

# Estimasi Harga Mobil Bekas Menggunakan Fuzzy Mamdani pada Domain Homogen

Rian Oktavianto Husain<sup>1</sup>, Nursetia Wati<sup>\*2</sup>,

<sup>1</sup>Program Studi Teknologi Informasi, Universitas Gadjah Mada, Yogyakarta, Indonesia

<sup>\*2</sup>Program Studi Teknik Informatika, Politeknik Gorontalo, Gorontalo, Indonesia

Email: <sup>1</sup>rianoktaviantohusain@mail.ugm.ac.id, <sup>\*2</sup>tia@poligon.ac.id

**Abstrak**—Ketidakpastian dalam penetapan harga kendaraan bekas sering menimbulkan asimetri informasi yang merugikan pembeli maupun penjual. Penelitian ini merancang sistem estimasi harga mobil bekas menggunakan *Fuzzy Inference System* metode Mamdani dengan tujuh variabel input, yaitu tahun kendaraan, jarak tempuh, status kepemilikan, jenis transmisi, kapasitas mesin, jenis bahan bakar, dan tipe penjual. Penelitian difokuskan pada kendaraan merek Hyundai, Honda, dan Toyota menggunakan 639 data hasil penyaringan dari dataset *Car Details V4 Kaggle* yang berjumlah 2059 data. Pembatasan domain dilakukan untuk meningkatkan presisi *knowledge base* terhadap karakteristik harga kendaraan segmen menengah. *Knowledge base* dibangun menggunakan 17 aturan *fuzzy* berlapis yang mencakup matriks inti tahun dan kapasitas mesin, penalti depresiasi, serta apresiasi kondisi tertentu. Proses defuzzifikasi menggunakan metode *centroid of area* dengan batas output 5.000.000 INR. Hasil evaluasi menunjukkan akurasi kategorisasi sebesar 54%, *hit rate* toleransi  $\pm 30\%$  sebesar 50,55%, serta *F1-score* tertinggi 0,70 pada kategori Sangat Murah dengan *precision* 0,85. Nilai *hit rate* tersebut lebih tinggi dibandingkan model generik yang mencakup 33 merek kendaraan dengan *hit rate*  $\pm 30\%$  sebesar 26,23%. Hasil penelitian menunjukkan bahwa pembatasan domain pada kendaraan dengan karakteristik harga serupa membantu meningkatkan konsistensi *knowledge base* dan performa estimasi sistem.

**Kata Kunci**—*Fuzzy Inference System*, Fuzzy Mamdani, Harga Mobil Bekas, Estimasi Harga, *Knowledge Base*

## I. PENDAHULUAN

PASAR kendaraan bekas merupakan segmen industri otomotif yang terus berkembang. Meningkatnya harga kendaraan baru akibat kebijakan pajak dan biaya produksi mendorong sebagian besar konsumen, terutama kalangan menengah, untuk beralih ke pasar kendaraan bekas sebagai alternatif yang lebih terjangkau [1]. Kondisi ini menciptakan ekosistem pasar yang dinamis namun rentan terhadap asimetri informasi, di mana pembeli dan penjual sering kali tidak memiliki acuan harga yang objektif dan terstandar.

Penetapan harga kendaraan bekas merupakan proses yang kompleks karena dipengaruhi oleh berbagai faktor seperti tahun produksi, jarak tempuh, kapasitas mesin, jenis bahan bakar, transmisi, status kepemilikan, dan tipe penjual [2],[3]. Kompleksitas hubungan antarvariabel

tersebut menyebabkan proses estimasi harga sering bersifat subjektif dan tidak konsisten.

Berbagai pendekatan komputasional telah dikembangkan untuk mengatasi permasalahan ini. Metode berbasis machine learning seperti *Random Forest* [4], regresi linear [5], dan *K-Nearest Neighbor* [6] telah banyak diaplikasikan dan menunjukkan akurasi prediksi numerik yang tinggi. Namun, metode-metode tersebut umumnya bersifat *black-box*, sehingga sulit diinterpretasikan dan dijelaskan kepada pengguna awam. Di sisi lain, logika *fuzzy* menawarkan pendekatan yang lebih transparan melalui representasi pengetahuan dalam bentuk aturan linguistik *IF-THEN* yang dapat dipahami secara intuitif [7].

Metode *Fuzzy Inference System* (FIS) Mamdani merupakan salah satu pendekatan inferensi *fuzzy* yang banyak digunakan dalam sistem pendukung keputusan karena mampu merepresentasikan pengetahuan pakar melalui aturan linguistik *IF-THEN* yang mudah dipahami [9]. Dibandingkan FIS Sugeno yang menghasilkan keluaran berupa fungsi matematis dan FIS Tsukamoto yang mensyaratkan fungsi keanggotaan output monoton, FIS Mamdani memberikan fleksibilitas yang lebih tinggi dalam menjelaskan hubungan antarvariabel secara transparan, sehingga sesuai untuk kasus estimasi harga kendaraan bekas.

Penelitian mengenai estimasi harga kendaraan bekas telah berkembang pesat dalam lima tahun terakhir dengan berbagai pendekatan yang berfokus pada peningkatan akurasi prediksi. Maryam dkk. [8] pada tahun 2021 menerapkan metode Fuzzy Mamdani dan Fuzzy Tsukamoto untuk menentukan harga mobil bekas dan menunjukkan bahwa pendekatan berbasis *fuzzy* mampu merepresentasikan pertimbangan pakar dalam proses penentuan harga. Pada tahun yang sama, Andrari dkk. [10] menerapkan metode Fuzzy Mamdani untuk menentukan harga jual ponsel pintar bekas dan memperoleh hasil estimasi yang cukup baik pada domain perangkat elektronik.

Selanjutnya, Siva dan A.M. [5] pada tahun 2022 menggunakan metode *Linear Regression* dan *Support Vector Machine* untuk prediksi harga kendaraan dan menunjukkan bahwa pendekatan statistik mampu menghasilkan akurasi prediksi yang tinggi. Pada tahun 2023, Kumar dkk. [11] mengembangkan model prediksi harga kendaraan bekas berbasis *Multiple Linear Regression* dan melaporkan performa prediksi yang baik pada data pengujian. Li

[4] pada tahun 2024 mengembangkan model prediksi harga kendaraan berbasis *machine learning* dan menunjukkan performa prediksi numerik yang sangat baik pada dataset kendaraan bekas berskala besar. Penelitian terbaru oleh Ahmed dkk. [2] pada tahun 2025 melakukan analisis komparatif berbagai algoritma *machine learning* untuk prediksi harga kendaraan dan menunjukkan bahwa pendekatan berbasis data mampu menghasilkan tingkat akurasi yang tinggi pada berbagai kondisi pasar.

Meskipun berbagai penelitian berhasil meningkatkan performa prediksi harga kendaraan, sebagian besar masih berorientasi pada optimasi akurasi numerik dan kurang menekankan aspek interpretabilitas model. Penelitian berbasis *fuzzy* yang telah dilakukan umumnya menggunakan jumlah variabel terbatas atau diterapkan pada domain kendaraan yang heterogen sehingga *knowledge base* menjadi lebih kompleks dan berpotensi menimbulkan inkonsistensi inferensi. Berdasarkan kondisi tersebut masih diperlukan penelitian yang mengkaji strategi pembatasan domain kendaraan homogen untuk meningkatkan konsistensi *knowledge base* dan transparansi proses estimasi harga.

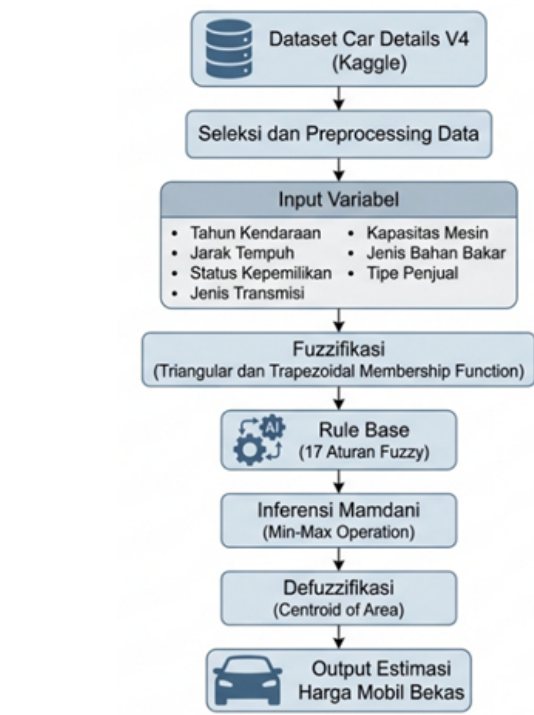
Oleh karena itu, penelitian ini mengusulkan sistem estimasi harga mobil bekas berbasis FIS Mamdani dengan tujuh variabel input dan pembatasan domain pada tiga merek dominan dalam dataset, yaitu Hyundai, Honda, dan Toyota. Pemilihan ketiga merek tersebut didasarkan pada volume data yang memadai, karakteristik harga yang relatif homogen, serta dominasi pada pasar kendaraan bekas kawasan Asia Selatan sehingga memungkinkan kalibrasi *knowledge base* yang lebih presisi.

