

BAB I

PENDAHULUAN

Bab ini menyajikan mengenai latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan pada skripsi yang berjudul prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*.

1.1. Latar Belakang

Salah satu permasalahan yang dihadapi manusia adalah kebutuhan hidupnya yaitu kebutuhan akan mobil. Mobil yang umumnya digunakan sebagai alat transportasi untuk mengangkut penumpang, barang dan perjalanan baik jarak jauh maupun jarak dekat. Sejalan dengan kemajuan teknologi, menyebabkan semakin banyak pula jenis kendaraan roda empat yang ditawarkan di pasaran untuk memenuhi kebutuhan masyarakat (Puteri & Safitri, 2020). Pasar mobil merupakan industri yang signifikan sebagai hasil dari perluasan luar biasa dalam produksi kendaraan dalam beberapa tahun terakhir (Dhabe dkk., 2023). Beberapa orang tidak mampu membeli mobil baru karena kekurangan dana (D'Costa dkk., 2020). Sebagian besar orang lebih suka membeli mobil lama karena harganya terjangkau dan dapat dijual kembali setelah beberapa tahun digunakan sehingga dapat mendatangkan keuntungan (Samruddhi & Kumar, 2020).

Tingginya minat masyarakat terhadap mobil lama telah menyebabkan bisnis ini berkembang pesat, yang terlihat dari peningkatan jumlah *showroom* mobil lama. Banyaknya pelaku usaha di bidang ini menciptakan persaingan ketat di antara penjual untuk tetap bertahan dalam bisnis mobil. Salah satu tantangan utama yang dihadapi semua *showroom* adalah kemampuan menetapkan harga dengan cepat dan akurat, agar mereka bisa menjual stok mobil dan segera menghasilkan pendapatan. Saat ini, prediksi harga mobil masih menjadi kebutuhan yang populer (Pandey dkk., 2020). Berbagai *showroom* bersaing dalam hal harga untuk menarik pelanggan.

Harga mobil lama dapat dipengaruhi oleh spesifikasi mobil itu sendiri, seperti tahun produksi, jarak tempuh, jenis transmisi, dan sebagainya. Faktor-faktor ini dapat

menyebabkan harga jual mobil lama menjadi tidak stabil atau bervariasi. Untuk memperkirakan harga mobil lama secara akurat, diperlukan metode yang mampu menganalisis dan memodelkan hubungan antara variabel-variabel tersebut (Dewi dkk., 2020).

Salah satu metode yang dapat digunakan untuk memprediksi harga adalah *Machine Learning*. *Machine Learning* merupakan cabang dari *Artificial Intelligence* (AI) yang mampu membuat keputusan berdasarkan data (Kaur & Kumari, 2022). *Machine Learning* dapat mengoptimalkan proses pengambilan keputusan dengan menyediakan wawasan prediktif yang berjalan secara otomatis. *Machine Learning* merupakan studi yang terus berkembang mengenai konsep pengenalan pola dan pembelajaran komputasi dalam kecerdasan buatan. Metode ini menggunakan algoritma pembelajaran untuk memprediksi dan mendukung pengambilan keputusan secara otomatis berdasarkan kumpulan data (Abellera & Bulusu, 2017).

Penelitian mengenai regresi prediksi harga mobil telah dilakukan dengan berbagai metode seperti *Linear Regression* dan *K-Nearest Neighbors Regression*. Dalam studi oleh Noor & Jan (2017) menggunakan metode *Linear Regression* dengan *dataset* dari *www.pakwheels.com* yang mencakup atribut seperti model kendaraan, kota, warna, dan *mileage*, diperoleh *R-Squared* sebesar 98,61%. Penelitian lain oleh D'Costa dkk. (2020) menggunakan *Multiple Linear Regression* pada *dataset* dari Mandovi Motors Pvt Ltd dengan atribut seperti produsen, model, tipe bahan bakar, dan tahun registrasi, menghasilkan *R-Squared* sebesar 98,34%.

Samruddhi & Kumar (2020) menggunakan *K-Nearest Neighbors Regression* dengan *dataset* dari *www.kaggle.com* di India yang mencakup atribut seperti lokasi, mesin, dan tahun, mencapai *R-Squared* sebesar 85%. Di Indonesia, Puteri & Safitri (2020) menerapkan *Linear Regression* pada *dataset* dari *www.mobil123.com* dengan atribut seperti usia kendaraan dan jenis transmisi, dan memperoleh *R-Squared* sebesar 76,2%. Dhabe dkk. (2023) juga menggunakan *Linear Regression* dengan *dataset* dari *www.kaggle.com* di India, menghasilkan *R-Squared* sebesar 76% dengan atribut seperti nama perusahaan, model, tahun pembelian, dan tipe bahan bakar.

