

Sistem Informasi Prediksi Harga Mobil Bekas Menggunakan Metode KNN

Haya Fadhillah¹, Salamun Rohman Nudin²

D4 Manajemen Informatika, Fakultas Vokasi, Universitas Negeri Surabaya
Kampus Unesa 1, Jalan Ketintang, Surabaya

¹haya.20005@mhs.unesa.ac.id

²salamunrohman@unesa.ac.id

Abstrak— Penetapan harga mobil bekas yang akurat merupakan tantangan bagi penjual dan pembeli mobil bekas. Metode K-Nearest Neighbor (KNN) pada machine learning dipilih karena algoritma ini melakukan pengklasifikasian data berdasarkan jarak terdekat dengan tujuan membantu pengguna dalam menentukan harga terbaik pada mobil bekas. Data yang digunakan berasal dari situs kaggle. Instrumen penelitian berupa dataset yang terdiri dari fitur-fitur mobil seperti merek, model, tahun, jarak tempuh, dan harga. Teknik analisis data menggunakan algoritma KNN untuk melakukan prediksi harga berdasarkan fitur-fitur tersebut. Sistem dikembangkan menggunakan bahasa pemrograman Python dan framework Laravel untuk membuat website. Evaluasi menggunakan metrik Root Mean Squared Error (RMSE) untuk menunjukkan kinerja sistem prediksi yang akurat. Dengan RMSE yang hanya sebesar 1324544.102, menunjukkan bahwa rata-rata kesalahan prediksi harga mobil bekas dapat meleset sekitar angka tersebut dari harga aktual.

Kata kunci— Dataset, Machine Learning, Harga mobil bekas, KNN, RMSE

Abstract— Accurate used car pricing is a challenge for used car sellers and buyers. The K-Nearest Neighbor (KNN) method in machine learning was chosen because this algorithm classifies data based on the closest distance with the aim of helping users determine the best price for used cars. The data used comes from the kaggle site. The research instrument is a dataset consisting of car features such as make, model, year, mileage, and price. The data analysis technique uses the KNN algorithm to predict prices based on these features. The system was developed using Python programming language and Laravel framework to create a website. The evaluation uses Root Mean Squared Error (RMSE) metric to show the accurate performance of the prediction system. With an RMSE of only 1324544.102, it shows that the average used car price prediction error can be off by about that number from the actual price.

Keywords— Dataset, Machine Learning, Used car price, KNN, RMSE

I. PENDAHULUAN

Dalam beberapa tahun terakhir, industri otomotif telah berkembang pesat, terutama di pasar mobil bekas. Harga mobil bekas yang lebih murah dibandingkan dengan mobil baru merupakan nilai jual utama bagi banyak pembeli. [2]. Namun, menentukan harga yang wajar untuk mobil bekas seringkali menjadi tantangan bagi konsumen maupun penjual. Harga mobil bekas dapat dipengaruhi oleh berbagai faktor, seperti tahun produksi, jarak tempuh, kondisi mobil, dan tren pasar [8].

Salah satu pendekatan yang mampu dipakai untuk memperkirakan harga mobil bekas adalah dengan

memanfaatkan teknik machine learning. Salah satu teknik pembelajaran mesin yang terkenal dan ampuh untuk mengklasifikasikan dan memprediksi dari data masa lalu adalah pendekatan K-Nearest Neighbor (KNN) [8]. Dengan menggunakan metode KNN, sistem dapat menganalisis data mobil bekas yang telah terjual sebelumnya dan mencari kemiripan dengan mobil yang ingin diprediksi .

II. KAJIAN PUSTAKA

A. Machine Learning

Tujuan *machine learning* yang merupakan subbidang AI, adalah untuk menciptakan algoritma dan sistem yang dapat "belajar" dari data untuk membuat penilaian atau prediksi tanpa campur tangan manusia. Pendekatan *machine learning* digunakan untuk memprediksi harga mobil bekas dalam penelitian ini. Dapat pula diartikan sebagai kemampuan sistem komputer untuk belajar dan membangun model prediktif dari data, bukan dengan instruksi yang diprogram secara eksplisit [10].

