



Sistem Informasi Prediksi Harga Pembelian Mobil Bekas Menggunakan Algoritma Random Forest

Muhammad Farhan Azmi Darus*, Triase

Fakultas Sains dan Teknologi, Program Studi Sistem Informasi, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email: ^{1,*}farhanazmi393@gmail.com, ²triase@uinsu.ac.id

Email Penulis Korespondensi: farhanazmi393@gmail.com

Abstrak—Perkembangan pasar mobil bekas di Indonesia yang semakin kompetitif menuntut adanya sistem yang mampu memberikan estimasi harga secara objektif dan berbasis data. Permasalahan utama yang dihadapi adalah penentuan harga pembelian mobil bekas yang masih bersifat subjektif dan bergantung pada pengalaman individu, sehingga sering terjadi ketidaksesuaian harga antara penjual dan pembeli. Penelitian ini bertujuan untuk mengembangkan sistem informasi prediksi harga pembelian mobil bekas menggunakan algoritma Random Forest. Dataset yang digunakan terdiri dari 301 data kendaraan dengan atribut meliputi merek, model, tahun produksi, kapasitas mesin, jarak tempuh, jenis bahan bakar, transmisi, dan wilayah. Tahapan penelitian meliputi pengumpulan data, preprocessing data, pembagian dataset menjadi data latih dan uji, pembangunan model, serta evaluasi model. Evaluasi dilakukan menggunakan metrik Mean Absolute Error (MAE) dan R-Squared (R^2). Hasil penelitian menunjukkan bahwa model Random Forest menghasilkan nilai MAE sebesar 22.654 dan R^2 sebesar 0.958, yang menunjukkan bahwa model mampu menjelaskan variasi data sebesar 95,8% dengan tingkat kesalahan prediksi yang relatif rendah. Sistem yang dikembangkan mampu memberikan estimasi harga secara cepat dan akurat, sehingga dapat membantu pengguna dalam pengambilan keputusan serta meningkatkan transparansi dalam transaksi mobil bekas.

Kata Kunci: Prediksi Harga; Mobil Bekas; Random Forest; Sistem Informasi; Algoritma

Abstract—The rapid growth of the used car market in Indonesia has increased the need for an objective and data-driven approach to determine appropriate purchase prices. One of the main challenges in used car transactions is the subjectivity in price determination, which often leads to discrepancies between sellers and buyers. This study aims to develop an information system for predicting used car purchase prices using the Random Forest algorithm. The dataset used in this study consists of 301 records with attributes including brand, model, production year, engine capacity, mileage, fuel type, transmission, and region. The research methodology involves data collection, preprocessing, dataset splitting, model development, and evaluation. The model performance is evaluated using Mean Absolute Error (MAE) and R-Squared (R^2). The results show that the Random Forest model achieves an MAE value of 22.654 and an R^2 value of 0.958, indicating that the model can explain 95.8% of the variance in the data with relatively low prediction error. The developed system implements an input-process-output mechanism, where user-provided vehicle data are processed by the trained model to generate price predictions automatically. The system is able to provide fast, accurate, and objective price estimates, thereby supporting better decision-making and improving transparency and efficiency in used car transactions.

Keywords: Price Prediction; Used Cars; Random Forest; Information System; Algorithm

1. PENDAHULUAN

Perkembangan industri otomotif di Indonesia semakin pesat, seiring dengan meningkatnya kebutuhan masyarakat akan kendaraan pribadi. Hal ini mendorong pertumbuhan pasar mobil bekas yang semakin kompetitif. Pembeli mobil bekas menghadapi tantangan dalam menentukan harga yang sesuai dengan kondisi pasar. Salah satu kendala utama dalam transaksi mobil bekas adalah adanya perbedaan harga yang cukup signifikan antara penjual dan pembeli. Penjual sering kali menetapkan harga yang lebih tinggi dengan harapan mendapatkan keuntungan maksimal, sementara pembeli cenderung mencari harga yang lebih rendah sesuai dengan anggaran mereka (E. Dewi et al., 2020).

Mobil bekas merupakan bagian besar dari pasar otomotif, di mana harga jualnya dipengaruhi oleh berbagai faktor seperti usia mobil, kondisi mesin, kilometer yang telah ditempuh, serta faktor pasar dan tren. Penentuan layak tidaknya harga beli atau jual mobil bekas, merupakan salah satu kendala yang dihadapi oleh masyarakat dalam pengambilan keputusan saat akan membeli atau menjual mobil atau kendaraan. Informasi yang ditemukan di internet pun terkadang sangat beragam tanpa ada kejelasan faktor yang mempengaruhi penentuan harga tersebut (Hasibuan et al., 2022).

Agen SQL adalah Agen SQL mobil bekas yang berfokus pada penyediaan kendaraan berkualitas dengan harga yang kompetitif. Agen SQL ini melayani berbagai segmen pelanggan, mulai dari individu yang mencari mobil pertama hingga kolektor yang menginginkan kendaraan spesifik. Agen SQL juga menawarkan layanan penjualan mobil bekas dengan harga kompetitif serta konsultasi harga berdasarkan data pasar terkini. Agen SQL ini juga menyediakan layanan *test drive* bagi pelanggan yang ingin mencoba kendaraan sebelum membeli.

Saat ini, banyak Agen SQL mobil bekas termasuk Agen SQL masih menggunakan metode manual atau berdasarkan pengalaman subjektif dalam menentukan beli kendaraan. Hal ini dapat menyebabkan harga yang tidak sesuai dengan kondisi pasar, baik terlalu tinggi yang dapat menghambat penjualan maupun terlalu rendah yang dapat merugikan penjual. Selain itu, tanpa sistem prediksi harga yang akurat, pembeli juga kesulitan mengetahui apakah harga yang ditawarkan sesuai dengan nilai pasar kendaraan tersebut. Oleh karena itu, diperlukan solusi yang dapat memberikan estimasi harga secara objektif dan berbasis data guna meningkatkan transparansi dalam transaksi beli mobil bekas.