Penelitian ini bertujuan merancang sistem estimasi harga mobil bekas berbasis FIS Mamdani dengan tujuh variabel input, membangun *knowledge base* berlapis yang konsisten, mengevaluasi performa sistem menggunakan *hit rate* toleransi, *confusion matrix*, dan *classification report*, serta menyediakan antarmuka interaktif untuk pengujian inferensi pada data nyata.

II. DESKRIPSI SISTEM

Sistem yang dikembangkan merupakan sistem estimasi harga mobil bekas berbasis *Fuzzy Inference System* (FIS) Mamdani. Metode ini dipilih karena mampu menangani ketidakpastian dalam penentuan harga kendaraan bekas yang dipengaruhi oleh banyak variabel serta dapat merepresentasikan pengetahuan pakar melalui aturan linguistik yang mudah diinterpretasikan [8], [9].

Penelitian difokuskan pada kendaraan merek Hyundai, Honda, dan Toyota yang memiliki karakteristik harga relatif serupa pada segmen kendaraan menengah. Sistem menerima tujuh parameter input yang terdiri atas tahun kendaraan, jarak tempuh, status kepemilikan, jenis transmisi, kapasitas mesin, jenis bahan bakar, dan tipe penjual. Ketujuh variabel dipilih berdasarkan penelitian terdahulu yang menunjukkan bahwa tahun kendaraan, jarak tempuh, kapasitas mesin, transmisi, bahan bakar, status kepemilikan, dan tipe penjual merupakan faktor yang paling berpengaruh terhadap harga kendaraan bekas



Gambar 1. Arsitektur Sistem Estimasi Harga Mobil Bekas Menggunakan FIS Mamdani

[2], [4], [11]. Variabel tersebut merepresentasikan aspek depresiasi kendaraan, spesifikasi teknis, preferensi pasar, dan tingkat kepercayaan konsumen. Variabel-variabel tersebut kemudian digunakan sebagai dasar dalam proses inferensi *fuzzy* untuk menentukan estimasi harga kendaraan bekas. Variabel input yang digunakan pada sistem estimasi harga mobil bekas ditunjukkan pada Tabel I.

TABLE I  
VARIABEL INPUT SISTEM

| Variabel           | Keterangan                           |
|--------------------|--------------------------------------|
| Tahun Kendaraan    | Tahun Produksi Kendaraan             |
| Jarak Tempuh       | Total Kilometer Penggunaan Kendaraan |
| Status Kepemilikan | Status Oemikik Kendaraan             |
| Jenis Transmisi    | Manual Atau Autimetic                |
| Kapasitas Mesin    | Kapasitas Mesin Kendaraan (Cc)       |
| Jenis Bahan Bakar  | Petrol, Diesel, Cng, Dan Lainnya     |
| Tipe Penjual       | Dealer Atau Individual               |

Pada tahap fuzzifikasi, setiap variabel direpresentasikan ke dalam beberapa himpunan *fuzzy* menggunakan fungsi keanggotaan berbentuk triangular dan trapezoidal [9]. Tahap ini bertujuan untuk mengubah data crisp menjadi nilai *fuzzy* sehingga dapat diproses dalam mekanisme inferensi *fuzzy*. Variabel tahun kendaraan direpresentasikan ke dalam kategori lama, sedang, dan baru, sedangkan variabel jarak tempuh dan kapasitas mesin dibagi ke dalam beberapa kategori linguistik sesuai karakteristik data kendaraan.

Secara matematis, fungsi keanggotaan triangular (trimf) yang digunakan dalam penelitian ini didefinisikan sebagai

berikut [9]:

$$\mu(x) = \max \left( \min \left( \frac{x-a}{b-a}, \frac{c-x}{c-b} \right), 0 \right) \quad (1)$$

dengan  $a$  merupakan batas bawah,  $b$  titik tengah, dan  $c$  batas atas dari fungsi keanggotaan *fuzzy*. Selain fungsi triangular, penelitian ini juga menggunakan fungsi keanggotaan trapezoidal (*trapmf*) untuk merepresentasikan himpunan *fuzzy* yang memiliki rentang nilai dominan pada interval tertentu. Secara matematis, fungsi *trapmf* dirumuskan sebagai berikut [9]:

$$\mu(x) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x \leq b, \\ 1, & b < x \leq c, \\ \frac{d-x}{d-c}, & c < x \leq d, \\ 0, & x > d. \end{cases} \quad (2)$$

Dengan  $a$ ,  $b$ ,  $c$ , dan  $d$  merupakan parameter batas fungsi keanggotaan trapesium. Fungsi ini digunakan pada himpunan *fuzzy* yang memiliki daerah keanggotaan penuh dalam suatu rentang nilai tertentu, seperti kategori Lama, Baru, Rendah, dan Tinggi pada variabel input sistem.

Nilai *fuzzy* hasil fuzzifikasi selanjutnya diproses menggunakan metode Mamdani berdasarkan 17 aturan *fuzzy* yang disusun dalam *knowledge base* berlapis [8], [9]. Struktur aturan dibangun menggunakan kombinasi variabel utama berupa tahun kendaraan dan kapasitas mesin sebagai matriks inti penentuan harga mobil bekas. Beberapa contoh aturan *fuzzy* yang digunakan pada *knowledge base* sistem ditunjukkan pada Tabel II.

TABLE II  
CONTOH RULE BASE FUZZY

| Rule | Aturan                                                                     |
|------|----------------------------------------------------------------------------|
| R1   | IF Tahun Baru AND Kilometer Rendah AND Mesin Besar THEN Harga Sangat Mahal |
| R2   | IF Tahun Lama AND Kilometer Tinggi THEN Harga Murah                        |
| R3   | IF Tahun Sedang AND Automatic AND First Owner THEN Harga Sedang            |
| R4   | IF Tahun Baru AND Dealer THEN Harga Mahal                                  |

Hasil inferensi *fuzzy* kemudian diproses menggunakan metode defuzzifikasi *centroid of area* untuk menghasilkan nilai estimasi harga kendaraan dalam bentuk crisp output [8]. Persamaan defuzzifikasi centroid dinyatakan sebagai berikut:

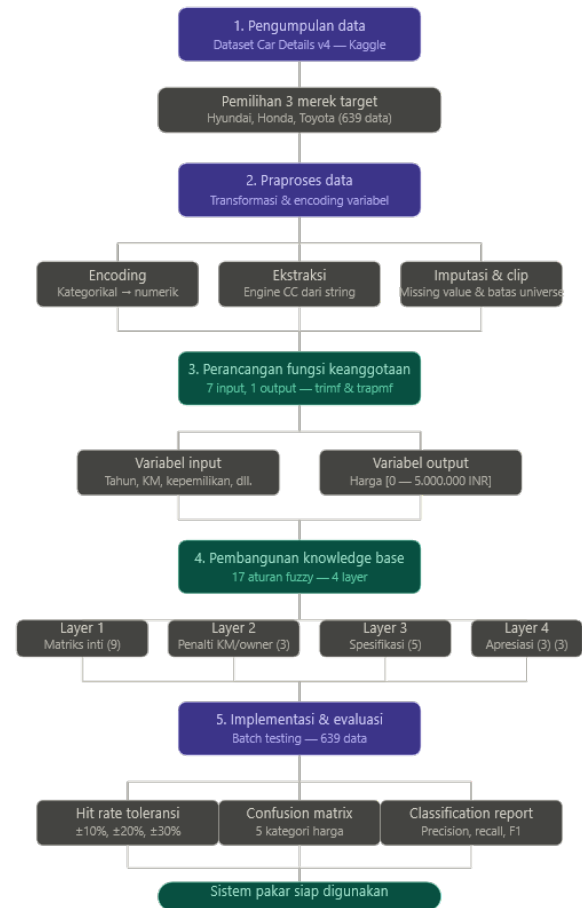
$$z^* = \frac{\int z \mu(z) dz}{\int \mu(z) dz} \quad (3)$$

dengan  $z^*$  merupakan nilai crisp output hasil defuzzifikasi dan  $\mu(z)$  merupakan derajat keanggotaan himpunan *fuzzy* output yang telah diagregasi [8]. Output sistem

berupa estimasi harga kendaraan bekas dalam satuan INR berdasarkan kombinasi karakteristik kendaraan yang diberikan sebagai input sistem.