B. KNN (K-Nearest Neighbor)

Algoritma K-Nearest Neighbor (KNN) adalah metode yang sangat efektif dalam melaksanakan prediksi dan klasifikasi menurut data historis, dalam konteks prediksi harga mobil bekas KNN akan mencari k data mobil bekas yang mirip dengan data yang ingin diprediksi harganya, kemudian mengambil rata-rata harga dari k tetangga terdekat tersebut sebagai prediksi harga [10]. Pemilihan fitur yang relevan dan berpengaruh terhadap harga mobil bekas sangat penting untuk meningkatkan akurasi prediksi.

C. RMSE (Root Mean Squared Error)

Root Mean Squared Error (RMSE) adalah metode yang digunakan untuk mengukur sejauh mana perbedaan antara nilai yang diprediksi oleh suatu model dengan nilai yang diamati. Ketepatan estimasi suatu nilai ditunjukkan oleh rendahnya nilai RMSE. Semakin kecil nilai RMSE, semakin tinggi tingkat akurasi estimasi dibandingkan dengan RMSE yang lebih besar [6].

D. Laravel

Laravel merupakan salah satu framework PHP terpopuler dalam pengembangan aplikasi web. Framework ini dikembangkan untuk mempercepat dan menyederhanakan proses pembuatan web dengan menyediakan berbagai fitur serta alat bawaan [3]. Mengadopsi arsitektur MVC (*Model-*

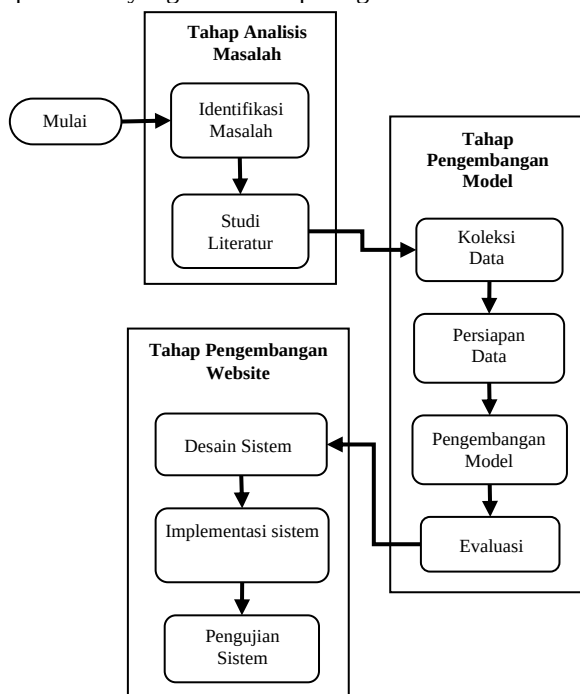
View-Controller), Laravel memisahkan logika aplikasi, tampilan, dan data, sehingga mempermudah pengelolaan serta pengembangannya.

E. Website

Website merupakan sekumpulan halaman informasi yang dapat diakses melalui internet atau jaringan komputer. Halaman-halaman ini saling terhubung dan umumnya dibuka menggunakan peramban web (browser) seperti Google Chrome, Mozilla Firefox, atau Microsoft Edge. Website berfungsi untuk menampilkan berbagai jenis konten, termasuk teks, gambar, animasi, dan suara, baik dalam bentuk statis maupun dinamis, yang tersusun dalam satu struktur yang saling berkaitan [1].

III. METODOLOGI PENELITIAN

Pada Metodologi penelitian ini menggambarkan mengenai alur penelitian yang tercantum pada gambar 1.



Gambar 1 Rancangan Penelitian

1. Tahap Analisis Masalah

Tahap masalah masalah terbagi menjadi 2 langkah yaitu identifikasi masalah dan studi literatur.

a. Identifikasi Masalah

Identifikasi masalah pada sistem informasi prediksi harga mobil bekas dengan metode KNN sering meliputi akurasi prediksi yang rendah dan pemilihan fitur yang tidak tepat, selain itu, masalah dalam pengolahan data dan pemilihan parameter K juga dapat mempengaruhi hasil prediksi. Dalam konteks sistem prediksi harga mobil bekas menggunakan metode KNN (K-Nearest Neighbors), identifikasi masalah untuk memahami kebutuhan dan tantangan dalam menentukan harga mobil bekas secara akurat dan

objektif agar dapat memberikan manfaat bagi penjual, pembeli, dan industri mobil bekas secara keseluruhan.

b. Studi Literatur

Pada tahap ini peneliti mengambil sumber-sumber informasi berdasarkan 2 jurnal terindex sinta 2 dan 2 jurnal internasional yang dari keempat jurnal tersebut dapat disimpulkan bahwa kombinasi metode KNN dan RMSE dapat menjadi pendekatan yang efektif dalam meningkatkan akurasi prediksi harga mobil bekas.

