritma *Random Forest* dipilih karena kemampuannya dalam menangani *data* yang kompleks, tidak memerlukan asumsi linearitas, dan memiliki interpretabilitas yang baik. Langkah-langkah pengembangan model meliputi:

1. Pembagian *data* membagi dataset menjadi *data training* (80%) dan *data test* (20%).
2. Pelatihan model melatih model *Random Forest* pada data latih dengan menentukan jumlah pohon sebanyak 200 pohon ($n_estimators=200$)

3. Evaluasi model mengevaluasi kinerja model pada data uji menggunakan metrik seperti *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, dan *R-squared (R²)*. Metrik-metrik ini memberikan gambaran mengenai seberapa baik model dalam memprediksi harga berdasarkan data yang ada. Pada Tabel 6 ditunjukkan hasil metrik evaluasi dari model yang sudah dikembangkan.

Tabel 6. Hasil Metrik Evaluasi

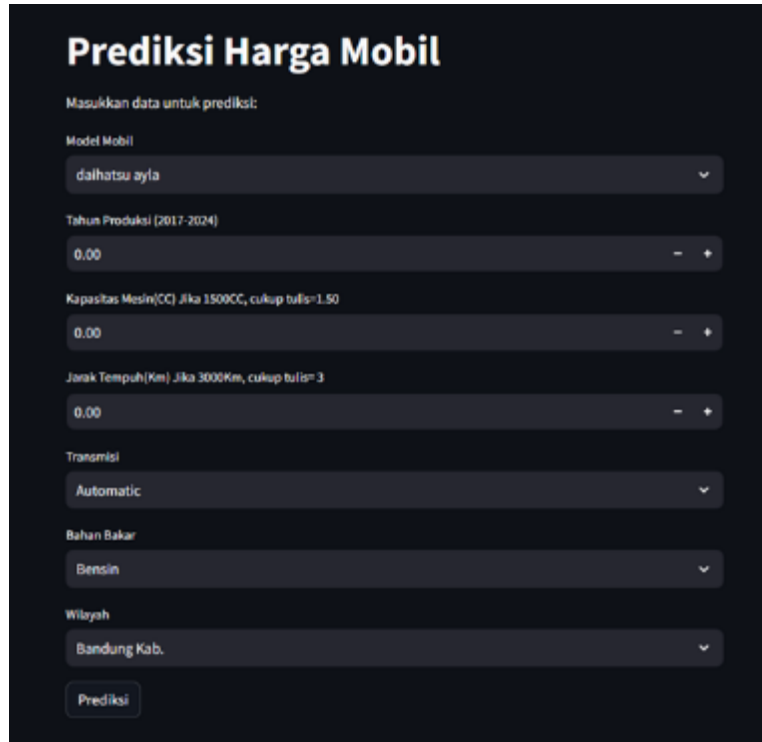
MAE	RMSE	R-SQUARE
3.053.701	4.751.339	0.90010494

Evaluasi model *Random Forest Regressor* menunjukkan performa yang baik dalam memprediksi harga mobil bekas. Hal ini dibuktikan dengan nilai *Mean Absolute Error (MAE)* sebesar 3.053.701 yang menunjukkan bahwa rata-rata selisih prediksi harga yang dihasilkan model dengan harga aktual relatif kecil, yaitu sekitar 3 juta rupiah. Meskipun nilai *Root Mean Squared Error (RMSE)* sebesar 4.751.339 tergolong cukup besar, namun hal ini masih dalam batas wajar mengingat RMSE menggunakan satuan mata uang yang sama dengan data harga dan lebih sensitif terhadap fluktuasi data. Lebih lanjut, nilai *R-squared (R²)* yang tinggi, yaitu sebesar 0,90 mengindikasikan bahwa model mampu menjelaskan sekitar 90% variabilitas harga aktual berdasarkan variabel input yang digunakan. Hasil evaluasi ini menunjukkan bahwa model *Random Forest Regressor* memiliki akurasi prediksi yang baik dan mampu memperhitungkan faktor-faktor yang relevan dalam memprediksi harga mobil bekas.

Pengembangan WebApp untuk Prediksi Harga Mobil Bekas

Hasil pengembangan *webApp* untuk prediksi harga mobil bekas ditunjukkan pada Gambar 8. Gambar tersebut menunjukkan tampilan dari halaman utama *webapp*. Aplikasi ini dirancang untuk membantu pengguna dalam memperkirakan harga jual mobil bekas berdasarkan beberapa parameter penting yang dapat mempengaruhi nilai jual mobil di pasaran. Fitur *webApp* ini terdiri dari Input Parameter Mobil dan Tombol Prediksi.