Sebagai solusi atas permasalahan tersebut, penelitian ini mengusulkan pengembangan Sistem Informasi Prediksi Harga Pembelian Mobil Bekas Menggunakan Algoritma *Random Forest*. Algoritma *Random Forest* merupakan salah satu metode *machine learning* yang efektif dalam melakukan prediksi harga karena kemampuannya dalam menangani dataset yang kompleks dan menghasilkan estimasi yang akurat (Santoso et al., 2023). Dengan menggunakan pendekatan ini, sistem informasi yang dikembangkan dapat memprediksi harga beli yang optimal bagi pembeli berdasarkan data historis dan parameter kendaraan (Saadah & Salsabila, 2021). Algoritma *Random Forest* dipilih karena memiliki keunggulan dalam mengolah data dengan berbagai variabel dan mampu menghasilkan prediksi harga yang lebih akurat dibandingkan metode konvensional seperti *regresi linear* (Mu'tashim et al., 2021) (Asbullah & Samsudin, 2024). Dengan memanfaatkan data historis harga mobil bekas, sistem ini dapat mengidentifikasi pola harga berdasarkan faktor-faktor yang memengaruhi nilai kendaraan, sehingga dapat memberikan estimasi harga jual yang lebih realistis bagi penjual dan harga beli yang lebih sesuai bagi pembeli.

Beberapa penelitian terdahulu telah mengembangkan model prediksi harga mobil bekas menggunakan berbagai pendekatan *machine learning*. Seperti, penelitian yang dilakukan oleh (Kriswantara et al., 2021) mengembangkan model prediksi harga mobil bekas menggunakan metode DNN dengan tiga lapisan tersembunyi. Hasilnya menunjukkan bahwa DNN menghasilkan *Mean Absolute Error* (MAE) sebesar 501.232 dan R^2 sebesar 0,88, lebih baik dibandingkan dengan metode *Random Forest* yang menghasilkan MAE sebesar 521.947 dan R^2 sebesar 0,82. Penelitian sebelumnya oleh (B. E. S. Dewi et al., 2024) mengimplementasikan algoritma *Random Forest* 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. Penelitian yang dilakukan oleh (Alkarim & Zuama, 2024) menerapkan algoritma KNN untuk memprediksi harga mobil bekas. Hasil penelitian menunjukkan bahwa KNN memberikan prediksi dengan tingkat kesalahan yang rendah, dengan *Root Mean Square Error* (RMSE) sebesar 7.852.159, *Mean Absolute Error* (MAE) sebesar 6.172.682, dan *Mean Squared Error* (MSE) sebesar 61.656.409.756.097,56. Dari berbagai penelitian di atas, terlihat bahwa algoritma *machine learning* seperti DNN, KNN, regresi linear, dan *Random Forest* telah banyak digunakan untuk memprediksi harga mobil bekas. Namun, penelitian yang mempertimbangkan perbedaan perspektif antara penjual dan pembeli dalam prediksi harga masih terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan sistem informasi yang mampu memberikan prediksi harga yang lebih akurat dengan mempertimbangkan faktor spesifik dari kedua pihak menggunakan algoritma *Random Forest*.

Penyelesaian dari permasalahan ini melibatkan pengembangan sistem informasi yang dapat mengotomatisasi proses prediksi harga mobil bekas di Agen SQL. Sistem ini akan mengintegrasikan data historis harga mobil bekas dengan algoritma *Random Forest*, sehingga dapat memberikan rekomendasi harga yang lebih akurat dan berbasis data. Dengan penerapan sistem ini, diharapkan Agen SQL dapat menetapkan harga jual yang lebih kompetitif, sementara pembeli mendapatkan transparansi dalam harga yang ditawarkan. Selain itu, sistem ini juga dapat meningkatkan efisiensi dalam proses negosiasi, mengurangi kesenjangan harga antara penjual dan pembeli, serta membantu Agen SQL dalam meningkatkan volume transaksi. Dengan demikian, penerapan sistem informasi ini diharapkan dapat membawa manfaat bagi seluruh pihak yang terlibat dalam jual beli mobil bekas..

2. METODOLOGI PENELITIAN

2.1 Metode Penelitian

Dalam penelitian ini penulis menggunakan pendekatan penelitian kuantitatif. Metode kuantitatif adalah pendekatan penelitian yang menggunakan data numerik dan teknik statistik untuk menguji hipotesis atau menjawab pertanyaan penelitian. Data dikumpulkan melalui instrumen seperti kuesioner, survei, atau pengukuran langsung, dan dianalisis menggunakan metode statistik untuk menemukan pola atau hubungan dalam data (Waruwu, 2023).

Dalam penelitian ini, metode kuantitatif digunakan untuk menganalisis dan memprediksi harga pembelian mobil bekas menggunakan algoritma *Random Forest*. Proses dimulai dengan pengumpulan data berupa harga jual dan beli mobil, serta variabel independen seperti merek, model, tahun pembuatan, jarak tempuh, jenis bahan bakar, kapasitas mesin, dan kondisi kendaraan (Ridwan et al., 2024). Data yang diperoleh kemudian melalui tahap preprocessing, termasuk pembersihan data, normalisasi, dan encoding variabel kategorikal agar dapat digunakan dalam model prediksi. Selanjutnya, dataset dibagi menjadi data latih dan data uji untuk memastikan model dapat melakukan prediksi secara akurat (Farhanuddin et al., 2024).

2.2 Tahapan Penelitian

Tahapan penelitian dalam pengembangan sistem prediksi harga mobil bekas ini disusun secara sistematis agar proses penelitian lebih terstruktur. Adapun tahapan penelitian yang dilakukan adalah sebagai berikut (Muna et al., 2022).

1. Pengumpulan Data

Data dikumpulkan dari dataset mobil bekas yang mencakup atribut seperti merek, model, tahun produksi, kapasitas mesin, jarak tempuh, bahan bakar, transmisi, dan wilayah.

2. Preprocessing Data

Tahap ini meliputi pembersihan data (*data cleaning*), penanganan missing value, encoding variabel kategorikal, serta normalisasi data agar siap digunakan dalam proses pemodelan.

3. Pembagian Dataset

Dataset dibagi menjadi data latih (*training*) dan data uji (*testing*) untuk menghindari *overfitting* dan menguji performa model.

4. Pembangunan Model

Model dibangun menggunakan algoritma *Random Forest* dengan memanfaatkan teknik *ensemble learning* untuk meningkatkan akurasi prediksi.

5. Evaluasi Model

Model dievaluasi menggunakan metrik *Mean Absolute Error* (MAE) dan *R-Squared* (R^2) untuk mengukur tingkat kesalahan dan kemampuan model dalam menjelaskan variasi data.

6. Implementasi Sistem

Model yang telah dievaluasi kemudian diimplementasikan ke dalam sistem informasi berbasis web untuk digunakan oleh pengguna.