### III. METODE

Pemilihan Hyundai, Honda, dan Toyota didasarkan pada volume data yang memadai dalam dataset, karakteristik harga yang relatif homogen pada segmen menengah, serta kebutuhan untuk membangun *knowledge base* yang lebih terkalibrasi dan konsisten.

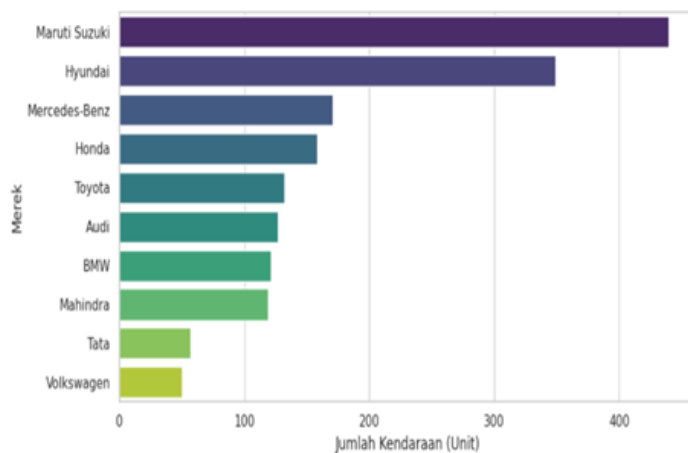


Gambar 2. Diagram Alur Metodologi Penelitian, Sumber : Hasil Analisis (2026)

#### A. Dataset dan Pemilihan Data

Dataset yang digunakan adalah *Car Details v4* yang bersumber dari platform Kaggle, memuat data kendaraan bekas yang beredar di pasar India. Dari keseluruhan data yang berjumlah 2059, dilakukan penyaringan terhadap tiga merek target: Hyundai, Honda, dan Toyota, sehingga diperoleh 639 baris data yang digunakan dalam seluruh proses pengembangan dan evaluasi sistem. Pemilihan ketiga merek ini dilandasi oleh tiga pertimbangan: (1) ketiganya merupakan merek dengan volume transaksi terbesar di pasar kendaraan bekas Asia Selatan dalam dataset; (2) ketiganya berada dalam segmen harga menengah sehingga karakteristik harga lebih homogen; serta (3)

pembatasan domain memungkinkan kalibrasi *knowledge base* yang lebih presisi.



Gambar 3. Top 10 Kendaraan Berdasarkan Dataset, Sumber : Hasil Analisis (2026)

Gambar 3 menampilkan distribusi 10 merek teratas berdasarkan jumlah data dalam dataset *Car Details v4*. Hyundai (349 unit) dan Honda (158 unit) menempati posisi kedua dan keempat, sementara Toyota (132 unit) berada di posisi kelima. Ketiga merek ini dipilih karena selain memiliki volume data yang memadai untuk evaluasi statistik, ketiganya merupakan merek segmen menengah yang secara konsisten mendominasi pasar kendaraan bekas di kawasan Asia Selatan. Maruti Suzuki yang mendominasi secara populasi tidak disertakan karena termasuk merek lokal India dengan karakteristik harga yang sangat spesifik dan tidak representatif untuk pasar yang lebih luas.

TABLE III  
STATISTIK DESKRIPTIF DATASET

| Atribut             | Nilai                                    |
|---------------------|------------------------------------------|
| Total Data          | 639 kendaraan                            |
| Merk                | Hyundai (349), Honda (158), Toyota (132) |
| Rentang Tahun       | 2002 - 2022                              |
| Rentang Harga       | 100.000 - 11.500.000 IDR                 |
| Rata-Rata Harga     | 906.122 IDR                              |
| Median Harga        | 660.000 IDR                              |
| Jarak Tempuh (Mean) | 58.029 km                                |
| Kapasitas Mesin     | 814 - 2.982 cc (rata-rata 1.572 cc)      |
| Bahan Bakar         | Petrol: 391, Diesel: 236, Lainnya: 12    |
| Transmisi           | Manual: 446, Otomatis: 193               |
| Kepemilikan         | First: 525, Second: 93, Third+: 21       |
| Tipe Penjual        | Individual: 622, Corporate: 17           |

Sumber: Hasil Analisis (2026)

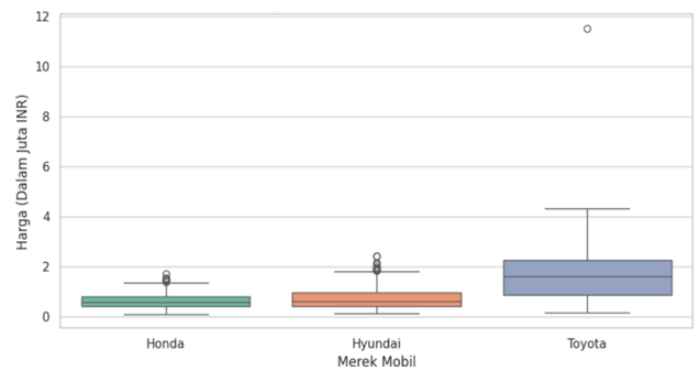
Distribusi data berdasarkan era produksi menunjukkan dominasi era sedang (2015–2019) dengan 342 kendaraan (53,5%), diikuti era lama (2010–2014) dengan 210 kendaraan (32,9%), dan era baru (2020–2026) dengan 87 kendaraan (13,6%). Distribusi ini mencerminkan kondisi pasar kendaraan bekas yang umumnya didominasi oleh kendaraan berusia lima hingga sepuluh tahun.

TABLE IV  
DISTRIBUSI KATEGORI HARGA PER MERK

| Merk    | Sangat Murah (<500RB) | Murah (500RB -1JT) | Normal (1-1,8JT) | Mahal (1,8-3JT) | Sangat Mahal (>3JT) | Total |
|---------|-----------------------|--------------------|------------------|-----------------|---------------------|-------|
| Honda   | 62                    | 81                 | 15               | 0               | 0                   | 158   |
| Hyundai | 129                   | 137                | 72               | 11              | 0                   | 349   |
| Toyota  | 17                    | 25                 | 34               | 40              | 16                  | 132   |
| Total   | 208                   | 243                | 121              | 51              | 16                  | 639   |

Sumber: Hasil Analisis (2026)

Untuk mempermudah proses evaluasi dan interpretasi hasil inferensi, harga kendaraan pada dataset dikelompokkan ke dalam lima kategori linguistik, yaitu Sangat Murah, Murah, Normal, Mahal, dan Sangat Mahal. Penentuan batas setiap kategori dilakukan berdasarkan analisis distribusi harga aktual kendaraan pada dataset, dengan mempertimbangkan rentang harga minimum, nilai tengah distribusi, kepadatan data pada setiap rentang harga, serta karakteristik pasar kendaraan bekas yang menjadi objek penelitian. Pendekatan ini digunakan agar fungsi keanggotaan output yang dibangun tidak bersifat subjektif, tetapi tetap merepresentasikan kondisi pasar secara proporsional dan sesuai dengan pola distribusi data yang diamati.



Gambar 4. Distribusi Harga dan Analisis *Outlier*, Sumber : Hasil Analisis (2026)

Gambar 4 memperlihatkan distribusi harga ketiga merek melalui diagram kotak (*box plot*). Honda dan Hyundai memiliki IQR yang sangat kompak di bawah 2 juta IDR dengan median masing-masing sekitar 577.500 dan 595.000 IDR. Toyota memiliki sebaran yang lebih lebar dengan median 1.589.500 IDR namun tetap dengan IQR yang sebagian besar berada di bawah 4 juta IDR. Keberadaan *outlier* di atas 5 juta IDR bersifat sangat terbatas (kurang dari 1% data), sehingga penetapan batas atas semesta output *fuzzy* pada nilai 5.000.000 IDR dapat dipertanggungjawabkan secara statistik tanpa mengorbankan representasi data yang signifikan.