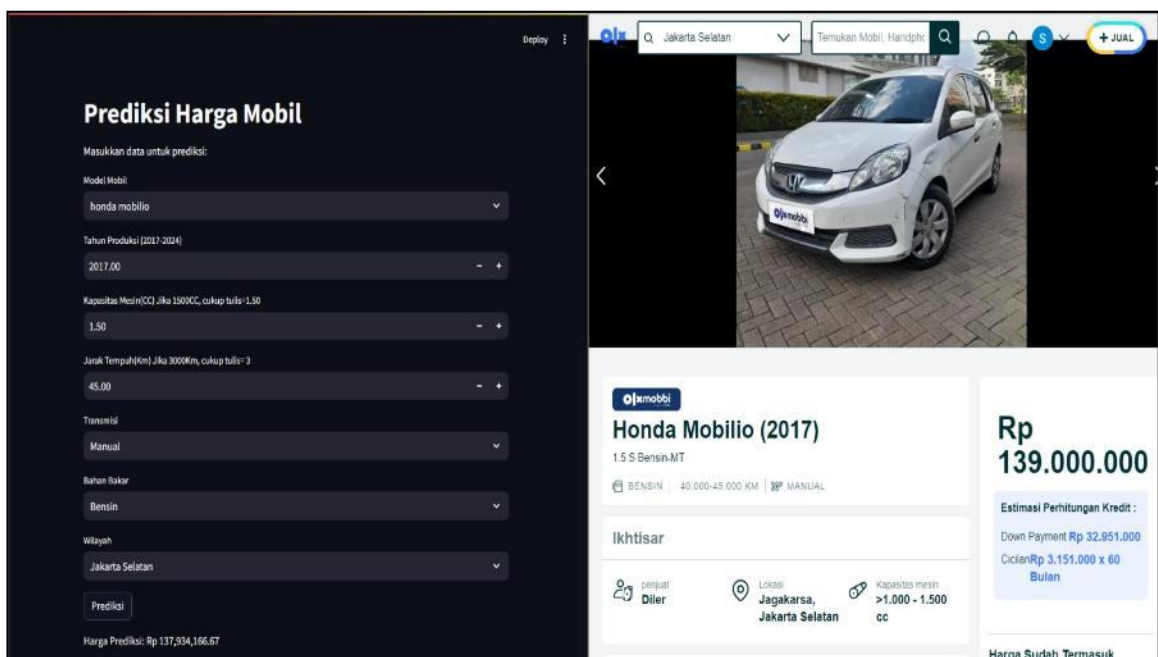
Input Parameter Mobil terdiri dari parameter model mobil, tahun produksi mobil, kapasitas mesin, ajrak tempuh, transmisi pilihan, bahan bakar, dan wilayah. Setelah semua parameter dimasukkan, pengguna dapat menekan tombol prediksi untuk mendapatkan estimasi harga jual mobil berdasarkan model yang telah dilatih menggunakan *Random Forest*.



Gambar 8. *WebApp* Prediksi Harga Mobil

Implementasi Hasil Prediksi pada *WebApp*

Gambar 9 menunjukkan implementasi model *Random Forest* dalam sebuah *WebApp* yang digunakan untuk memprediksi harga mobil bekas. Pada aplikasi tersebut, pengguna memasukkan parameter, yaitu model mobil, tahun produksi, kapasitas mesin, jarak tempuh, transmisi, bahan bakar, dan wilayah, untuk mendapatkan estimasi harga dari mobil yang dimaksud.



Gambar 9. Implementasi *WebApp* pada prediksi harga mobil bekas

Aplikasi ini berhasil memberikan estimasi harga yang mendekati harga pasar aktual. Sebagai contoh, untuk Honda Mobilio tahun 2017 dengan spesifikasi tertentu,

aplikasi memprediksi harga sebesar Rp 137,934,166.67, sementara harga pasar aktual dari mobil yang sama adalah Rp 139,000,000. Selisih yang sangat kecil antara harga prediksi dan harga pasar ini menunjukkan bahwa model *Random Forest* yang digunakan memiliki tingkat akurasi yang tinggi.

Dengan demikian, aplikasi ini dapat menjadi alat yang sangat berguna bagi pengguna yang ingin mengetahui estimasi harga mobil bekas mereka sebelum melakukan transaksi jual-beli. Selain itu, hasil yang akurat ini juga menegaskan bahwa model *Random Forest* adalah pilihan yang tepat untuk tugas prediksi harga, terutama dalam konteks data yang kompleks seperti harga mobil bekas.

PENUTUP

Kesimpulan

Penelitian ini berhasil mengembangkan model prediksi harga mobil bekas dengan algoritma *Random Forest* yang memiliki tingkat akurasi tinggi. Model yang dihasilkan mampu menjelaskan 90% variabilitas harga berdasarkan data yang relevan, sehingga model ini terbukti efektif untuk diterapkan di pasar otomotif Indonesia, khususnya pada kisaran harga mobil bekas antara 100-150 juta rupiah. Pengembangan *webApp* sudah berhasil dilakukan. *WebApp* ini berfungsi sebagai *interface* bagi *user* yang ingin melakukan prediksi harga mobil bekas. Dengan adanya *webApp* prediksi harga mobil yang dikembangkan dengan memanfaatkan *machine learning* dan algoritma *Random Forest*, dapat membantu penjual dan pembeli dalam mengambil keputusan yang lebih tepat dengan estimasi harga yang mendekati pasar.

