

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan kendaraan listrik di wilayah Eropa menunjukkan peningkatan yang signifikan dalam beberapa tahun terakhir seiring dengan kebijakan transisi energi dan target pengurangan emisi karbon yang diterapkan oleh Uni Eropa. Laporan *Global EV Outlook 2025* menunjukkan bahwa pangsa pasar kendaraan listrik di Eropa terus meningkat dalam total penjualan kendaraan baru [1].

Peningkatan adopsi kendaraan listrik baru mendorong pertumbuhan pasar kendaraan listrik bekas sebagai alternatif yang lebih terjangkau bagi konsumen [2]. Penentuan harga kendaraan listrik bekas memiliki tingkat kompleksitas yang lebih tinggi dibandingkan kendaraan konvensional karena dipengaruhi oleh karakteristik teknis kendaraan dan faktor usia kendaraan [3]. Faktor-faktor teknis seperti kapasitas baterai, efisiensi energi, serta performa kendaraan berkontribusi secara langsung terhadap nilai jual kendaraan listrik di pasar sekunder.

Kompleksitas dalam penentuan harga kendaraan listrik bekas menuntut pendekatan analisis berbasis data yang mampu memodelkan hubungan antar variabel secara kuantitatif. Pendekatan machine learning banyak digunakan dalam permasalahan prediksi harga karena kemampuannya dalam memanfaatkan data historis untuk menghasilkan estimasi yang objektif [4]. Salah satu metode yang sering digunakan dalam prediksi harga adalah regresi linear karena kesederhanaan model dan kemudahan interpretasi hasilnya [5].

Regresi linear banyak digunakan dalam prediksi harga kendaraan karena mampu menjelaskan pengaruh masing-masing variabel independen terhadap harga kendaraan secara langsung. Penelitian mengenai prediksi harga mobil bekas menunjukkan bahwa regresi linear masih relevan digunakan sebagai pendekatan prediktif yang efektif dan interpretable [6]. Dalam kasus kendaraan listrik bekas, harga kendaraan dipengaruhi oleh lebih dari satu variabel, sehingga pendekatan *Multiple Linear Regression* (MLR) menjadi sesuai untuk digunakan.

Multiple Linear Regression memungkinkan pemodelan hubungan linier antara harga kendaraan sebagai variabel dependen dengan berbagai spesifikasi teknis kendaraan sebagai variabel independen secara simultan. Pendekatan ini telah

diterapkan dalam berbagai penelitian prediksi harga kendaraan berbasis machine learning [7]. Penggunaan MLR juga memberikan transparansi dalam interpretasi model melalui koefisien regresi yang dihasilkan.

Estimasi parameter dalam model MLR umumnya dilakukan menggunakan metode *Ordinary Least Squares* (OLS). Metode OLS bertujuan untuk meminimalkan jumlah kuadrat selisih antara nilai aktual dan nilai prediksi sehingga menghasilkan estimasi parameter yang optimal [8]. Pendekatan ini banyak digunakan dalam regresi linear karena kestabilan estimasi dan kemudahan implementasinya.

Dalam pembangunan model regresi linear, pemilihan fitur merupakan tahapan penting untuk menghindari permasalahan multikolinearitas antar variabel independen. Multikolinearitas dapat menyebabkan ketidakstabilan nilai koefisien regresi dan menurunkan interpretabilitas model [9]. Salah satu metode yang digunakan untuk mendeteksi dan mengendalikan multikolinearitas adalah *Variance Inflation Factor* (VIF) [10].

Evaluasi kinerja model regresi diperlukan untuk mengukur tingkat akurasi prediksi dan kemampuan generalisasi model. Metrik *Mean Absolute Error* (MAE) digunakan untuk mengukur rata-rata kesalahan absolut antara nilai aktual dan nilai prediksi [11]. *Root Mean Squared Error* (RMSE) digunakan untuk memberikan penalti yang lebih besar terhadap kesalahan prediksi bernilai ekstrem [12]. Koefisien determinasi (R^2) digunakan untuk mengukur proporsi variasi harga kendaraan yang dapat dijelaskan oleh model regresi [13].

Selain metrik error tersebut, analisis *Confidence Interval* (CI) untuk prediksi digunakan untuk menilai seberapa pasti atau stabil prediksi model terhadap data baru. CI memberikan rentang nilai di mana harga aktual diperkirakan berada dengan tingkat keyakinan tertentu, misalnya 95%, sehingga pengguna model dapat memahami tingkat ketidakpastian dalam prediksi [14]. Dalam konteks prediksi harga kendaraan listrik, CI membantu mengevaluasi apakah model mampu menghasilkan estimasi yang konsisten dan realistis dibandingkan harga aktual, terutama ketika menghadapi data yang belum pernah dilihat sebelumnya. Penggunaan CI juga mendukung pengambilan keputusan berbasis risiko, karena memberikan indikasi probabilistik mengenai rentang nilai harga yang mungkin terjadi.