2. Tahap Pengembangan Model

Tahapan ini terbagi menjadi 4 langkah yaitu koleksi data, persiapan data, pengembangan model dan evaluasi.

a. Koleksi Data

Sumber data merupakan dataset yang diambil dari sumber platform online Kaggle dengan kata kunci "Car Price From monil123.com (May 2020)". Jumlah keseluruhan data terdiri dari 103227 baris dan 11 kolom.

b. Persiapan Data

Persiapan data (data preparation) melibatkan proses pembersihan data, seleksi fitur, normalisasi data dan penanganan data yang hilang, dalam hal ini peneliti menghasilkan dataset akhir sejumlah 55086 baris dan 7 kolom

c. Pengembangan Model

Pada tahap ini data yang telah disiapkan selanjutnya digunakan untuk melatih dan membangun model prediksi menggunakan metode K-Nearest Neighbors (KNN). melalui tahapan seperti, penentuan parameter, pelatihan model dan validasi model. Jarak antara data tetangga dihitung menggunakan rumus 1 Euclidean berikut:

$$euc = \sqrt{\sum_{k=1}^n (a_k - b_k)^2} \dots \dots \dots (2.1)$$

Keterangan

a = data latih

k = variable data

b = data uji

d = dimensi data

d. Evaluasi

Tahap evaluasi dalam penelitian ini menggunakan RMSE sebagai pengukur keakuratan sistem yang dibangun.

$$RMSE = \sqrt{\frac{1}{n} (\sum_{j=1}^n (y_j - \hat{y}_j)^2)} \dots \dots \dots (2.2)$$

Keterangan

y = data uji yang sebenarnya

n = jumlah data uji

$\hat{y}$ = data hasil prediksi

3. Tahap Pengembangan Website

Tahapan pengembangan website berisi tahap desain system, implementasi system dan pengujian system.

a. Desain Sistem

Desain sistem dikembangkan dengan framework Laravel 10 yang dikombinasikan dengan API flask. Laravel akan berfungsi sebagai aplikasi utama yang menangani antarmuka pengguna dan logika bisnis, sementara Flask akan digunakan untuk layanan prediksi menggunakan model machine learning.

c. Implementasi Sistem

Implementasi sistem prediksi harga mobil bekas, kami mengintegrasikan Laravel sebagai framework backend dan Flask sebagai API untuk layanan prediksi. Proses dimulai dengan menyiapkan lingkungan pengembangan, di mana Laravel diinstal dan dikonfigurasi, sementara Flask dibangun dengan endpoint `/api/predict` yang menerima data dan mengembalikan prediksi harga. Pada Laravel, kami membuat model dan controller yang menggunakan HTTP client untuk mengirim data ke API Flask dan menerima respons prediksi, dengan demikian sistem ini berhasil menggabungkan kekuatan Laravel dan Flask, memungkinkan prediksi harga mobil bekas yang efisien dan terintegrasi.

d. Pengujian Sistem

Pengujian menggunakan *black box* pada sistem prediksi harga mobil bekas ini bertujuan untuk memastikan bahwa sistem berfungsi sesuai harapan dari perspektif pengguna akhir. Berikut langkah yang diambil saat pengujian,

- Identifikasi fitur yang diuji
- Periapan kasus uji
- Pelaksanaan pengujian
- Dokumentasi hasil