Penelitian menggunakan metode *Ensemble Voting Regressor* belum ditemukan pada penelitian prediksi harga mobil namun ditemukan pada penelitian prediksi limbah

medis oleh Erdebilli & Devrim-Içtenbas (2022) menggunakan model *Ensemble Voting Regressor* (EVR) berisi gabungan 3 model regresi yaitu *Random Forest*, *Adaptive Boosting*, dan *Gradient Boosting Machine* dimana EVR memiliki performa lebih baik dari 3 model regresi yang digabungkan. Penelitian dengan judul *Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey* dilakukan berdasarkan *dataset* dari *Open Data Portal Istanbul Metropolitan Municipality Department* dan *Population Information Table by the Turkish Statistics Institute for the years 2004–2020* mendapatkan *R-Squared* sebesar 97%.

Berdasarkan penelitian-penelitian tersebut, maka pada penelitian ini melakukan prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*. Bagian utama dari penelitian ini adalah penerapan dan evaluasi metode *Ensemble Voting Regressor* dalam memprediksi harga mobil lama. Luaran dari penelitian ini berupa model prediksi harga mobil lama yang akurat dan dapat diandalkan. Penelitian ini dapat membantu penjual dan pembeli mobil lama, *dealer* mobil, serta pihak lain yang berkepentingan dalam industri otomotif untuk mendapatkan estimasi harga yang lebih akurat. Penelitian ini berfokus dalam mengembangkan dan menguji metode *Ensemble Voting Regressor*, serta membandingkan performanya dengan metode regresi individu yang telah digabungkan yaitu *K-Nearest Neighbors Regression* dan *Linear Regression*.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah yang dibahas dalam penelitian ini adalah:

1. Bagaimana membangun model prediksi harga mobil lama yang akurat menggunakan metode *Ensemble Voting Regressor* berisi gabungan model regresi individu seperti *K-Nearest Neighbors Regression* dan *Linear Regression*?
2. Bagaimana performa model *Ensemble Voting Regressor* dibandingkan dengan model regresi individu seperti *K-Nearest Neighbors Regression* dan *Linear Regression*?

1.3. Tujuan dan Manfaat

Tujuan dari penelitian ini adalah:

1. Menghasilkan model prediksi harga mobil lama menggunakan *Ensemble Voting Regressor* berisi gabungan model regresi individu seperti *K-Nearest Neighbors Regression* dan *Linear Regression*
2. Membandingkan performa model *Ensemble Voting Regressor* dengan model regresi individu seperti *K-Nearest Neighbors Regression* dan *Linear Regression*.

Manfaat dari penelitian ini adalah digunakan untuk membantu penjual dan pembeli mobil lama dalam menentukan harga jual dan harga beli yang wajar dengan memberikan alternatif metode prediksi yang lebih akurat dibandingkan metode konvensional.

1.4. Ruang Lingkup

Penelitian ini memerlukan batasan-batasan agar penelitian lebih terarah. Adapun batasan-batasan tersebut sebagai berikut:

1. *Dataset* yang dipakai merupakan *dataset* yang diambil dari *website kaggle* berisi 609 data mobil lama dan memiliki 20 atribut untuk setiap datanya.
2. Pelatihan dan pengujian model *Regression* dilakukan di *Jupyter Notebook* pada platform *Google Colaboratory* dengan *Library* yang digunakan dalam pelatihan adalah *Pandas* untuk membaca serta merapikan *dataset*, dan *Scikit-learn* untuk *Principal Component Analysis*, pembagian rasio *dataset*, pelatihan model regresi, *Grid Search*, metrik evaluasi.

1.5. Sistematika Penulisan

Sebagai gambaran yang runtut dan jelas tentang pembahasan dari penyusunan laporan dari penelitian ini, sistematika penulisan dibuat sebagai berikut:

BAB I PENDAHULUAN

Bab ini menyajikan mengenai latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan pada skripsi yang berjudul prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*.

BAB II TINJAUAN PUSTAKA

Bab ini menyajikan hasil studi literatur mengenai teori yang berhubungan dengan topik atau masalah yang dibahas pada penelitian yang berjudul prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*. Pada bab ini terdiri dari beberapa subbab antara lain *literature review*, pasaran mobil lama, *machine learning*, prediksi harga, data *preprocessing*, pembagian *dataset*, model regresi, metrik evaluasi, *grid search*, serta *tools* dan *library*.

BAB III METODOLOGI PENELITIAN

Bab ini menyajikan garis besar penyelesaian masalah, pengumpulan data, data *preprocessing*, pembagian *dataset*, pelatihan prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*, evaluasi prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*, serta *grid search*.

BAB IV HASIL DAN PEMBAHASAN

Bab ini membahas tentang lingkungan dan perangkat yang digunakan, data evaluasi, skenario evaluasi prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*, hasil dan analisa evaluasi prediksi harga mobil lama menggunakan *Ensemble Voting Regressor*, dan pembahasan hasil penelitian.

BAB V PENUTUP

Bab ini menyajikan tentang kesimpulan dari penjabaran bab-bab yang telah dibahas sebelumnya dan saran untuk pengembangan atau penelitian selanjutnya.