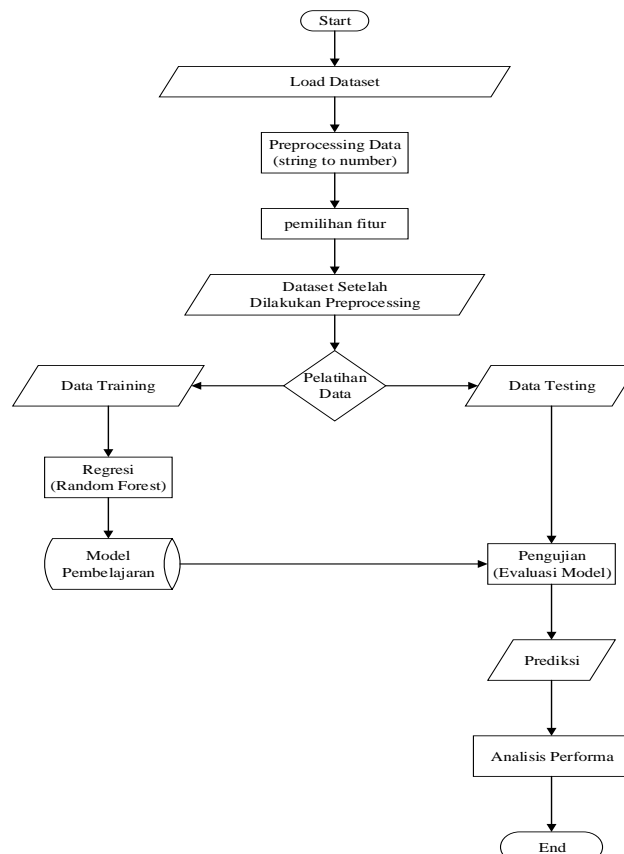
2.3 Algoritma *Random Forest*

Random Forest adalah algoritma *machine learning* berbasis *ensemble* yang digunakan untuk tugas klasifikasi dan *regresi*. Algoritma ini bekerja dengan membangun sejumlah besar pohon keputusan (*decision trees*) dan menggabungkan hasilnya untuk meningkatkan akurasi prediksi serta mengurangi risiko *overfitting*. Setiap pohon keputusan dilatih menggunakan subset data yang dipilih secara acak dengan teknik *bagging* (*bootstrap aggregating*), sehingga setiap pohon memiliki variasi dalam proses pembelajarannya (Aprilia et al., 2021).

Pada tahap prediksi, untuk kasus klasifikasi, hasil akhir ditentukan berdasarkan mayoritas suara (*majority voting*) dari seluruh pohon. Sementara itu, dalam tugas *regresi*, nilai akhir diperoleh dengan menghitung rata-rata hasil prediksi dari semua pohon. Metode ini membuat *Random Forest* lebih stabil dibandingkan dengan satu pohon keputusan tunggal yang cenderung rentan terhadap perubahan kecil dalam data pelatihan (Hadi & Benedict, 2024).

Keunggulan utama *Random Forest* meliputi kemampuannya dalam menangani data dengan banyak fitur, bekerja dengan baik pada dataset yang memiliki nilai yang hilang (*missing values*), serta memiliki ketahanan terhadap *overfitting*. Namun, kelemahannya adalah konsumsi komputasi yang lebih tinggi dibandingkan dengan model sederhana seperti *Decision Tree*, serta sulitnya dalam melakukan interpretasi terhadap model yang dihasilkan (B. E. S. Dewi et al., 2024).

Dalam implementasi nyata, *Random Forest* sering digunakan dalam berbagai bidang, seperti prediksi harga, deteksi penyakit, klasifikasi citra, dan analisis pasar. Dengan pendekatan berbasis statistik dan *ensemble learning*, algoritma ini menjadi salah satu metode yang populer dalam dunia data *science* dan kecerdasan buatan (Warjiyono et al., 2024).



Gambar 1. Flowchart *Random Forest*

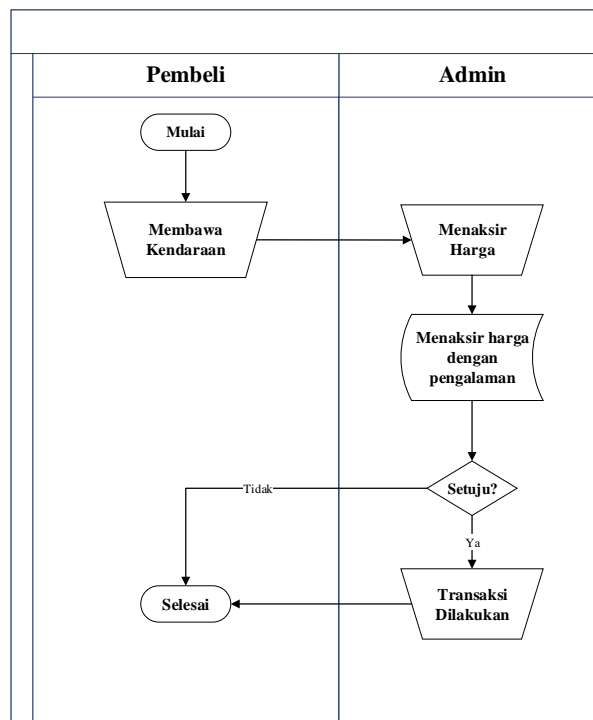
Gambar 2 merupakan proses prediksi harga pembelian menggunakan metode *Random Forest*. Proses dimulai dengan memuat dataset pembelian, diikuti oleh preprocessing data untuk memastikan semua data berada dalam format yang dapat digunakan, mengubah string menjadi angka. Selanjutnya, fitur-fitur signifikan seperti *Previous Close*, *Open Price*, *High*, *Low*, *Change*, dan *Volume* dipilih untuk digunakan dalam model. Dataset kemudian dibagi menjadi data training dan testing untuk melatih dan menguji model (Muhamad Malik Matin, 2023). Model *Random Forest* dilatih dengan data training untuk mempelajari pola-pola yang ada, lalu diuji dengan data testing untuk mengevaluasi akurasi model menggunakan metrik seperti *Mean Absolute Percentage Error* (MAPE) dan *R-squared*. Setelah model divalidasi, digunakan untuk melakukan prediksi harga pembelian, yang hasilnya dianalisis untuk menilai kinerja dan akurasi model (Riskiyono & Mahdiana, 2024).