IV. HASIL DAN PEMBAHASAN

1. Hasil Penelitian

a. Implementasi Algoritma KNN

Implementasi algoritma K-Nearest Neighbors (KNN) dimulai dari data yang digunakan terdiri 7 atribut yang telah disesuaikan terdiri dari model, tahun, harga, transmisi, jarak tempuh, pajak dan kapasitas mesin yang kemudian dibagi menjadi set pelatihan dan set pengujian. Proses dimulai dengan normalisasi data agar semua fitur berada dalam skala yang seragam, diikuti dengan pemilihan nilai K yang optimal melalui teknik validasi silang. Setelah model KNN dilatih menggunakan set pelatihan, kinerja model dievaluasi dengan menghitung akurasi menggunakan RMSE. Hasil penelitian menunjukkan bahwa algoritma KNN mampu menghasilkan perkiraan harga yang tepat dengan tingkat kesalahan yang masih dalam batas wajar, sehingga membuktikan efektivitasnya dalam aplikasi prediksi harga mobil bekas..

b. Implementasi Sistem Prediksi

Pada tahap ini peneliti mengembangkan sistem informasi berdasarkan rancangan yang telah dibangun sebelumnya menggunakan framework laravel serta pengkodean system menggunakan *text editor Visual Code*. Berikut tampilan website yang telah dikembangkan:

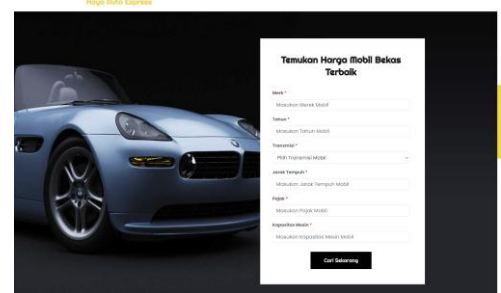
- Halaman Utama (Home)



Gambar 2 Landing Page

Pada Gambar 6 merupakan tampilan awal saat membuka website prediksi harga mobil bekas, yang selanjutnya usap kebawah untuk membuka halaman form pengisian prediksi.

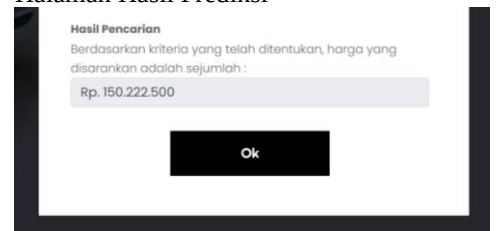
- Halaman Form Prediksi



Gambar 3 Halaman Pengisian Form Prediksi

Pengguna akan memasukkan kategori-kategori tertera pada kolom gambar 7 yang selanjutnya digunakan untuk memprediksi harga sesuai dengan input yang diperintahkan.

- Halaman Hasil Prediksi



Gambar 4 Output Prediksi Harga

Gambar 8 tersebut merupakan hasil prediksi dari data yang sudah diinput oleh pengguna pada form prediksi sebelumnya. Hasil prediksi ditampilkan tepat dibawah form input untuk prediksi.

c. Pengujian Sistem

Tahap Pengujian system prediksi ini peneliti menguji system berdasarkan metode *black box testing* dan *usability testing*. Berikut pengujian yang dilakukan:

1) Black Box Testing

Pengujian ini dilakukan dengan melihat juga mengamati hasil dari eksekusi sistem dengan memeriksa fungsionalitas dari perangkat lunak yang telah dibangun.

TABEL I
UKURAN HURUF UNTUK MAKALAH

Pengujian	Masukan	Hasil yang Diharapkan	Hasil yang Diperoleh
Pengujian Homepage	Menampilkan informasi singkat website	Tampilan Homepege yang sesuai	Berhasil
Pengujian Form Data Mobil Bekas	Menjalankan fungsi isi pada tiap form sesuai dengan yang ditentukan	Sistem berhasil menampilkan data yang telah diinput	Berhasil
Pengujian Hasil Input Prediksi Harga Mobil Bekas	Menjalankan perhitungan dari input form yang telah dilakukan pengguna	Sistem berhasil menampilkan hasil prediksi berupa harga dalam nilai rupiah dari mobil bekas	Berhasil