Dataset yang digunakan dalam penelitian ini bersumber dari pasar kendaraan listrik Eropa dengan harga dalam mata uang Euro [15]. Penelitian ini berfokus pada pembangunan model prediksi harga kendaraan listrik bekas melalui

tahapan pra-pemrosesan data, pemilihan fitur, pembangunan model Multiple Linear Regression menggunakan pendekatan OLS, serta evaluasi kinerja model menggunakan metrik evaluasi regresi yang telah ditetapkan.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk membangun dan mengevaluasi model prediksi harga kendaraan listrik bekas di pasar Eropa menggunakan metode Multiple Linear Regression berbasis machine learning. Model yang dibangun diharapkan mampu memberikan estimasi harga yang transparan, terukur, dan dapat dipertanggungjawabkan secara ilmiah.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimana membangun model regresi linear untuk memprediksi harga kendaraan listrik berdasarkan spesifikasi teknis kendaraan?
2. Bagaimana hasil pemodelan regresi linear dalam merepresentasikan hubungan antara spesifikasi teknis kendaraan dan harga kendaraan listrik?

1.3 Batasan Permasalahan

Agar penelitian ini lebih terarah dan sesuai dengan tujuan yang ingin dicapai, maka batasan masalah yang diterapkan adalah sebagai berikut:

1. Dataset yang digunakan merupakan data kendaraan listrik yang beredar di kawasan Eropa dengan satuan harga dalam mata uang Euro.
2. Penelitian ini hanya membahas prediksi harga mobil listrik bekas di pasar Eropa dan tidak membandingkan dengan harga kendaraan di negara atau wilayah lain.
3. Metode machine learning yang digunakan dalam penelitian ini dibatasi pada regresi linear.
4. Variabel yang digunakan dalam pemodelan meliputi spesifikasi teknis kendaraan, seperti jarak tempuh maksimum, efisiensi energi, kecepatan pengisian daya, performa kendaraan, serta usia kendaraan.

5. Faktor eksternal seperti kebijakan pajak, subsidi pemerintah, inflasi, kurs mata uang, dan kondisi ekonomi makro tidak dimasukkan ke dalam model.
6. Hasil prediksi yang dihasilkan bersifat estimasi berdasarkan data historis dan tidak dimaksudkan sebagai penentuan harga pasar secara mutlak.

1.4 Tujuan Penelitian

Beberapa tujuan utama yang ingin dicapai dari penelitian ini adalah sebagai berikut:

1. Membangun model regresi linear untuk memprediksi harga kendaraan listrik berdasarkan spesifikasi teknis kendaraan.
2. Menganalisis hubungan antara spesifikasi teknis kendaraan dan harga kendaraan listrik berdasarkan hasil pemodelan regresi linear.

1.5 Manfaat Penelitian

Manfaat utama yang ingin didapatkan dari penelitian ini adalah sebagai berikut:

1. Memberikan pemahaman mengenai penerapan metode regresi linear dalam membangun model prediksi harga kendaraan listrik berdasarkan spesifikasi teknis kendaraan.
2. Memberikan gambaran mengenai hubungan antara spesifikasi teknis kendaraan dan harga kendaraan listrik berdasarkan hasil pemodelan regresi linear.
3. Menjadi referensi awal bagi penelitian selanjutnya yang berkaitan dengan prediksi harga kendaraan listrik menggunakan pendekatan *machine learning*.

1.6 Sistematika Penulisan

Berisikan uraian singkat mengenai struktur isi penulisan laporan penelitian, dimulai dari Pendahuluan hingga Simpulan dan Saran.

Sistematika penulisan laporan adalah sebagai berikut:

- **Bab 1 PENDAHULUAN**

berisi penjelasan mengenai latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian, serta manfaat yang diharapkan. Bab ini memberikan gambaran umum mengenai alasan penelitian dilakukan dan arah penelitian yang akan dicapai.

- **Bab 2 LANDASAN TEORI**

membahas kajian pustaka dan landasan teori yang menjadi dasar penelitian. Bab ini mencakup teori terkait machine learning, konsep regresi linear, penelitian-penelitian sebelumnya, serta penjelasan mengenai variabel dan konsep yang digunakan dalam pemodelan prediksi harga mobil bekas.

- **Bab 3 METODOLOGI PENELITIAN**

menjelaskan metodologi penelitian yang digunakan dalam penyusunan model. Isi bab mencakup alur penelitian, teknik pengolahan dan pembersihan data, pemilihan fitur, metode analisis, serta alasan pemilihan algoritma Linear Regression sebagai model utama.

- **Bab 4 HASIL DAN DISKUSI**

berisi proses implementasi dan analisis data. Bab ini mencakup tahapan data preprocessing, eksplorasi data, visualisasi fitur, pembangunan model Linear Regression, serta evaluasi awal terhadap model yang telah dibuat.

- **Bab 5 SIMPULAN DAN SARAN**

berisi kesimpulan dari penelitian yang telah dilakukan serta saran yang dapat diberikan untuk penelitian selanjutnya. Bab ini merangkum temuan penting dan potensi pengembangan lebih lanjut dari model prediksi harga mobil bekas.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A