3. HASIL DAN PEMBAHASAN

3.1 Analisis Sistem Berjalan

Dataset yang digunakan dalam penelitian ini terdiri dari 301 data mobil bekas yang diperoleh dari data historis transaksi. Setiap data memiliki beberapa atribut penting yang memengaruhi harga kendaraan, yaitu merek mobil, model, tahun produksi, kapasitas mesin, jarak tempuh, jenis bahan bakar, transmisi, dan wilayah penjualan.

Dataset ini digunakan sebagai dasar dalam proses pelatihan model *Random Forest* untuk menemukan pola hubungan antara atribut kendaraan dengan harga mobil bekas. Sebelum digunakan, dataset melalui tahap preprocessing untuk memastikan kualitas data tetap terjaga dan sesuai untuk proses pemodelan.



Gambar 3. Analisis Sistem Berjalan

Gambar 3 menunjukkan analisis sistem berjalan pada proses penentuan harga mobil bekas di Agen SQL. Analisis sistem berjalan proses penentuan harga beli mobil bekas pada Agen SQL yang masih dilakukan secara manual. Pada sistem yang sedang berjalan, terdapat dua aktor utama yaitu pembeli dan admin. Proses dimulai ketika pembeli membawa kendaraan ke Agen SQL untuk dinilai harganya. Selanjutnya, admin melakukan proses penaksiran harga berdasarkan kondisi fisik kendaraan seperti tahun pembuatan, jarak tempuh, serta kondisi mesin dan body. Penentuan harga ini masih bergantung pada pengalaman pribadi admin, tanpa menggunakan sistem prediksi atau perhitungan berbasis data yang terukur. Setelah harga ditentukan, admin menyampaikan hasil penilaian kepada pembeli untuk dimintai persetujuan. Apabila pembeli menyetujui harga yang ditawarkan, maka transaksi dilakukan dan proses berakhir. Namun, jika pembeli tidak setuju, maka transaksi dibatalkan dan kegiatan dinyatakan selesai. Berdasarkan analisis ini, dapat disimpulkan bahwa sistem yang berjalan masih bersifat manual dan subjektif, sehingga sering menimbulkan ketidaksesuaian harga antara pembeli dan admin. Oleh karena itu, diperlukan pengembangan sistem informasi prediksi harga beli mobil bekas yang dapat memberikan estimasi harga lebih objektif, efisien, dan akurat berdasarkan data historis menggunakan algoritma *Random Forest*.

Data yang diperoleh untuk pelatihan dan pengujian model harus dijelaskan dengan tepat sebelum membuat model sehingga model dapat mempelajari polanya dengan lebih cepat. Tahap pre-processing akan membuat dataset akhir yang akan dimasukkan ke dalam model yang dibuat. Data yang diolah dalam penelitian ini sebanyak 301 data.

merk	model	cc_mesin	kondisi	kilometer	tahun	harga
Mitsubishi	Xpander	1500	Bekas Sangat Baik	126958	2023	372
Mitsubishi	Xpander	1500	Bekas Baik	124879	2019	320
Suzuki	Ertiga	1300	Bekas Sangat Baik	115268	2013	92
Daihatsu	Sigra	1200	Bekas Sangat Baik	173266	2017	206
Mitsubishi	Pajero	2400	Bekas Baik	92498	2022	375
Daihatsu	Xenia	1500	Bekas Sangat Baik	21023	2006	89
Honda	Civic	1600	Bekas Cukup	46090	2006	87
Honda	Civic	1600	Bekas Sangat Baik	10311	2018	273
Toyota	Avanza	1300	Bekas Cukup	108355	2018	156
....
Daihatsu	Sigra	1200	Bekas Baik	178880	2017	134
Daihatsu	Xenia	1000	Bekas Baik	189943	2018	87
Toyota	Avanza	1500	Bekas Cukup	6252	2004	87

Gambar 4. Preprocessing Data

3.2 Algoritma Random Forest

Algoritma *Random Forest* merupakan algoritma *ensemble learning*. *Ensemble learning* adalah suatu teknik dengan menggabungkan beberapa model untuk melakukan prediksi. *Random forest* merupakan bentuk implementasi dari *homogeneous ensemble learning* yang menggabungkan beberapa model sejenis yaitu *decision tree* (pohon keputusan) (Haryanto et al., 2023).

Klasifikasi *random forest* dilakukan dengan menggabungkan pohon-pohon melalui pelatihan, yang menghasilkan pohon klasifikasi dengan banyak versi, yang kemudian digabungkan untuk mendapatkan klasifikasi akhir. Dalam *Random Forest*, proses pengacakan untuk membentuk pohon klasifikasi dilakukan untuk data sampel dan variabel prediktor. Penelitian ini terdiri dari beberapa tahap: pengumpulan dataset saham, *preprocessing* data, *feature selection*, pembagian dataset, pembentukan pohon keputusan dan prediksi menggunakan *random forest* (Kriswantara et al., 2021).

3.3 Evaluasi Model

Evaluasi model dilakukan untuk mengetahui tingkat akurasi model *Random Forest* dalam memprediksi harga mobil bekas. Pada penelitian ini digunakan dua metrik evaluasi yaitu *Mean Absolute Error* (MAE) dan *R-Squared* (R²). MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai aktual dan nilai prediksi, sedangkan R² digunakan untuk mengetahui seberapa besar kemampuan model dalam menjelaskan variasi data.

1. Perhitungan Mean Absolute Error (MAE)

Rumus *Mean Absolute Error* (MAE) dapat dinyatakan sebagai berikut:

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

Pada rumus tersebut, y_i menyatakan nilai aktual, \hat{y}_i merupakan nilai prediksi yang dihasilkan oleh model, sedangkan \bar{y} adalah nilai rata-rata dari seluruh data aktual yang digunakan sebagai pembanding dalam perhitungan koefisien determinasi (R²). Tabel 1 menunjukkan contoh hasil prediksi model *Random Forest* pada data pengujian.

Tabel 1. Contoh Hasil Prediksi Model *Random Forest*

No	Harga Aktual (y _i)	Harga Prediksi (ŷ _i)	Error	(y _i - ŷ _i) ²
1	372	350	22.00	484
2	320	298	22.00	484
3	92	70	22.00	484
4	206	183	23.00	529
5	375	352	23.00	529
6	89	66	23.00	529
7	87	64	23.00	529
8	273	250	23.00	529
9	156	133	23.00	529
10	134	110.46	22.54	507.33

Berdasarkan tabel tersebut diperoleh total nilai error sebagai berikut:

$$\sum |y_i - \hat{y}_i| = 22 + 22 + 22 + 23 + 23 + 23 + 23 + 23 + 23 + 22.54$$

$$\sum |y_i - \hat{y}_i| = 226.54$$

Selanjutnya nilai MAE dihitung menggunakan rumus sebagai berikut:

$$MAE = \frac{226.54}{10}$$

$$MAE = 22.654$$

Nilai MAE sebesar 22.654 menunjukkan bahwa rata-rata kesalahan prediksi harga mobil yang dihasilkan oleh model Random Forest adalah sekitar 22.654 juta rupiah.