2. Pembahasan

Bagian pembahasan ini akan menjabarkan mengenai alur sistem prediksi harga mobil bekas berbasis website yang sudah disesuaikan dengan alur penelitian. Langkah selanjutnya merupakan dilakukannya pengembangan model menggunakan platform Google Colab. Langkah awal pada proses pengembangan model adalah pengumpulan data yang mana pada penelitian ini data bersumber dari website Kaggle yang berisikan harga-harga mobil bekas beserta fitur-fitur pada mobil sebagai penunjang untuk memprediksi harga. Kemudian dilakukannya langkah penembangan model yang, lalu membagi jumlah data menjadi 2 yaitu untuk *Test* dan *Train*. Implementasi metode KNN dengan melakukan training data yang kemudian melakukan pengecekan akurasi dengan RMSE. Proses dimulai ketika memasukkan data mobil yang ingin diprediksi harganya pada dalam form pengisian. Setelah semua form terisi lalu klik 'Cari Sekarang' dan tekan 'Ya' pada pop up tampilkan hasil predisksi, hal ini dilakukan untuk mengirimkan data ke API Flask untuk dilakukan pemrosesan data. Selanjutnya API Flask diterapkan untuk *me-load* permodean menggunakan modul *pickle* pada Python yang berfungsi sebagai pengelola prediksi dengan lebih efisien.

```
MSE: 1754417078956.0833
RMSE: 1324544.1023069345
MAPE: 0.00018263553720448378
```

Gambar 5 Output Akurasi Data

Berdasarkan gambar 9 dapat dilihat bahwa nilai MSE yang besar ini bisa disebabkan oleh outlier atau

variabel yang tidak terduga dalam dataset, namun disamping itu secara rinci akurasi RMSE sebesar 1324544.1023 yang berarti bahwa prediksi harga mobil bekas dapat meleset sekitar 1.324.544,- dari harga aktual, selanjutnya yang terakhir MAPE sebesar 0.00018263553720448378 (atau 0.018%) yang menunjukkan bahwa model memiliki kesalahan prediksi yang relatif kecil dalam persentase sehingga dapat disimpulkan bahwa model ini memprediksi harga dengan cukup baik.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Mengembangkan sistem informasi prediksi harga mobil bekas menggunakan metode KNN ini peneliti menggunakan dataset yang diunduh dari Kaggle dan lakukan eksplorasi data untuk memahami fitur yang ada sehingga sistem informasi prediksi harga mobil bekas ini dapat diimplementasikan dengan baik.

1. Mengembangkan model prediksi harga mobil bekas menggunakan metode KNN. Penggunaan metode KNN dalam penelitian ini dapat diimplementasikan dengan baik dan terbukti akurat dalam melakukan prediksi harga mobbil bekas dengan hasil akurasi RMSE 1324544.1023069345 yang dapat diartikan bahwa prediksi harga dapat meleset hanya sejauh 1.324.544,- dan jika dipresentasikan dapat dilihat melalui besarnya nilai MAPE sejumlah 0.00018263553720448378 atau berkisar 0.018%. Pemanfaatan 7 fitur dalam dataset yang digunakan untuk penelitian ini membuktikan bahwa metode KNN (K-Nearest Neighbor) mampu mengklasifikasi harga mobil bekas sesuai dengan parameter K yang digunakan dengan optimal
2. Merancang dan mengembangkan sistem informasi prediksi harga mobil bekas menggunakan framework laravel berbasis website. istem Informasi berbasis website dikembangkan menggunakan framework laravel berhasil diimplementasikan sesuai dengan rancangan yang dibangun. Sistem informasi ini menghasilkan output berupa hasil prediksi bernilai rupiah yang ditampilkan sesudah form pengisian fitur mobil oleh pengguna sehingga mudah untuk dipahami.

B. Saran

Penelitian ini juga membuka peluang untuk eksplorasi lebih lanjut mengenai:

1. Peningkatan data, dalam hal ini disarankan untuk mengumpulkan lebih banyak data serta fitur tambahan yang relevan guna meningkatkan akurasi.
2. Pengujian metode lain, selain menggunakan KNN, sebaiknya dilakukan evaluasi dengan algoritma lain seperti regresi linier atau random forest untuk membandingkan hasil dan menentukan model yang paling efektif.
3. Monitoring dan pembaruan model, model perlu dipantau dan diperbarui secara berkala untuk

memastikan bahwa akurasi prediksi tetap tinggi seiring dengan perubahan data dan kondisi pasar.