Saran

Sebagai pengembangan lebih lanjut, disarankan untuk meningkatkan akurasi prediksi dengan menambahkan variabel eksternal seperti kondisi ekonomi atau fluktuasi permintaan pasar, mengintegrasikan model prediksi ke dalam platform e-commerce agar pengguna di Indonesia dapat menentukan harga jual mobil bekas secara akurat dan kompetitif dan memperluas cakupan dataset dengan menyertakan variasi model mobil dan wilayah geografis yang lebih luas agar prediksi harga lebih representatif.

REFERENSI

- Aleryani, A. Y. (2024). Analyzing Data Flow: A Comparison between Data Flow Diagrams (DFD) and User Case Diagrams (UCD) in Information Systems Development. *European Modern Studies Journal*, 8(1), 313–320. [https://doi.org/10.59573/emsj.8\(1\).2024.28](https://doi.org/10.59573/emsj.8(1).2024.28)
- Al-Turjman, F., Hussain, A. A., Alturjman, S., & Altrjman, C. (2022). Vehicle Price Classification and Prediction Using Machine Learning in the IoT Smart Manufacturing Era. *Sustainability (Switzerland)*, 14(15). <https://doi.org/10.3390/su14159147>
- Arribe, E., Safitri, E., & Tsabitah, N. (2023). *PERANCANGAN SISTEM PENDAFTARAN PASIEN RAWAT JALAN RUMAH SAKIT PMC BERBASIS WEB*. 10(2).
- Gede Iwan Sudipa, I., Andi Putra, T. E., Jurnaidi Wahidin, A., Alfa Syukrilla, W., Khrisna Wardhani, A., Heryana, N., Indriyani, T., Willyanto Santoso Tutuk Indriyani, L., & Willyanto Santoso, L. (2023). *DATA MINING*. www.globaleksekitifteknologi.co.id
- Ihzaniah, L. S., Setiawan, A., & Wijaya, R. W. N. (2023). Perbandingan Kinerja Metode Regresi K-Nearest Neighbor dan Metode Regresi Linear Berganda pada Data Boston Housing. *Jambura Journal of Probability and Statistics*, 4(1), 17–29. <https://doi.org/10.34312/jjps.v4i1.18948>

- Khairunnisa, A. (2023). PERBANDINGAN MODEL RANDOM FOREST DAN XGBOOST UNTUK PREDIKSI KEJAHATAN KESUSILAAN DI PROVINSI JAWA BARAT. *JIKO (Jurnal Informatika Dan Komputer)*, 7(2), 202. <https://doi.org/10.26798/jiko.v7i2.799>
- Lahon, M., & Sharma, U. (2019). Complexity assessment based on UML-activity diagram. *International Journal of Recent Technology and Engineering*, 8(2), 2391–2397. <https://doi.org/10.35940/ijrte.B1596.078219>
- Liu, G. R. (Gui-R. (2024). Machine Learning Theory and Applications. In *Machine Learning Theory and Applications*. World Scientific Publishing Co. Pte. Ltd. <https://doi.org/10.1002/9781394220649>
- MA'ARIF ALFIAN. (2020). *BUKU AJAR PEMROGRAMAN LANJUT BAHASA PEMROGRAMAN PYTHON*.
- Marisa, F., Maukar, A. L., & Akhriza, T. M. (2021). Data Mining Konsep dan Pembahasannya. *Mining of Massive Datasets*, 1, 1–127. <https://repository.widyagama.ac.id/files/original/9e2d3a13ce7e4845e9597b2039f43ff3.pdf>
- MUTI, S., & YILDIZ, K. (2023). Using Linear Regression For Used Car Price Prediction. *International Journal of Computational and Experimental Science and Engineering*, 9(1), 11–16. <https://doi.org/10.22399/ijcesen.1070505>
- Raharjo, B. (2021). *Pembelajaran Mesin (Machine Learning)*.
- Raj, A. (2019). A Review on Machine Learning Algorithms. *International Journal for Research in Applied Science and Engineering Technology*, 7(6), 792–796. <https://doi.org/10.22214/ijraset.2019.6138>
- Tholib Abu. (2023). *Implementasi Algoritma Machine Learning Berbasis Web dengan Framework Streamlit (1st ed.)*. Pustaka Nurja. <https://pustakanurja.unuja.ac.id>
- Zahra Lubis, A., Lastian Nahulae, L., Marliana Anggraini, N., Adawiyah, R., & Islam Negeri Sumatera Utara, U. (2024). *ANALISIS FAKTOR-FAKTOR YANG MEMENGARUHI PENETAPAN HARGA. Volume 9*. <https://doi.org/10.30651/jms.v9i1.21412>