## B. Praproses Data

Praproses data mencakup beberapa tahap transformasi. Pertama, variabel transmisi dipetakan ke representasi biner (Manual=0, Otomatis=1). Kedua, variabel

kepemilikan dipetakan ke skala ordinal (*First*=1, *Second*=2, *Third*=3, *Fourth*=4). Ketiga, kapasitas mesin diekstraksi dari format string "X cc" menggunakan operasi penggantian karakter. Keempat, variabel bahan bakar dikelompokkan menjadi biner (Diesel=1, selainnya=0), dan tipe penjual dikelompokkan menjadi biner (Dealer/Corporate=1, Individual=0). Seluruh nilai yang hilang diimputasi dengan nilai 0. Berbeda dari versi generik, variabel merek tidak digunakan sebagai input inferensi karena ketiga merek yang dianalisis berada dalam segmen yang sama sehingga pengaruh merek sudah terwakili dalam kalibrasi fungsi keanggotaan.

C. Perancangan Fungsi Keanggotaan

Sistem menggunakan tujuh variabel *input* dan satu variabel *output*. *Universe of discourse* dan himpunan *fuzzy* masing-masing variabel dirancang berdasarkan distribusi data aktual dan pertimbangan domain pasar kendaraan bekas India. Fungsi keanggotaan yang digunakan adalah kombinasi *trapmf* untuk himpunan yang memiliki batas tegas di satu sisi, dan *trimf* untuk himpunan dengan puncak tunggal.

TABLE V  
VARIABEL INPUT DAN FUNGSI KEANGGOTAAN

| Varia<br>bel | Uni<br>verse | Him<br>punan<br>Fuzzy | Fung<br>si | Parameter                            |
|--------------|--------------|-----------------------|------------|--------------------------------------|
| Tahun        | [2010, 2026] | Lama                  | trapmf     | [2010, 2010, 2014, 2017]             |
|              |              | Sedang Baru           | trimf      | [2015, 2018, 2022]                   |
| Kilometer    | [0, 200.000] | Rendah                | trapmf     | [2020, 2024, 2026, 2026]             |
|              |              |                       | trapmf     | [0, 0, 30.000, 60.000]               |
| Kepemilikan  | [1, 4]       | Sedang                | trimf      | [40.000, 80.000, 120.000]            |
|              |              | Tinggi                | trapmf     | [100.000, 150.000, 200.000, 200.000] |
| Transmisi    | [0, 1]       | First                 | trimf      | [1, 1, 2]                            |
|              |              | Second+               | trapmf     | [1, 2, 4, 4]                         |
| Mesin (Cc)   | [500, 4.000] | Manual                | trimf      | [0, 0, 1]                            |
|              |              | Otomatis              | trimf      | [0, 1, 1]                            |
| Bahan Bakar  | [0, 1]       | Kecil                 | trapmf     | [500, 500, 1.100, 1.400]             |
|              |              | Sedang                | trimf      | [1.300, 1.500, 2.000]                |
| Tipe Penjual | [0, 1]       | Besar                 | trapmf     | [1.800, 2.500, 4.000, 4.000]         |
|              |              | Bensin                | trimf      | [0, 0, 1]                            |
|              |              | Diesel                | trimf      | [0, 1, 1]                            |
|              |              | Individual            | trimf      | [0, 0, 1]                            |
|              |              | Dealer                | trimf      | [0, 1, 1]                            |
|              |              |                       | trimf      | [0, 1, 1]                            |

Sumber: Hasil Analisis (2026)

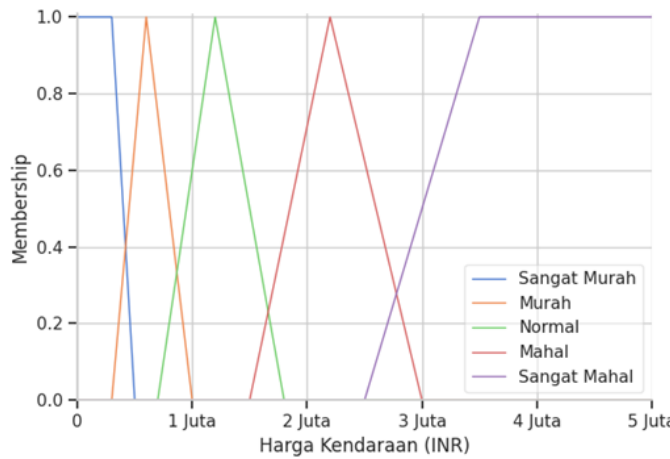
Variabel output harga dikalibrasi dengan universe [0, 5.000.000] INR, yang mencerminkan rentang harga realitis untuk tiga merek yang dikaji. Batas atas 5 juta INR dipilih karena 99,4% data berada di bawah nilai tersebut, sehingga informasi ekstrem yang berpotensi mendistorsi centroid defuzzifikasi dapat dieliminasi.

Gambar 5 mengilustrasikan rancangan kurva fungsi keanggotaan untuk variabel output harga kendaraan yang dibangun menggunakan kombinasi kurva trapesium (*trapmf*) dan segitiga (*trimf*). Berdasarkan parameter

TABLE VI  
FUNGSI KEANGGOTAAN OUTPUT HARGA

| Kategori Linguistik | Fungsi | Parameter (INR)                              |
|---------------------|--------|----------------------------------------------|
| Sangat Murah        | trapmf | [0, 0, 300.000, 500.000]                     |
| Murah               | trimf  | [300.000, 600.000, 1.000.000]                |
| Normal              | trimf  | [700.000, 1.200.000, 1.800.000]              |
| Mahal               | trimf  | [1.500.000, 2.200.000, 3.000.000]            |
| Sangat Mahal        | trapmf | [2.500.000, 3.500.000, 5.000.000, 5.000.000] |

Sumber: Hasil Analisis (2026)

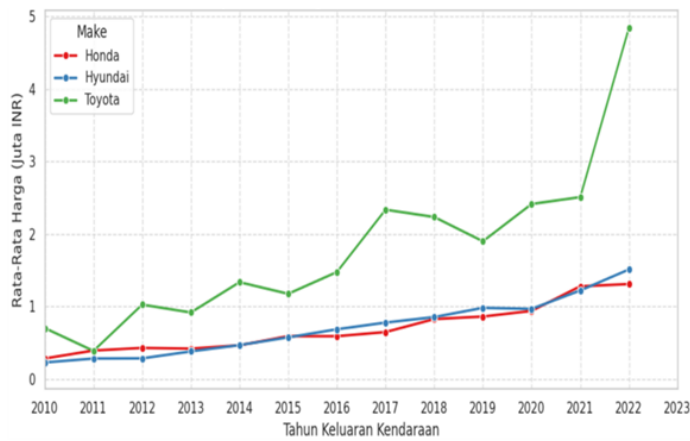


Gambar 5. Visualisasi Fungsi Keanggotaan Output Harga, Sumber: Hasil Analisis (2026)

pada Tabel VI, kalibrasi batas kurva pada setiap himpunan linguistik sengaja dirancang secara asimetris untuk merefleksikan kepadatan data aktual pasar. Resolusi kurva disusun lebih rapat pada rentang harga di bawah 2.000.000 INR untuk kategori 'Sangat Murah', 'Murah', dan 'Normal', guna mengakomodasi dan membedakan variasi harga secara sensitif pada segmen populasi terpadat. Sebaliknya, rentang kurva untuk kategori 'Mahal' dan 'Sangat Mahal' diregangkan lebih lebar guna menangkap varians harga pada frekuensi data yang lebih langka, dengan batas absolut semesta (*universe of discourse*) dikunci pada angka 5.000.000 INR. Selain itu, proporsi persimpangan (*overlap*) antar-kurva memvisualisasikan transisi mulus derajat keanggotaan ganda, yang menjadi karakteristik utama logika *fuzzy* dalam menangani ketidakpastian persepsi nilai pasar dibandingkan pendekatan logika klasik (hitam-putih).

D. Pembangunan Knowledge Base

*Knowledge base* dirancang dalam tiga lapisan aturan yang merepresentasikan faktor utama, faktor penalti depresiasi, dan faktor spesifikasi kendaraan. Dengan pembatasan pada tiga merek *mid-tier* yang homogen, variabel merek dieliminasi dari input inferensi sehingga jumlah aturan dapat disederhanakan secara signifikan dari 55 aturan (desain generik) menjadi 17 aturan tanpa kehilangan representasi pengetahuan yang esensial.