2. Perhitungan *R-Squared* (R^2)

R-Squared digunakan untuk mengetahui seberapa baik model dalam menjelaskan variasi data harga mobil. Rumus *R-Squared* dapat dinyatakan sebagai berikut:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \tag{2}$$

Berdasarkan rumus diatas, \bar{y} merupakan rata-rata nilai aktual.

a. Langkah pertama adalah menghitung rata-rata harga aktual:

$$\bar{y} = \frac{372+320+92+206+375+89+87+273+156+134}{10} = 210.4$$

b. Selanjutnya dihitung jumlah kuadrat error:

$$\sum(y_i - \hat{y}_i)^2 = 5104.33$$

c. dan jumlah kuadrat total:

$$\sum(y_i - \bar{y})^2 = 121871$$

d. Sehingga nilai R^2 dapat dihitung sebagai berikut:

$$R^2 = 1 - \frac{5104.33}{121871}$$

$$R^2 = 1 - 0.0418$$

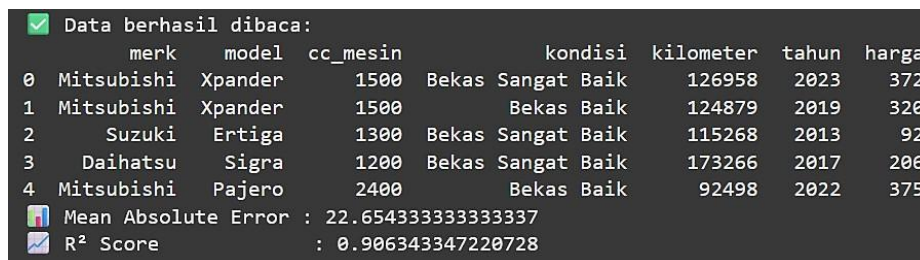
$$R^2 = 0.958$$

Berdasarkan hasil perhitungan evaluasi model diperoleh nilai MAE dan *R-Squared* sebagai berikut.

Tabel 2. Hasil Evaluasi Model

Metode	MAE	R^2
Random Forest	22.654	0.958

Nilai MAE sebesar 22.654 menunjukkan bahwa rata-rata kesalahan prediksi harga mobil yang dihasilkan oleh model adalah sekitar 22.654 juta rupiah. Selain itu, nilai R^2 sebesar 0.958 menunjukkan bahwa model Random Forest mampu menjelaskan variasi data harga mobil bekas sebesar 95.8%. Hal ini menunjukkan bahwa model yang dibangun memiliki tingkat akurasi yang sangat baik dalam melakukan prediksi harga mobil bekas.



merk	model	cc_mesin	kondisi	kilometer	tahun	harga
0 Mitsubishi	Xpander	1500	Bekas Sangat Baik	126958	2023	372
1 Mitsubishi	Xpander	1500	Bekas Baik	124879	2019	320
2 Suzuki	Ertiga	1300	Bekas Sangat Baik	115268	2013	92
3 Daihatsu	Sigra	1200	Bekas Sangat Baik	173266	2017	206
4 Mitsubishi	Pajero	2400	Bekas Baik	92498	2022	375
Mean Absolute Error : 22.654333333333337						
R ² Score : 0.906343347220728						

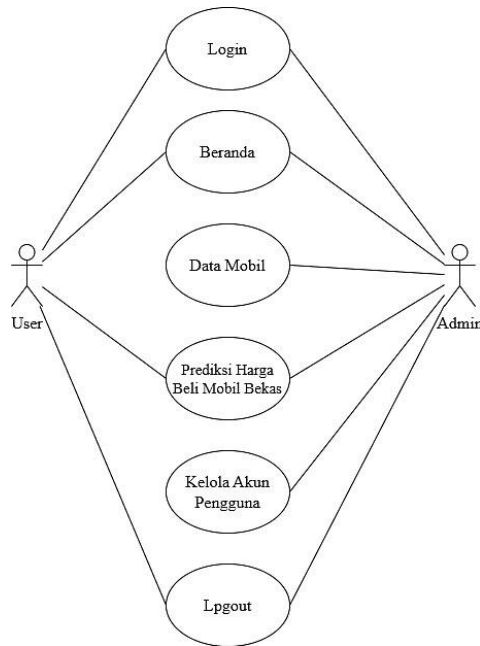
Gambar 5. Evaluasi Model

3.4 Unified Modelling Language (UML)

Tahap ini akan menjelaskan proses perancangan sistem dengan menggunakan *use case diagram*, *activity diagram*, dan *class diagram* yang merupakan pemodelan diagram *Unifield Modelling Language* (UML) (Ayunandita & Dadi Riskiono, 2021).

1. *Use Case Diagram*

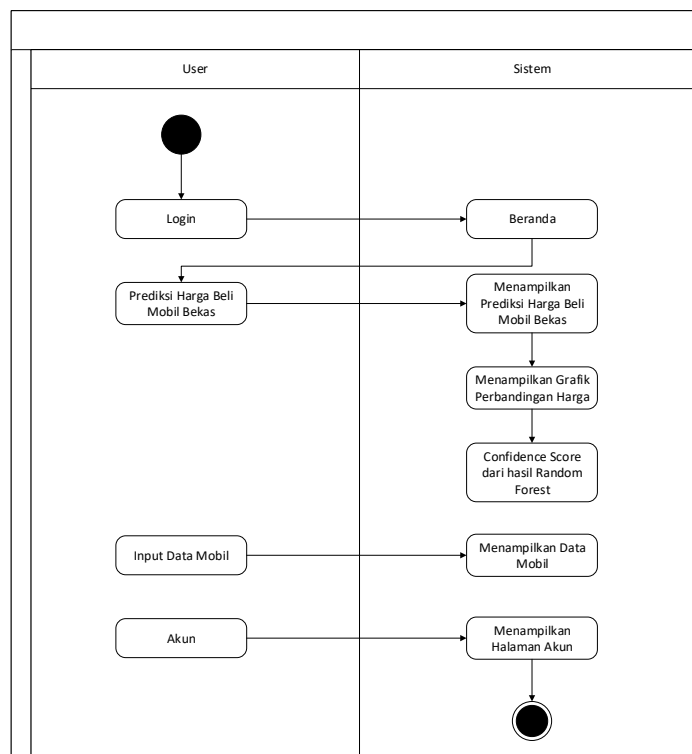
Berikut merupakan *use case diagram* yang digunakan untuk menjelaskan interaksi antar aktor yaitu *admin* dan pembeli dan penjual mobil bekas dalam penggunaan sistem. Dalam membangun sistem informasi prediksi di tunjukan pada Gambar dibawah ini:



Gambar 6. Use Case Diagram

2. Activity Diagram

Activity diagram adalah sebuah flowchart tentang aktivitas apa saja yang akan dilakukan oleh user Agen SQL.

