

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi *smartphone* mendorong tren pasar yang dinamis, di mana model-model baru terus bermunculan dengan fitur lebih canggih. Hal ini mendorong konsumen untuk sering mengganti perangkat, menghasilkan volume besar ponsel bekas yang tersedia di pasaran. Fenomena ini menghadirkan peluang bagi konsumen yang mencari alternatif ekonomis dengan membeli ponsel bekas yang berkualitas baik namun dengan harga lebih terjangkau dibandingkan ponsel baru. Para penjual juga dapat memperpanjang siklus hidup produk dan mengurangi dampak lingkungan dengan memanfaatkan pasar ponsel bekas (Hanif dan Best, 2024). Berbagai platform *e-commerce* kini berperan besar dalam mengakomodasi perdagangan perangkat bekas, menciptakan permintaan yang signifikan akan penilaian harga yang akurat dan adil guna memfasilitasi transaksi yang efisien.

Penelitian yang berfokus pada data ponsel bekas tidak hanya memberikan manfaat ekonomis tetapi juga mendorong aspek keberlanjutan. Analisis data harga ponsel bekas membantu menstabilkan harga di pasar, mengurangi praktik harga yang tidak realistis, dan menciptakan transparansi yang lebih baik dalam transaksi. Pemahaman mengenai nilai dan depresiasi harga perangkat dari sisi konsumen membantu dalam membuat keputusan pembelian yang lebih terinformasi, sehingga dapat meningkatkan kepuasan pengguna. Platform *e-commerce* dapat menggunakan model prediktif untuk memperkirakan harga yang sesuai berdasarkan faktor-faktor seperti usia, merek, dan kondisi fisik ponsel, yang pada akhirnya memberikan keuntungan kompetitif di pasar (Gong *et al.*, 2024).

Algoritma *Random Forest*, khususnya *Random Forest Regression*, sangat cocok untuk mengestimasi harga ponsel bekas karena kemampuannya dalam menangani data dengan banyak variabel kompleks. *Random Forest Regression* bekerja dengan menggabungkan hasil dari beberapa pohon keputusan sehingga dapat menghasilkan prediksi yang stabil dan akurat. Algoritma ini juga memiliki kemampuan menangani data yang tidak terstruktur dengan baik yang umumnya terdapat dalam data *e-commerce* ponsel bekas (Wiguna *et al.*, 2025).

Keunggulan lain dari *Random Forest Regression* adalah kemampuannya dalam melakukan *feature selection*, yaitu memilih variabel-variabel yang paling relevan dan signifikan untuk prediksi. Variabel seperti merek, kapasitas penyimpanan, usia perangkat, dan kondisi fisik berpotensi besar memengaruhi harga akhir ponsel bekas. *Random Forest Regression* dapat mengevaluasi pentingnya setiap variabel-variabel sehingga memungkinkan untuk memahami faktor mana yang paling berpengaruh terhadap harga. Keunggulan ini memberikan fleksibilitas tinggi dalam model prediksi dan meningkatkan *interpretabilitas* hasil, membuatnya sangat berguna bagi pengguna dan pengelola platform (Yin *et al.*, 2024).

Beberapa penelitian sebelumnya telah menunjukkan keberhasilan *Random Forest Regression* dalam berbagai kasus prediksi harga dan penilaian produk bekas di sektor *e-commerce*. *Random Forest Regression* efektif digunakan dalam menganalisis data *e-commerce* dengan hasil yang lebih baik dibandingkan metode prediksi lainnya. Hal ini menjadikan *Random Forest Regression* pilihan ideal untuk aplikasi prediksi harga di *e-commerce*, termasuk untuk produk bekas seperti ponsel (Rokhimakhumullah *et al.*, 2024). Beberapa studi sebelumnya telah menunjukkan

keunggulan *Random Forest Regression* dalam prediksi harga barang bekas dan penilaian produk di berbagai sektor. Penelitian oleh Samosir dan Hery (2024), *Random Forest Regression* digunakan untuk menganalisis faktor-faktor yang memengaruhi harga laptop bekas di platform *e-commerce*, yang menghasilkan prediksi harga dengan akurasi tinggi dibandingkan model regresi linier tradisional. Penelitian ini menggarisbawahi fleksibilitas *Random Forest Regression* dalam mengolah data dari berbagai sumber *e-commerce* dan membangun model yang dapat diandalkan (Samosir dan Hery, 2024). Penelitian lain seperti Lubis *et al.* (2019) menunjukkan bahwa *Random Forest Regression* efektif dalam peramalan harga produk dengan menggunakan data *e-commerce* sebagai sumber data utama. Lubis *et al.* (2019) juga menunjukkan bahwa *Random Forest Regression* mampu memberikan prediksi harga yang lebih stabil dibandingkan algoritma lain seperti *Support Vector Regression*, khususnya dalam kondisi di mana data memiliki variabilitas tinggi.