REFERENSI

- [1] Cherniaieva, A. A. 2021. “Частота Асимптоматической Гиперурикемии Среди Взрослых Больных Сахарным Диабетом 1-Го И 2-Го Типа.” *INTERNATIONAL JOURNAL OF ENDOCRINOLOGY (Ukraine)* 16(4):327–32. doi: 10.22141/2224-0721.16.4.2020.208486.
- [2] Dahiya, Himanshu, Chetan Aggarwal, Shubh Goyal, dan Mini Agarwal. 2021. “Used Car Price Prediction Using Machine Learning.” *International journal of multidisciplinary advanced scientific research and innovation* 1(10):246–51. doi: 10.53633/ijmasri.2021.1.10.002.
- [3] Elmas, Ç. (2021). Predicting Used Car Prices Using Machine Learning Techniques. *PESA International Journal of Social Studies*, 7(3), 61-74
- [4] Hadinata, Wira, dan Lilis Stianingsih. 2024. “Analisis Perbandingan Performa Restfull Api Antara Express.Js Dengan Laravel Framework.” *Jurnal Informatika dan Teknik Elektro Terapan* 12(1):531–40. doi: 10.23960/jitet.v12i1.3845.
- [5] Jiang, J., & Lian, Y. (2021). A Survey on Used Car Price Prediction with Machine Learning Techniques. In *Proceedings of the 2021 International Conference on Computer Information and Big Data Applications (CIBDA)* (pp. 211-216). IEEE
- [6] Khandelwal, A., & Sawant, S. (2021). Used Car Price Prediction Using Machine Learning. *International Journal of Engineering Research & Technology (IJERT)*, 10(5), 450-454
- [7] Krithika, S., & Nagarajan, N. (2020). Prediction of Used Car Prices Using Machine Learning Techniques. *International Journal of Engineering Research & Technology (IJERT)*, 9(5), 818-822.
- [8] Lau, L. (2021). The Rise of Visual Studio Code. *IEEE Software*, 38(2), 104-107. <https://doi.org/10.1109/MS.2021.3051899>
- [9] Nurfauzan, D., dan T. Fatimah. 2022. “Implementasi Algoritma K-Nearest Neighbors Regression Dalam Memprediksi Harga Saham.” *Seminar Nasional Mahasiswa ...* (September):576–84.
- [10] Pandey, S., & Singh, K. (2020). A Systematic Literature Review of Machine Learning Techniques for Used Car Price Prediction. In *Proceedings of the International Conference on Computing and Communication Systems* (pp. 199-208). Springer, Singapore
- [11] Rahardja, Chandra Arief, Try Juardi, dan Halim Agung. 2019. “Implementasi Algoritma K-Nearest Neighbor Pada Website Rekomendasi Laptop.” *Jurnal Buana Informatika* 10(1):75–84. doi: 10.24002/jbi.v10i1.1847.
- [12] Samruddhi, K., dan R. Ashok Kumar. 2020a. “Used Car Price Prediction using K-Nearest Neighbor Based Model.” *International Journal of Innovative Research in Applied Sciences and Engineering* 4(2):629–32. doi: 10.29027/ijirase.v4.i2.2020.629-632.
- [13] Saputra, Wilson, Joko Santoso, dan Patricia Ardanari. 2021. “Penerapan Metode K-Nearest Neighbor Untuk Mendeteksi Penyakit Kulit.” *Jurnal Informatika Atma Jogja* 2(1):63–72.
- [14] Suyanto. 2022. *MACHINE LEARNING TINGKAT DASAR DAN TINGKAT LANJUT*. 2 ed. Bandung: INFORMATIKA.
- [15] Xu, Z., Zu, J., & Li, Y. (2020). A Web-Based Used Car Price Prediction System Using Machine Learning Techniques. *International Journal of Machine Learning and Computing*, 10(3), 376-382. <https://doi.org/10.18178/ijmlc.2020.10.3.948>
- [16] YILMAZ, Seda, dan İhsan Hakan SELVİ. 2023. “Price Prediction Using Web Scraping and Machine Learning Algorithms in the Used Car Market.” *Sakarya University Journal of Computer and Information Sciences* 6(2):140–48. doi: 10.35377/saucis...1309103.
- [17] Zhou, Y., Wang, Y., Hou, Y., & Wang, W. (2020). Used car price prediction based on machine learning techniques. *IEEE Access*, 8, 128369-128380.