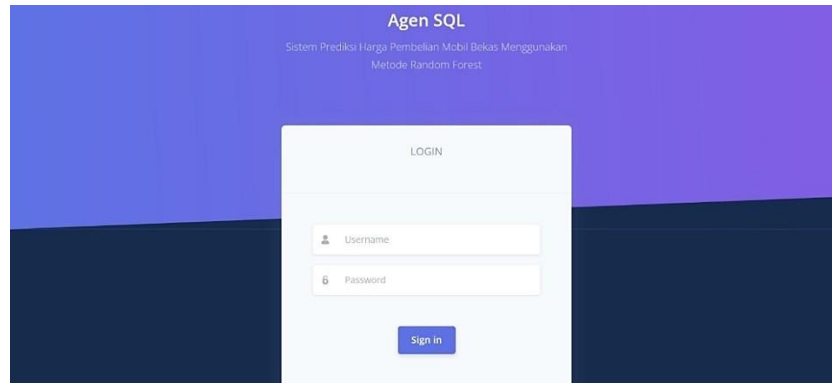


Gambar 7. Activity Diagram

3.5 Implementasi Prediksi Harga Pembelian Mobil Bekas

Hasil implementasi sistem prediksi harga pembelian mobil bekas menunjukkan tampilan halaman utama aplikasi yang dirancang untuk membantu pengguna dalam memperkirakan harga kendaraan secara lebih objektif. Aplikasi ini memungkinkan pengguna untuk memasukkan beberapa parameter penting yang memengaruhi nilai mobil di pasaran, seperti model kendaraan, tahun produksi, kapasitas mesin, jarak tempuh, jenis transmisi, bahan bakar, dan wilayah penjualan. Setelah seluruh data dimasukkan, pengguna dapat menekan tombol prediksi untuk memperoleh estimasi harga mobil bekas yang dihitung berdasarkan model algoritma Random Forest. Melalui fitur ini, sistem diharapkan dapat memberikan hasil prediksi yang akurat, sehingga membantu pengguna dalam mengambil keputusan pembelian yang lebih tepat dan efisien.

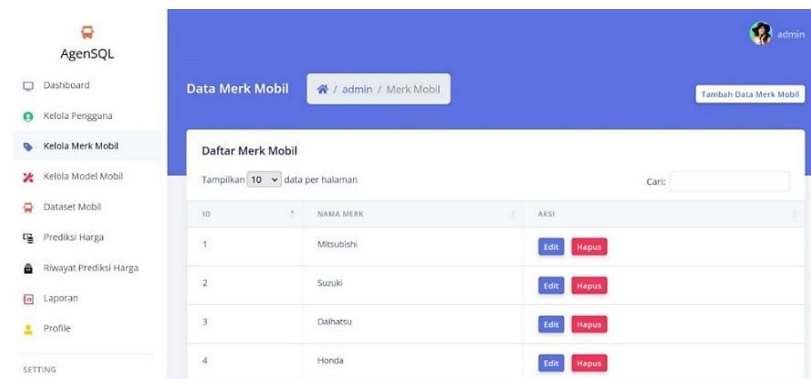
1. Halaman Login



Gambar 8. Tampilan Halaman Login

Gambar 8 menampilkan halaman login sistem informasi prediksi harga pembelian mobil bekas yang dikembangkan untuk Agen SQL dengan menggunakan metode Random Forest. Halaman ini berfungsi sebagai gerbang awal bagi pengguna untuk mengakses sistem, di mana setiap pengguna diharuskan memasukkan username dan password yang valid sebagai bentuk autentikasi. Desain antarmuka login dibuat dengan tampilan sederhana dan profesional untuk memudahkan pengguna dalam proses masuk ke dalam sistem. Melalui halaman ini, sistem memastikan bahwa hanya pengguna yang memiliki hak akses tertentu, seperti admin atau pengguna terdaftar, yang dapat menggunakan fitur prediksi harga pembelian mobil bekas secara aman dan terkontrol.

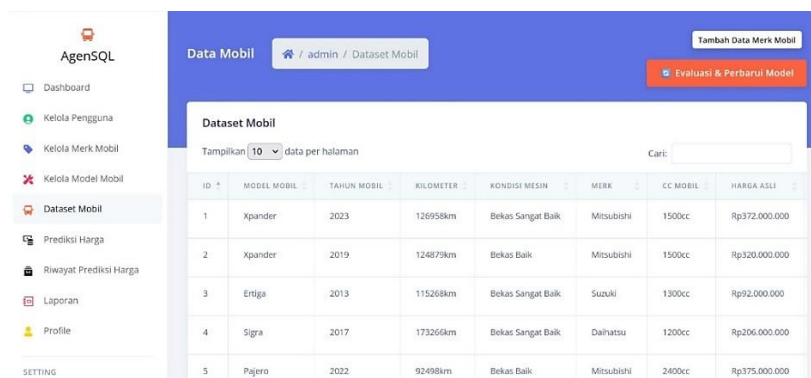
2. Halaman Merk Mobil



Gambar 9. Halaman Merk Mobil

Gambar 9 menampilkan tampilan halaman data merk mobil pada sistem informasi prediksi harga pembelian mobil bekas Agen SQL. Halaman ini berfungsi untuk mengelola daftar merk kendaraan yang menjadi salah satu parameter dalam proses prediksi harga. Terdapat fitur “Edit” untuk memperbarui data merk mobil dan tombol “Hapus” untuk menghapus data yang tidak lagi relevan. Selain itu, tersedia pula tombol “Tambah Data Merk Mobil” di bagian kanan atas untuk menambahkan entri baru. Fitur pencarian dan opsi jumlah data per halaman juga disediakan untuk memudahkan pengelolaan data secara efisien. Dengan adanya halaman ini, sistem dapat menjaga kelengkapan dan konsistensi data merk mobil yang akan digunakan dalam proses prediksi harga berbasis algoritma *Random Forest*.