Gambar 6. Tren Depresiasi Harga, Sumber: Hasil Analisis (2026)

Gambar 6 menampilkan tren rata-rata harga berdasarkan tahun keluaran kendaraan untuk ketiga merek. Pola kenaikan harga yang konsisten seiring bertambah barunya tahun produksi mengonfirmasi validitas empiris dari matriks inti *knowledge base* (Layer 1), di mana variabel tahun dijadikan penentu utama kategori harga dasar. Toyota secara konsisten menunjukkan nilai yang lebih tinggi dibandingkan Honda dan Hyundai pada tahun yang sama, memperkuat keputusan untuk menggunakan tiga merek ini dalam domain yang sama meskipun tanpa variabel merek sebagai input inferensi. Perbedaan harga antar merek masih berada dalam rentang yang dapat diwakili oleh satu set fungsi keanggotaan yang dikalibrasi bersama. Pola depresiasi yang tampak pada kendaraan era lama (2010-2014) secara langsung menjadi landasan empiris bagi aturan penalti pada Layer 2 *knowledge base*.

#### E. Metode Evaluasi

Evaluasi sistem dilakukan secara batch terhadap seluruh 639 data menggunakan tiga metrik. Pertama, hit rate toleransi, yaitu persentase data yang nilai prediksinya berada dalam rentang deviasi  $\pm 10\%$ ,  $\pm 20\%$ , dan  $\pm 30\%$  dari harga aktual. Kedua, *confusion matrix* lima kategori untuk menganalisis pola klasifikasi harga berdasarkan kategori Sangat Murah, Murah, Normal, Mahal, dan Sangat Mahal yang ditetapkan sesuai fungsi keanggotaan output. Ketiga, *classification report* yang mencakup *precision*, *recall*, *F1-score*, dan akurasi keseluruhan sebagai ukuran performa klasifikasi setiap kategori.

#### F. Alat dan Bahan Penelitian

Dalam pelaksanaan penelitian ini, instrumen yang digunakan diklasifikasikan ke dalam dua kategori utama, yakni bahan penelitian berupa data sekunder dan alat bantu komputasi berupa perangkat lunak.

##### 1) Bahan Penelitian

Bahan utama yang digunakan dalam penelitian ini adalah dataset publik kendaraan bekas yang bersumber dari repositori platform Kaggle dengan

TABLE VII  
KNOWLEDGE BASE - 17 ATURAN FUZZY

| No. | Layer            | Aturan (IF - THEN)                                                                     | Justifikasi                           |
|-----|------------------|----------------------------------------------------------------------------------------|---------------------------------------|
| 1   | Inti             | IF tahun=Lama AND mesin=Kecil THEN harga=Sangat Murah                                  | Depresiasi tinggi + kapasitas rendah  |
| 2   | Inti             | IF tahun=Lama AND mesin=Sedang THEN harga=Murah                                        | Depresiasi tinggi, kapasitas cukup    |
| 3   | Inti             | IF tahun=Lama AND mesin=Besar THEN harga=Normal                                        | Nilai residu mesin besar              |
| 4   | Inti             | IF tahun=Sedang AND mesin=Kecil THEN harga=Murah                                       | Depresiasi moderat + kapasitas rendah |
| 5   | Inti             | IF tahun=Sedang AND mesin=Sedang THEN harga=Normal                                     | Kondisi pasar rata-rata               |
| 6   | Inti             | IF tahun=Sedang AND mesin=Besar THEN harga=Mahal                                       | Mesin besar masih diminati            |
| 7   | Inti             | IF tahun=Baru AND mesin=Kecil THEN harga=Normal                                        | Usia muda kompensasi kapasitas kecil  |
| 8   | Inti             | IF tahun=Baru AND mesin=Sedang THEN harga=Mahal                                        | Kombinasi optimal pasar               |
| 9   | Inti             | IF tahun=Baru AND mesin=Besar THEN harga=Sangat Mahal                                  | Nilai tertinggi segmen ini            |
| 10  | Penalti KM/Owner | IF tahun=Lama AND (km=Tinggi OR owner=Second+) THEN harga=Sangat Murah                 | Depresiasi berlipat                   |
| 11  | Penalti KM/Owner | IF tahun=Sedang AND (km=Tinggi OR owner=Second+) THEN harga=Murah                      | Koreksi turun 1 kasta                 |
| 12  | Penalti KM/Owner | IF tahun=Baru AND (km=Tinggi OR owner=Second+) THEN harga=Normal                       | Usia muda menahan depresiasi          |
| 13  | Spesifikasi      | IF transmisi=Manual AND bbm=Bensin THEN harga=Murah                                    | Kombinasi standar, nilai moderat      |
| 14  | Spesifikasi      | IF tahun=Sedang AND km=Rendah AND transmisi=Auto THEN harga=Mahal                      | KM rendah + auto menambah nilai       |
| 15  | Spesifikasi      | IF tahun=Baru AND km=Rendah AND transmisi=Auto THEN harga=Sangat Mahal                 | Kondisi prima, nilai tertinggi        |
| 16  | Spesifikasi      | IF bbm=Diesel AND mesin=Besar AND (tahun=Baru OR tahun=Sedang) THEN harga=Sangat Mahal | Diesel besar diminati pasar           |
| 17  | Spesifikasi      | IF penjual=Dealer AND tahun=Baru THEN harga=Mahal                                      | Garansi dealer menambah nilai         |

Sumber: Hasil Analisis (2026)

nama "Vehicle dataset from cardekho" [12]. Dataset memuat informasi spesifikasi dan harga kendaraan bekas. Pada penelitian ini dilakukan penyaringan data untuk tiga merek target, yaitu Hyundai, Honda, dan Toyota, serta pemilihan atribut yang relevan dengan proses inferensi fuzzy.

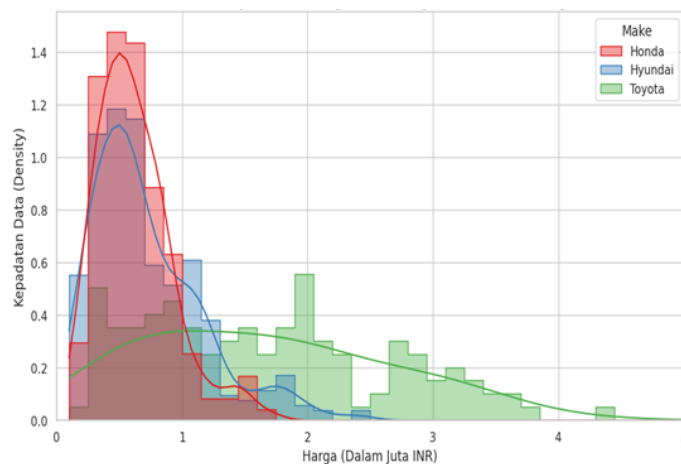
## 2) Alat Penelitian (Perangkat Lunak dan Lingkungan Komputasi)

Sistem dikembangkan menggunakan bahasa pemrograman Python pada lingkungan *Google Colaboratory*. Implementasi *Fuzzy Inference System* dilakukan menggunakan library *scikit-fuzzy*, sedangkan proses praproses data dan analisis menggunakan *pandas* dan *numpy*. Visualisasi data dibuat menggunakan *matplotlib* dan *seaborn*, sementara *ipywidgets* digunakan untuk membangun antarmuka pengujian inferensi secara interaktif.

## IV. HASIL DAN PEMBAHASAN

### A. Distribusi Data dan Karakteristik Segmen

Sebelum membahas hasil evaluasi sistem, penting untuk memahami karakteristik distribusi data yang digunakan. Sebagaimana ditampilkan pada Tabel IV, distribusi harga antar merek bersifat tidak merata. Honda dan Hyundai didominasi oleh kategori Sangat Murah dan Murah (masing-masing 91% dan 76% dari total data merek tersebut), sedangkan Toyota memiliki distribusi yang lebih merata dengan representasi signifikan pada kategori Mahal (40 unit) dan Sangat Mahal (16 unit). Ketidakseimbangan distribusi ini merupakan faktor utama yang perlu diperhatikan dalam interpretasi hasil evaluasi.