Penelitian ini bertujuan untuk membangun model prediktif berbasis *Random Forest Regression* yang mampu memperkirakan harga ponsel bekas secara efektif di platform *e-commerce*. Data yang digunakan mencakup informasi yang relevan dengan harga, seperti kondisi fisik, kapasitas penyimpanan, dan sebagainya. Penelitian ini diharapkan dapat membantu para pembeli dan penjual menentukan harga yang lebih adil dan akurat dengan membangun model prediktif yang *robust*. Hasil dari model ini tidak hanya relevan bagi pengguna akhir, tetapi juga bermanfaat bagi platform *e-commerce* dalam menyusun strategi harga yang kompetitif (Majid dan Sergeev, 2024).

Penelitian ini menggunakan *Streamlit* sebagai platform untuk *deployment* model prediksi harga ponsel bekas. Pengguna dapat mengakses model *Random Forest Regression* dan memprediksi harga perangkat secara langsung dengan memasukkan variabel seperti merek, kapasitas penyimpanan, dan kondisi fisik ponsel. *Streamlit* mempermudah pengembangan aplikasi web berbasis data dengan antarmuka yang interaktif dan intuitif sebagai *framework Python*. Keunggulan *Streamlit* terletak pada kemampuannya menampilkan visualisasi data dan hasil prediksi model dalam format yang mudah dipahami. Kemampuan ini menjadikan *Streamlit* sebagai alat yang efektif dalam mempercepat proses pengujian dan validasi model, baik di lingkungan penelitian maupun bisnis, melalui aplikasi berbasis web.

1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas berdasarkan latar belakang yang telah diuraikan pada penelitian ini, yaitu sebagai berikut:

1. Bagaimana akurasi yang dihasilkan dari algoritma *Random Forest Regression* berdasarkan *hyperparameter tuning* untuk memprediksi harga ponsel bekas di Indonesia?
2. Apakah fitur-fitur yang mempengaruhi harga ponsel bekas berdasarkan nilai *variable importance*?

1.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan masalah agar lebih spesifik dan terarah. Batasan masalah yang digunakan pada penelitian ini adalah sebagai berikut:

1. Data yang diambil merupakan data ponsel bekas di Indonesia yang diambil dari *website* Jagofon dengan *web scraping* pada 8 Oktober 2024 sebanyak 1.291 data.
2. Fitur yang digunakan untuk memprediksi harga ponsel bekas adalah merek, *details*, *conditions*, *camera*, *sensors*, *selfie*, *sim card*, *network*, *storage*, *curent price*, dan *old price*.
3. *Hyperparameter tuning* yang akan dicari untuk mendapatkan hasil terbaik yaitu menggunakan *Repeated K-Fold Cross Validation* dengan $n_splits = 10$ dan $n_repeats = 4$, serta parameter berikut :
 - a. $n_estimators$: 300, 400, dan 500
 - b. max_depth : 5, 20, dan 35
 - c. $min_samples_split$: 4, 7, dan 15
 - d. $min_samples_leaf$: 1, 3, dan 5

1.4 Tujuan Penelitian

Tujuan penelitian yang sesuai dengan rumusan masalah yang ada adalah sebagai berikut:

1. Memperoleh model untuk menentukan akurasi terbaik yang dihasilkan berdasarkan algoritma *Random Forest Regression* menggunakan *hyperparameter tuning GridSearchCV* dengan *Repeated K-Fold Cross Validation* untuk memprediksi harga ponsel bekas di Indonesia.
2. Menentukan faktor-faktor yang mempengaruhi harga ponsel bekas pada *website* Jagofon berdasarkan nilai *variable importances*.