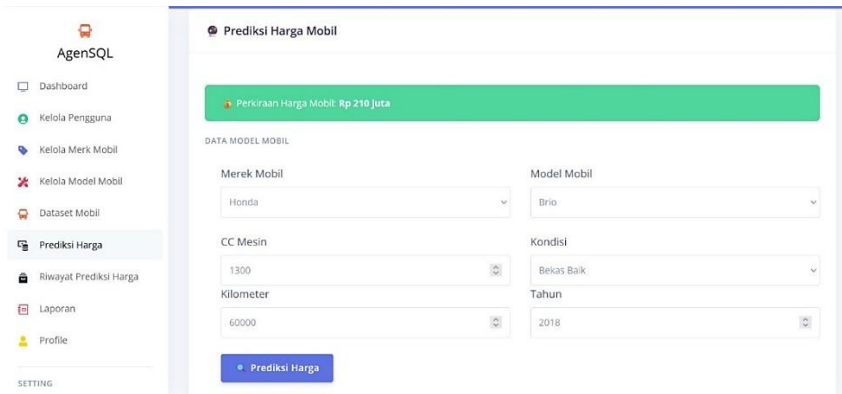
3. Halaman Dataset Mobil



Gambar 10. Halaman Dataset Mobil

Gambar 10 menampilkan halaman dataset mobil pada sistem informasi prediksi harga pembelian mobil bekas Agen SQL. Halaman ini berfungsi untuk menampilkan dan mengelola data kendaraan yang menjadi dasar dalam proses pelatihan dan prediksi harga menggunakan algoritma *Random Forest*. Selain menampilkan data, halaman ini juga menyediakan tombol “Evaluasi & Perbarui Model” yang berfungsi untuk memperbarui model prediksi berdasarkan dataset terbaru, serta tombol “Tambah Data Merek Mobil” untuk menambah entri kendaraan baru. Melalui fitur ini, admin dapat memastikan bahwa sistem selalu menggunakan data terkini dan relevan, sehingga hasil prediksi harga pembelian mobil bekas menjadi lebih tepat, objektif, dan adaptif terhadap perubahan pasar.

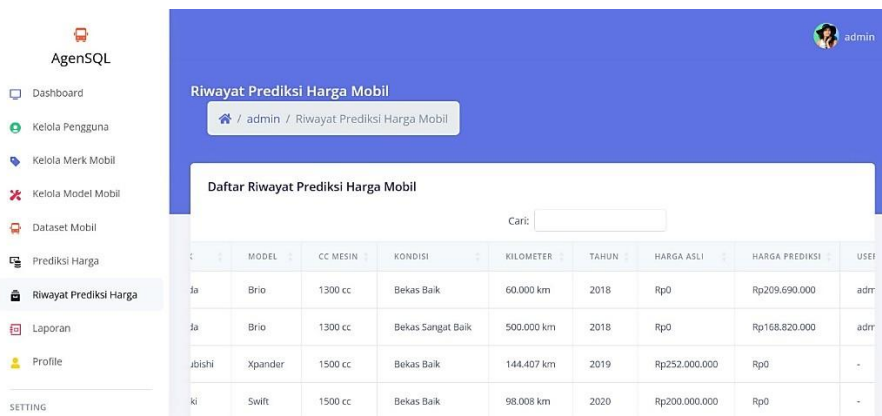
4. Halaman Prediksi Mobil



Gambar 11. Halaman Prediksi Harga

Gambar 11 menampilkan halaman prediksi harga mobil pada sistem Agen SQL. Pada halaman ini, pengguna dapat memasukkan data kendaraan seperti merek, model, kapasitas mesin, kondisi, jarak tempuh, dan tahun produksi untuk memperoleh perkiraan harga mobil bekas. Setelah semua parameter diisi, sistem menampilkan hasil prediksi harga secara otomatis berdasarkan model yang dilatih menggunakan algoritma *Random Forest*, seperti contoh pada gambar yang memperkirakan harga mobil sebesar Rp210 juta.

5. Halaman Riwayat Prediksi



MODEL	CC MESIN	KONDISI	KILOMETER	TAHUN	HARGA ASLI	HARGA PREDIKSI	USER
Brio	1300 cc	Bekas Baik	60.000 km	2018	Rp0	Rp209.690.000	admin
Brio	1300 cc	Bekas Sangat Baik	500.000 km	2018	Rp0	Rp168.820.000	admin
Xpander	1500 cc	Bekas Baik	144.407 km	2019	Rp252.000.000	Rp0	-
Swift	1500 cc	Bekas Baik	98.008 km	2020	Rp200.000.000	Rp0	-

Gambar 12. Halaman Riwayat Prediksi

Gambar 12 menampilkan halaman riwayat prediksi harga mobil pada sistem Agen SQL. Halaman ini berfungsi untuk menampilkan data hasil prediksi yang telah dilakukan sebelumnya, lengkap dengan informasi model mobil, kapasitas mesin, kondisi kendaraan, jarak tempuh, tahun produksi, serta hasil prediksi harga. Melalui fitur ini, admin dapat memantau dan mengevaluasi hasil prediksi yang telah dihasilkan oleh sistem berbasis algoritma *Random Forest*, sehingga memudahkan proses analisis, perbandingan, serta peningkatan akurasi model prediksi di masa mendatang..

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa pengembangan sistem informasi prediksi harga pembelian mobil bekas menggunakan algoritma *Random Forest* berhasil mencapai tujuan penelitian, yaitu menghasilkan estimasi harga yang objektif dan berbasis data. Model yang dibangun menunjukkan kinerja yang sangat baik dengan nilai Mean Absolute Error (MAE) sebesar 22.654 dan nilai R-Squared (R^2) sebesar 0.958, yang berarti model mampu menjelaskan variasi data harga mobil bekas sebesar 95,8% dengan tingkat kesalahan prediksi yang relatif rendah. Dari sisi sistem informasi, sistem yang dikembangkan telah berjalan secara fungsional dengan mekanisme input–proses–output. Pada tahap input, pengguna memasukkan data kendaraan seperti merek, model, tahun produksi,



kapasitas mesin, jarak tempuh, jenis bahan bakar, transmisi, dan wilayah. Pada tahap proses, sistem melakukan pengolahan data menggunakan model Random Forest yang telah dilatih dengan dataset historis. Sedangkan pada tahap output, sistem menghasilkan prediksi harga mobil bekas secara otomatis dan cepat. Dengan demikian, sistem ini mampu mengurangi subjektivitas dalam penentuan harga, meningkatkan akurasi estimasi, serta membantu pengguna dalam pengambilan keputusan yang lebih tepat. Selain itu, sistem juga memberikan kontribusi dalam meningkatkan transparansi dan efisiensi dalam proses transaksi mobil bekas.