Gambar 7. Distribusi Kategori Harga per Merek, Sumber: Hasil Analisis (2026)

Gambar 7 memvisualisasikan distribusi frekuensi kendaraan berdasarkan lima kategori harga linguistik untuk merek Hyundai, Honda, dan Toyota. Berdasarkan grafik tersebut, terlihat adanya perbedaan karakteristik profil kelas harga antar-merek. Populasi kendaraan Hyundai dan Honda terpusat secara masif pada kategori 'Sangat Murah' dan 'Murah', yang mengindikasikan

tingginya volume kendaraan berusia tua di dalam dataset untuk kedua merek tersebut. Sebaliknya, Toyota menunjukkan sebaran distribusi yang lebih merata dan secara absolut mendominasi segmen 'Mahal' serta 'Sangat Mahal'. Pola distribusi kategorikal ini memberikan konfirmasi empiris bahwa Toyota memiliki retensi nilai jual kembali (*resale value*) yang lebih tangguh dibandingkan dua kompetitornya di kelas menengah. Selain itu, visualisasi ini memperjelas adanya ketidakseimbangan kelas (*class imbalance*) yang condong ke kategori harga rendah, yang secara langsung menjadi faktor krusial penyebab rendahnya nilai metrik *precision* pada kategori 'Sangat Mahal' akibat terbatasnya jumlah representasi data.

### B. Evaluasi Sistem Generik

Untuk mengevaluasi dampak pembatasan domain terhadap performa sistem, dilakukan perbandingan antara model generik yang mencakup 33 merek kendaraan dan model terbatas yang hanya mencakup Hyundai, Honda, dan Toyota. Kedua model menggunakan metode *Fuzzy Inference System* Mamdani dengan struktur inferensi yang sama, namun berbeda pada cakupan domain kendaraan yang dianalisis.

Hasil evaluasi menunjukkan bahwa model terbatas menghasilkan performa yang lebih baik dibandingkan model generik pada berbagai metrik evaluasi. Pembatasan domain memungkinkan penyusunan fungsi keanggotaan dan *knowledge base* yang lebih spesifik terhadap karakteristik harga kendaraan yang dianalisis sehingga mengurangi kompleksitas inferensi dan meningkatkan konsistensi aturan *fuzzy*.

TABLE VIII  
EVALUASI SISTEM GENERIK

| Matrik              | Nilai  |
|---------------------|--------|
| Hit Rate $\pm 10\%$ | 9,66%  |
| Hit Rate $\pm 20\%$ | 17,44% |
| Hit Rate $\pm 30\%$ | 26,23% |

Sumber: Hasil Analisis (2026)

Berdasarkan Tabel VIII, model generik yang mencakup 33 merek kendaraan menunjukkan performa estimasi yang relatif rendah. Nilai hit rate  $\pm 30\%$  hanya mencapai 26,23%, yang berarti sebagian besar prediksi masih memiliki deviasi lebih dari 30% terhadap harga aktual. Hasil ini mengindikasikan bahwa tingginya heterogenitas karakteristik harga antar-merek menyulitkan proses kalibrasi fungsi keanggotaan dan penyusunan *knowledge base* yang mampu merepresentasikan seluruh populasi kendaraan secara konsisten. Oleh karena itu, dilakukan pembatasan domain pada tiga merek dengan karakteristik harga yang lebih homogen untuk mengevaluasi pengaruh homogenitas data terhadap performa sistem *fuzzy* yang diusulkan.

### C. Hasil Hit Rate Toleransi

Pengujian batch terhadap 639 sampel menghasilkan nilai *hit rate* toleransi sebagaimana ditampilkan pada

Tabel IX. Nilai *hit rate*  $\pm 30\%$  sebesar 50,55% menunjukkan bahwa lebih dari separuh sampel dapat diprediksi dengan deviasi kurang dari 30% terhadap harga aktual. Jika dibandingkan dengan model generik yang mencakup 33 merek kendaraan dan hanya memperoleh *hit rate*  $\pm 30\%$  sebesar 26,23%, terjadi peningkatan sebesar 24,32 poin persentase atau sekitar 92,8%. Hasil tersebut menunjukkan bahwa pembatasan domain pada kendaraan dengan karakteristik harga yang lebih homogen mampu meningkatkan konsistensi inferensi *fuzzy* dan memperbaiki kemampuan sistem dalam merepresentasikan pola harga pasar.

Meskipun nilai *hit rate*  $\pm 30\%$  belum menunjukkan presisi numerik yang sangat tinggi, hasil tersebut perlu diinterpretasikan sesuai karakteristik pasar kendaraan bekas yang memiliki tingkat variasi harga relatif besar. Pemilihan toleransi  $\pm 30\%$  didasarkan pada karakteristik pasar kendaraan bekas yang memiliki variasi harga cukup tinggi akibat perbedaan kondisi fisik, riwayat penggunaan, lokasi penjualan, serta faktor negosiasi yang tidak seluruhnya dapat direpresentasikan oleh variabel dalam model. Oleh karena itu, evaluasi berbasis rentang toleransi digunakan untuk mengukur kemampuan sistem dalam menghasilkan estimasi yang masih berada pada kisaran harga pasar yang dapat diterima. Pendekatan serupa umum digunakan pada sistem estimasi harga yang memiliki ketidakpastian tinggi, di mana kedekatan nilai prediksi terhadap harga aktual dianggap lebih relevan dibandingkan kesesuaian nilai secara absolut.

TABLE IX  
HASIL UJI HIT RATE TOLERANSI

| Toleransi  | Jumlah Tepat | Total Sampel | HitRate (%) |
|------------|--------------|--------------|-------------|
| $\pm 10\%$ | 121          | 639          | 18,94%      |
| $\pm 20\%$ | 249          | 639          | 38,97%      |
| $\pm 30\%$ | 323          | 639          | 50,55%      |

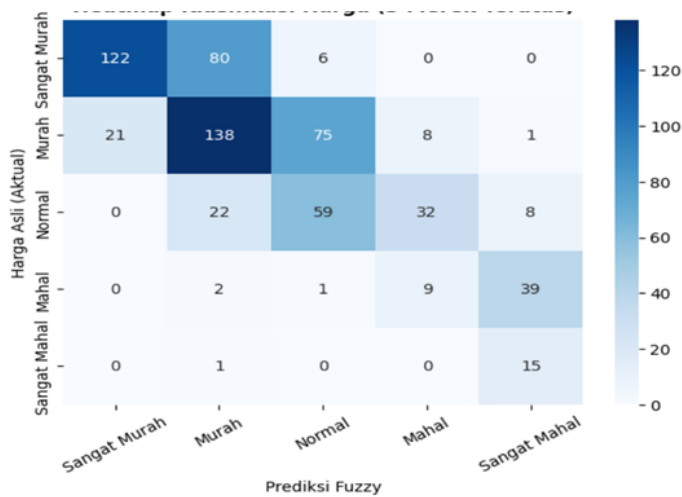
Sumber: Hasil Analisis (2026)

Nilai *hit rate*  $\pm 10\%$  sebesar 18,94% mencerminkan keterbatasan inheren metode *centroid* pada estimasi nilai numerik yang tepat. Metode *centroid* menghasilkan nilai tunggal berdasarkan pusat massa himpunan *fuzzy* teragregasi, yang secara alamiah bersifat *smooth* dan tidak dapat menangkap variasi harga individual secara presisi. Hal ini merupakan *trade-off* yang disadari dalam pemilihan metode Mamdani, yang mengutamakan interpretabilitas linguistik dibandingkan presisi numerik.

D. Analisis Confusion Matrix dan Classification Report

Selain evaluasi menggunakan *Hit Rate* toleransi, performa sistem juga dianalisis melalui *Confusion Matrix* dan *Classification Report* untuk mengetahui kemampuan sistem dalam mengklasifikasikan kendaraan ke dalam lima kategori harga linguistik. Evaluasi ini dilakukan menggunakan seluruh 639 sampel yang sama dengan pengujian *Hit Rate* sehingga dapat memberikan gambaran yang lebih

rinci mengenai pola kesalahan klasifikasi yang terjadi pada setiap kategori harga.



Gambar 8. Heatmap Confusion Matrix Klasifikasi Harga, Sumber: Hasil Analisis (2026)

Gambar 8 menyajikan *Confusion Matrix* yang memperlihatkan hubungan antara kategori harga aktual dan kategori harga hasil prediksi sistem. Secara umum terlihat bahwa sebagian besar prediksi terkonsentrasi pada diagonal utama matriks, yang menunjukkan bahwa sistem mampu mengidentifikasi kategori harga kendaraan dengan tingkat kesesuaian yang cukup baik. Konsentrasi nilai pada diagonal utama mengindikasikan bahwa mayoritas kendaraan berhasil diklasifikasikan ke dalam kategori harga yang sesuai dengan kondisi aktual pada dataset.

Kesalahan klasifikasi yang terjadi umumnya berada pada kategori yang saling berdekatan, seperti kendaraan dengan kategori aktual Murah yang diprediksi sebagai Normal atau sebaliknya. Tidak ditemukan pola kesalahan ekstrem, misalnya kendaraan kategori Sangat Murah yang diprediksi sebagai Sangat Mahal ataupun sebaliknya. Kondisi ini menunjukkan bahwa meskipun sistem belum mampu menghasilkan estimasi numerik yang sangat presisi, mekanisme inferensi *fuzzy* tetap mampu mempertahankan konsistensi logika penilaian harga sehingga hasil estimasi yang diberikan masih berada pada rentang kategori yang masuk akal secara pasar.

TABLE X  
CLASSIFICATION REPORT PER KATEGORI

| Kategori            | Precision | Recall | F1-Score | Support |
|---------------------|-----------|--------|----------|---------|
| Sangat Murah        | 0,85      | 0,59   | 0,70     | 208     |
| Murah               | 0,57      | 0,57   | 0,57     | 243     |
| Normal              | 0,42      | 0,49   | 0,45     | 121     |
| Mahal               | 0,18      | 0,18   | 0,18     | 51      |
| Sangat Mahal        | 0,24      | 0,94   | 0,38     | 16      |
| Akurasi Keseluruhan | -         | -      | 0,54     | 639     |
| Macro Avg           | 0,45      | 0,55   | 0,45     | 639     |
| Weighted Avg        | 0,59      | 0,54   | 0,55     | 639     |

Sumber: Hasil Analisis (2026)

Nilai *support* pada Tabel X menunjukkan jumlah sampel aktual pada masing-masing kategori harga yang digunakan dalam proses evaluasi. Kategori Sangat Murah memiliki 208 sampel, Murah 243 sampel, Normal 121 sampel, Mahal 51 sampel, dan Sangat Mahal 16 sampel. Distribusi tersebut menunjukkan adanya ketidakseimbangan kelas (*class imbalance*), di mana sebagian besar data terkonsentrasi pada kategori harga rendah, sedangkan kategori Mahal dan Sangat Mahal memiliki jumlah sampel yang relatif terbatas. Kondisi ini perlu diperhatikan dalam interpretasi nilai *precision*, *recall*, dan *F1-score* karena performa sistem pada kelas minoritas umumnya lebih sulit untuk dioptimalkan.

Berdasarkan Tabel X, kategori Sangat Murah memperoleh performa terbaik dengan *precision* sebesar 0,85 dan *F1-score* sebesar 0,70. Hasil ini menunjukkan bahwa sebagian besar kendaraan yang diprediksi sebagai Sangat Murah memang berasal dari kategori tersebut. Tingginya *precision* pada kategori ini mengindikasikan bahwa aturan penalti yang diterapkan pada knowledge base mampu mengidentifikasi kendaraan bernilai rendah secara konsisten dan andal.

Kategori Murah menunjukkan keseimbangan yang baik antara *precision* dan *recall* yang sama-sama bernilai 0,57 sehingga menghasilkan *F1-score* sebesar 0,57. Hasil ini menunjukkan bahwa sistem mampu mengenali kategori Murah secara relatif stabil tanpa kecenderungan yang kuat untuk melakukan *over-classification* maupun *under-classification*.

Pada kategori Normal diperoleh *precision* sebesar 0,42, *recall* sebesar 0,49, dan *F1-score* sebesar 0,45. Nilai tersebut menunjukkan bahwa kategori Normal berperan sebagai kelas transisi antara segmen harga rendah dan tinggi sehingga lebih rentan mengalami kesalahan klasifikasi. Kondisi ini wajar karena sebagian kendaraan pada kategori tersebut memiliki karakteristik yang masih tumpang tindih dengan kategori Murah maupun Mahal.

Kategori Mahal menghasilkan performa terendah dengan *precision* sebesar 0,18, *recall* sebesar 0,18, dan *F1-score* sebesar 0,18. Rendahnya performa pada kategori ini dipengaruhi oleh jumlah sampel yang relatif sedikit, yaitu hanya 51 data, sehingga pola harga yang terbentuk belum cukup kuat untuk direpresentasikan secara optimal oleh knowledge base yang digunakan. Selain itu, posisi kategori Mahal yang berada di antara kategori Normal dan Sangat Mahal menyebabkan peluang terjadinya misklasifikasi menjadi lebih besar.

Sebaliknya, kategori Sangat Mahal memiliki *recall* tertinggi sebesar 0,94 namun *precision* hanya sebesar 0,24. Hasil tersebut menunjukkan bahwa sistem mampu mendeteksi hampir seluruh kendaraan yang benar-benar termasuk kategori Sangat Mahal, tetapi pada saat yang sama masih terdapat cukup banyak kendaraan dari kategori lain yang ikut diprediksi sebagai Sangat Mahal. Kondisi ini menunjukkan bahwa sistem cenderung lebih sensitif dalam mengenali kendaraan bernilai tinggi dibandingkan menjaga ketepatan klasifikasinya.

Secara keseluruhan, *classification report* menghasilkan

akurasi kategorisasi sebesar 54% dengan *weighted F1-score* sebesar 0,55. Meskipun nilai tersebut belum setinggi pendekatan machine learning yang berorientasi pada optimasi prediksi numerik, hasil ini menunjukkan bahwa sistem *fuzzy* berbasis aturan masih mampu mempertahankan konsistensi klasifikasi pada domain kendaraan yang homogen. Temuan ini juga sejalan dengan hasil pengujian *hit rate* toleransi yang menunjukkan peningkatan signifikan dibandingkan model generik yang hanya memperoleh *hit rate*  $\pm 30\%$  sebesar 26,23%. Dengan demikian, pembatasan domain pada tiga merek kendaraan yang memiliki karakteristik harga relatif serupa terbukti memberikan kontribusi positif terhadap kualitas inferensi *fuzzy*, baik dari sisi akurasi kategorisasi maupun kemampuan sistem dalam menghasilkan estimasi yang tetap berada pada rentang harga yang rasional.

#### E. Pengujian Inferensi pada Contoh Kasus

Untuk memvalidasi kemampuan estimasi sistem secara individual, dilakukan pengujian inferensi pada kasus kenderaan Hyundai era sedang (2015–2019), jarak tempuh sedang (40.001–100.000 km), transmisi otomatis, mesin sedang (1.201–2.000 cc), BBM bensin, *first owner*, penjual individual. Sistem menghasilkan nilai centroid defuzzifikasi sebesar 1.233.333 INR dengan rentang harga pasar wajar  $\pm 10\%$  antara 1.110.000–1.356.667 INR

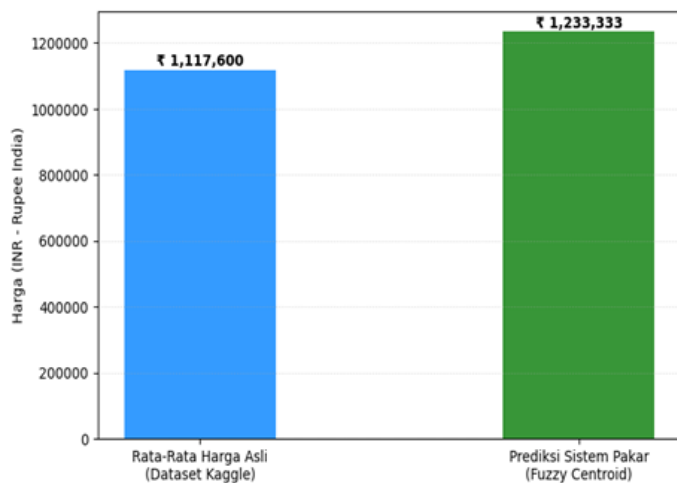
TABLE XI  
HASIL PENGUJIAN INFERENSI CONTOH KASUS

| Parameter                    | Nilai                                              |
|------------------------------|----------------------------------------------------|
| Merek                        | Hyundai                                            |
| Era Produksi                 | Sedang (2015–2019), nilai crisp: 2018              |
| Jarak Tempuh                 | Sedang (40.001–100.000 km), nilai crisp: 64.714 km |
| Transmisi                    | Otomatis                                           |
| Mesin                        | Sedang (1.201–2.000 cc), nilai crisp: 1.636 cc     |
| Bahan Bakar                  | Petrol/Lainnya                                     |
| Kepemilikan                  | First Owner                                        |
| Tipe Penjual                 | Individual                                         |
| Harga Centroid (Output)      | 1.233.333 INR                                      |
| Rentang Wajar ( $\pm 10\%$ ) | 1.110.000 – 1.356.667 INR                          |
| Jumlah Data Pembanding       | 10 unit (dari dataset)                             |
| Rata-Rata Harga Pasar Segmen | 1.117.600 INR                                      |
| Deviasi Prediksi Vs. Pasar   | 10,36 %                                            |

Sumber: Hasil Analisis (2026)

Deviasi sebesar 10,36% terhadap rata-rata harga pasar pada segmen yang sama menunjukkan bahwa sistem mampu menghasilkan estimasi yang mendekati kondisi pasar aktual pada kasus pengujian individual. Nilai tersebut masih berada dalam rentang toleransi  $\pm 10\%$ , sehingga menunjukkan bahwa *knowledge base* yang dibangun mampu merepresentasikan karakteristik kendaraan pada segmen yang diuji dengan cukup baik.