REFERENCES

- Alkarim, N., & Zuama, R. A. (2024). Implementasi Algoritma Knn (K-NEARS NEIGHBOR) Pada Prediksi Harga Mobil Bekas. *Jurnal IJCCS*, 3(1), 1–7.
- Apriliyah, W., Kurniawan, I., Baydhowi, M., & Haryati, T. (2021). Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest. *Sistemasi*, 10(1), 163. <https://doi.org/10.32520/stmsi.v10i1.1129>
- Asbullah, J., & Samsudin. (2024). Prediksi Harga Cryptocurrency Binance Berdasarkan Informasi Blockchain dengan Menggunakan Algoritma Random Forest. *Jurnal Media Informatika Budidarma*, 8(1), 260–271. <https://doi.org/10.30865/mib.v8i1.7100>
- Ayunandita, N., & Dadi Riskiono, S. (2021). Permodelan Sistem Informasi Akademik Menggunakan Extreme Programming Pada Madrasah Aliyah (Ma) Mambaul Ulum Tanggamus. *Jurnal Informatika Dan Rekayasa Perangkat Lunak (JATIKA)*, 2(2), 196–204.
- Dewi, B. E. S., Haikal, S., Sulistyowati, H. S., Fitriani, R., & K, D. P. (2024). Penerapan Machine Learning Menggunakan Algoritma Random Forest Untuk Prediksi. *JURNAL TRIDI Teknologi Informatika & Komputer*, 2(1), 20–31.
- Dewi, E., Mulyani, S., Mulady, F., Ramadhan, D., Ariyantono, A., & Ramdani, D. (2020). Estimasi Harga Jual Mobil Bekas Menggunakan Metode Regresi Linier Berganda. *E-Jurnal JUSITI (Jurnal Sistem Informasi Dan Teknologi Informasi)*, 9(1), 1–8. <https://doi.org/10.36774/jusiti.v9i1.649>
- Farhanuddin, Sarah Ennola Karina Sihombing, & Yahfizham. (2024). Komparasi Multiple Linear Regression dan Random Forest Regression Dalam Memprediksi Anggaran Biaya Manajemen Proyek Sistem Informasi. *Journal of Computers and Digital Business*, 3(2), 86–97. <https://doi.org/10.56427/jcbd.v3i2.408>
- Hadi, N., & Benedict, J. (2024). Implementasi Machine Learning Untuk Prediksi Harga Rumah Menggunakan Algoritma Random Forest. *Computatio : Journal of Computer Science and Information Systems*, 8(1), 50–61. <https://doi.org/10.24912/computatio.v8i1.15173>
- Haryanto, C., Rahaningsih, N., & Muhammad Basysyar, F. (2023). Komparasi Algoritma Machine Learning Dalam Memprediksi Harga Rumah. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 533–539. <https://doi.org/10.36040/jati.v7i1.6343>
- Hasibuan, E., Informasi, S., Ilmu, F., Informasi, T., Gunadarma, U., Margonda, J., No, R., Cina, P., & Jawa, D. (2022). Implementasi Machine Learning untuk Prediksi Harga Mobil Bekas dengan Algoritma Regresi Linear berbasis Web. *Jurnal Ilmiah Komputasi*, 21(4), 595–602. <https://doi.org/10.32409/jikstik.21.4.3327>
- Kriswantara, B., Kurniawati, & Pardede, H. F. (2021). Prediksi Harga Mobil Bekas Dengan Machine Learning. *Syntax Literate: Jurnal Ilmiah Indonesia*, 6(5).
- Mu'tashim, M. L., Muhayat, T., Damayanti, S. A., Zaki, H. N., & Wirawan, R. (2021). Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Multiple Linear Regression. *Informatik : Jurnal Ilmu Komputer*, 17(3), 238–245.
- Muhamad Malik Matin, I. (2023). Hyperparameter Tuning Menggunakan GridsearchCV pada Random Forest untuk Deteksi Malware. *Multinetics*, 9(1), 43–50. <https://doi.org/10.32722/multinetics.v9i1.5578>
- Muna, N., Herdayanti, D. M., Erawantini, F., & Yunus, M. (2022). Pemetaan Persebaran dan Pengukuran Balita Stunting di Kabupaten Jember Berbasis Sistem Informasi Geografis. 13(November), 153–160.
- Ridwan, A., Muzakir, U., & Nurhidayati, S. (2024). Optimizing E-commerce Inventory to prevent Stock Outs using the Random Forest Algorithm Approach. *International Journal Software Engineering and Computer Science (IJSECS)*, 4(1), 107–120. <https://doi.org/10.35870/ijsecs.v4i1.2326>
- Riskiyono, F., & Mahdiana, D. (2024). Implementation of Random Forest Algorithm for Graduation Prediction. *Sinkron*, 8(3), 1662–1670. <https://doi.org/10.33395/sinkron.v8i3.13750>
- Saadah, S., & Salsabila, H. (2021). Prediksi Harga Bitcoin Menggunakan Metode Random Forest. *Jurnal Komputer Terapan*, 7(1), 24–32. <https://doi.org/10.35143/jkt.v7i1.4618>
- Santoso, H., Putri, R. A., & Sahbandi, S. (2023). Deteksi Komentar Cyberbullying pada Media Sosial Instagram Menggunakan Algoritma Random Forest. *Jurnal Manajemen Informatika (JAMIKA)*, 13(1), 62–72. <https://doi.org/10.34010/jamika.v13i1.9303>
- Warjiyono, Rais, A. N., Alfarobi, I., Hadi, S. W., & Kurniawan, W. (2024). Analisa Prediksi Harga Jual Rumah Menggunakan Algoritma Random Forest Machine Learning. *JURSISTEKNI (Jurnal Sistem Informasi Dan Teknologi Informasi)*, 6(2), 416–423.
- Waruwu, M. (2023). Pendekatan Penelitian Pendidikan: Metode Penelitian Kualitatif, Metode Penelitian Kuantitatif dan Metode Penelitian Kombinasi (Mixed Method). *Jurnal Pendidikan Tambusai*, 7(1), 2896–2910.