Gambar 9 memperlihatkan perbandingan antara harga hasil prediksi sistem dan rata-rata harga pasar pada segmen kendaraan yang sama. Sistem menghasilkan estimasi sebesar 1.233.333 INR, sedangkan rata-rata harga pasar



Gambar 9. Komparasi Harga Prediksi vs. Rata-rata Pasar Segmen (Ujicoba), Sumber: Hasil Analisis (2026)

tercatat sebesar 1.117.600 INR. Selisih sebesar 10,36% menunjukkan bahwa hasil inferensi masih berada pada rentang harga yang wajar dan sesuai dengan karakteristik pasar untuk segmen kendaraan tersebut. Meskipun contoh kasus ini tidak merepresentasikan keseluruhan performa sistem, hasil ini menunjukkan bahwa mekanisme inferensi *fuzzy* yang digunakan mampu menghasilkan estimasi yang konsisten pada kendaraan yang sesuai dengan domain *knowledge base* yang dibangun.

#### F. Diskusi: Implikasi Pembatasan Domain

Implementasi sistem pakar pada domain yang dibatasi pada tiga merek kendaraan, yaitu Hyundai, Honda, dan Toyota, menghasilkan akurasi kategorisasi sebesar 54% serta *hit rate* toleransi  $\pm 30\%$  sebesar 50,55%. Jika dibandingkan dengan model generik yang mencakup 33 merek kendaraan dan hanya memperoleh *hit rate*  $\pm 30\%$  sebesar 26,23%, terjadi peningkatan sebesar 24,32 poin persentase. Hasil ini menunjukkan bahwa pembatasan domain pada kendaraan dengan karakteristik harga yang lebih homogen mampu meningkatkan konsistensi inferensi *fuzzy* dan mempermudah proses kalibrasi fungsi keanggotaan serta *knowledge base* yang digunakan dalam sistem.

Peningkatan performa tersebut menunjukkan bahwa kualitas inferensi pada sistem berbasis aturan tidak hanya dipengaruhi oleh jumlah data, tetapi juga oleh kesesuaian antara *knowledge base* dan karakteristik domain yang direpresentasikan. Dengan menghilangkan variasi harga yang terlalu ekstrem antar merek, aturan *fuzzy* dapat merepresentasikan hubungan antarvariabel secara lebih spesifik sehingga menghasilkan estimasi yang lebih stabil dan mudah diinterpretasikan. Meskipun performa numerik sistem masih berada di bawah pendekatan machine learning yang berorientasi pada optimasi akurasi, metode FIS Mamdani tetap memberikan keunggulan dalam aspek transparansi karena setiap hasil estimasi dapat ditelusuri melalui aturan linguistik yang digunakan.

## V. KESIMPULAN

Penelitian ini berhasil merancang dan mengimplementasikan sistem estimasi harga mobil bekas menggunakan *Fuzzy Inference System* (FIS) Mamdani dengan tujuh variabel input dan 17 aturan fuzzy pada domain kendaraan yang dibatasi pada tiga merek, yaitu Hyundai, Honda, dan Toyota. Hasil evaluasi terhadap 639 data menunjukkan akurasi kategorisasi sebesar 54%, *hit rate* toleransi  $\pm 30\%$  sebesar 50,55%, *precision weighted* sebesar 0,59, *recall weighted* sebesar 0,54, dan *F1-score weighted* sebesar 0,55. Dibandingkan model generik yang mencakup 33 merek kendaraan dan hanya menghasilkan *hit rate*  $\pm 30\%$  sebesar 26,23%, model dengan domain terbatas menunjukkan peningkatan performa yang signifikan. Hasil tersebut menunjukkan bahwa pembatasan domain pada kendaraan dengan karakteristik harga yang lebih homogen mampu meningkatkan konsistensi inferensi *fuzzy* melalui fungsi keanggotaan dan *knowledge base* yang lebih spesifik serta mudah diinterpretasikan.

## DAFTAR PUSTAKA

- [1] N. Monburinon, P. Chertchom, T. Kaewkiriya, S. Rungpheung, S. Buaya, and P. Boonpou, "Prediction of Prices for Used Car by Using Regression Models," in *Proceedings of the 5th International Conference on Business and Industrial Research (ICBIR)*, Bangkok, Thailand, 2018, pp. 1–5, doi: <https://doi.org/10.1109/ICBIR.2018.8391177>.
- [2] A. T. Ahmed, S. Raja, and B. T. Ahmed, "Comparative Analysis of Car Price Prediction Using Machine Learning and Beyond: A Data-Driven Approach," *Journal of Image Processing and Artificial Intelligence*, vol. 11, no. 1, pp. 43–53, 2025.
- [3] J. D. Wu, C. C. Hsu, and H. C. Chen, "An Expert System of Price Forecasting for Used Cars Using Adaptive Neuro-Fuzzy Inference," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7809–7817, 2009, doi: <https://doi.org/10.1016/j.eswa.2008.11.019>.
- [4] C. Li, "Machine Learning-Based Models for Accurate Car Prices Prediction," *Highlights in Business, Economics and Management*, vol. 40, pp. 416–421, 2024, doi: <https://doi.org/10.54097/9zcpv779>.
- [5] R. Siva and A. M., "Linear Regression Algorithm Based Price Prediction of Car and Accuracy Comparison with Support Vector Machine Algorithm," *ECS Transactions*, vol. 107, pp. 12953–12964, 2022.
- [6] S. Pudaruth, "Predicting the Price of Used Cars Using Machine Learning Techniques," *International Journal of Information & Computation Technology*, vol. 4, no. 7, pp. 753–764, 2014.
- [7] D. Sundararaman, "How Much Is My Car Worth? A Methodology for Predicting Used Cars Prices Using Random Forest," *arXiv preprint arXiv:1711.06970*, 2017. [Online]. Available: <https://arxiv.org/abs/1711.06970>
- [8] S. Maryam, E. Bu'ulolo, and E. Hatmi, "Penerapan Metode Fuzzy Mamdani dan Fuzzy Tsukamoto dalam Menentukan Harga Mobil Bekas," *Journal of Informatics Electrical and Electronics Engineering*, vol. 1, no. 1, pp. 10–14, 2021.
- [9] S. Kusumadewi and H. Purnomo, *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*, 2nd ed. Yogyakarta, Indonesia: Graha Ilmu, 2010.
- [10] F. R. Andrari, Maimunah, and N. D. Qadarsih, "Penerapan Metode Fuzzy Mamdani dalam Menentukan Harga Jual Ponsel Pintar Bekas," *Pixel: Jurnal Ilmiah Komputer Grafis*, vol. 14, no. 2, 2021. [Online]. Available: <https://journal.stekom.ac.id/index.php/pixel/article/view/585>
- [11] A. Kumar, D. Singh, K. Chauhan, and M. S. Guru Prasad, "An Effective Pre-Owned Car Price Prediction Based on Multi Linear Regression Technique," in *Proceedings of the International Conference on Computer Science, Engineering and Technology (CSET)*, 2023, pp. 1–5, doi: <https://doi.org/10.1109/CSET58993.2023.10346830>.

- [12] N. Birla, *Vehicle Dataset from Cardekho*. Kaggle, 2020. [Online]. Available: <https://www.kaggle.com/datasets/nehalbirla/vehicle-dataset-from-cardekho/data>

#### BIOGRAFI PENULIS

**Rian Oktavianto Husain** merupakan mahasiswa peneliti di Program Studi Magister Teknologi Informasi, Universitas Gadjah Mada,

Yogyakarta, Indonesia. Lulusan terakhir dari Universitas Negeri Gorontalo, Indonesia pada bidang keahlian Data science dan jaringan komputer.

**Nursetia Wati** bekerja sebagai Dosen di Politeknik Gorontalo, Indonesia. Menyelesaikan studi S2 di Universitas Dian Nuswantoro dalam bidang Data mining. Minat penelitiannya mencakup teknik ekstraksi pola, prediksi, dan analisis data.