

BAB IV PERENCANAAN

Perencanaan dari penelitian terdiri dari langkah-langkah penelitian, dan rancangan pengujian yang akan dilakukan oleh peneliti. Oleh karena itu, penjelasan dari perencanaan penelitian adalah berikut ini.

4.1 Langkah-langkah Penelitian

Langkah-langkah penelitian yang dilakukan oleh peneliti dalam membandingkan akurasi algoritma K-Nearest Neighbors dan Decision Tree dalam pemodelan harga mobil bekas adalah sebagai berikut ini.

1. Pertama adalah melakukan studi pustaka mengumpulkan serta menganalisis penelitian-penelitian terdahulu. Dengan studi pustaka tersebut peneliti mendapatkan sebuah subjek atau studi kasus baru yang dapat diuji dengan urgensi yang cukup sehingga dapat dijadikan penelitian. Analisis terdahulu diperlukan untuk mengetahui persamaan dan perbedaan dari penelitian terdahulu dengan penelitian yang dilakukan. Dengan adanya persamaan dan perbedaan dapat memberikan informasi mengenai pengembangan kasus yang diangkat pada penelitian. Tabel persamaan dan perbedaan dari penelitian terdahulu ditunjukkan pada tabel 4.1.

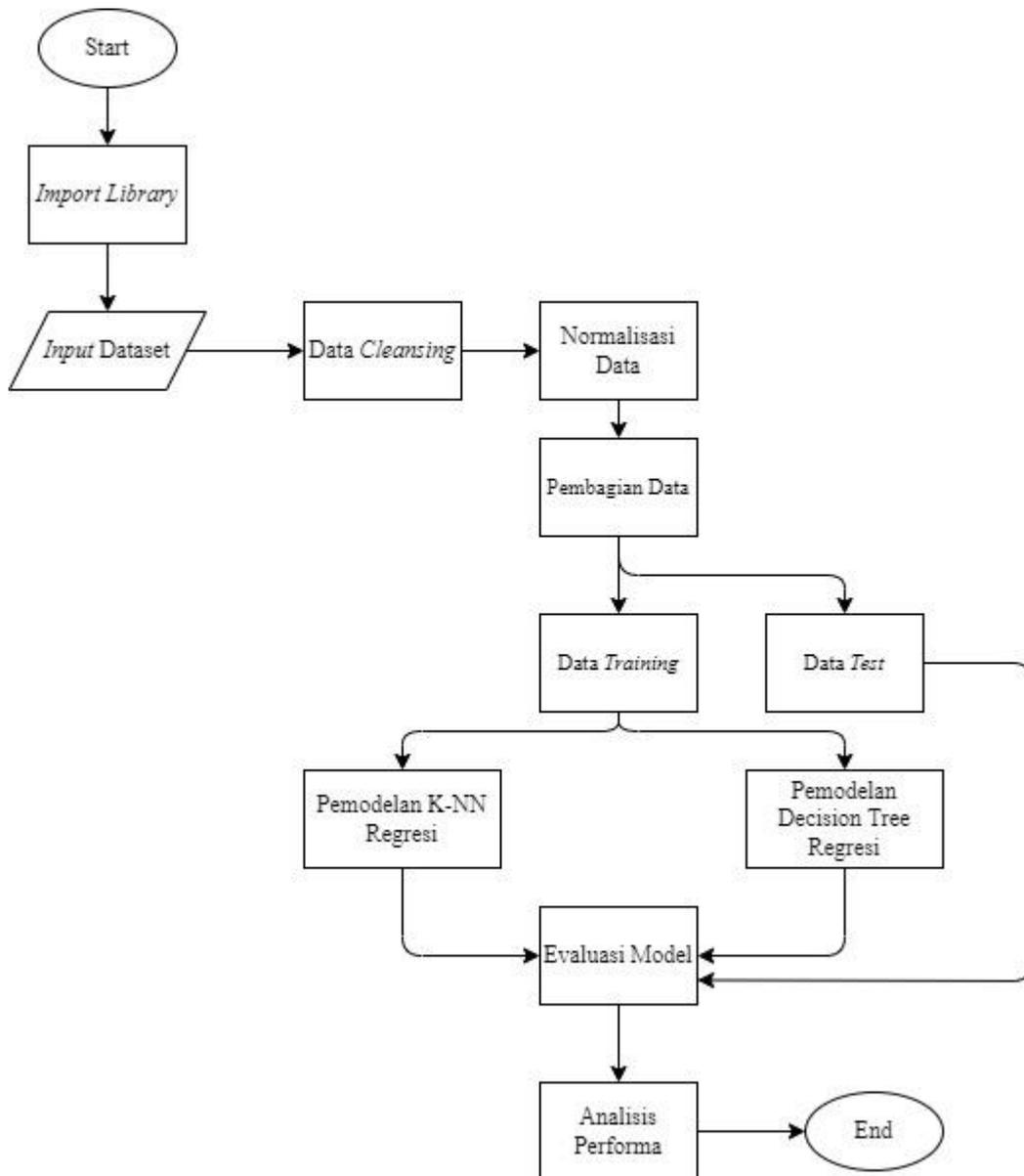
Tabel 4.1 Persamaan dan Perbedaan Penelitian Terdahulu

Penelitian Terdahulu	Persamaan	Perbedaan
<i>“The prediction of scholarship recipients in higher education using K-Nearest Neighbor algorithm”</i> (Kurniadi et al., 2018)	Penggunaan algoritma K-Nearest Neighbors untuk melakukan prediksi. Kemudian hasil akurasi K-Nearest Neighbors yang diterapkan.	Subjek atau studi kasus yang diteliti berbeda. Pada penelitian ini penerapan K-NN untuk melakukan analisis, prediksi, dan pengelompokkan siswa yang mempunyai potensi untuk mendapatkan beasiswa.
<i>“An optimized K-Nearest Neighbors based breast cancer detection”</i> (Assegie, 2021)	Penggunaan algoritma K-Nearest Neighbors. Kemudian menentukan K terbaik untuk menghasilkan akurasi yang baik.	Subjek atau studi kasus yang diteliti berbeda. Pada penelitian ini pencarian K terbaik dilakukan untuk mendapatkan hasil akurasi yang optimal dari K-Nearest Neighbors pada deteksi kanker payudara.

“Penerapan Metode K-Nearest Neighbors Untuk Sistem Rekomendasi Pemilihan Mobil” (Luh Gede Pivin Suwirmayanti, 2017)	Penggunaan algoritma K-Nearest Neighbors.	Subjek penelitian sedikit berbeda, yaitu sistem rekomendasi pemilihan mobil.
“ICA Learning Approach for Predicting RNA-Seq Data Using KNN and Decision Tree Classifiers” (Adebiyi et al., 2020)	Penggunaan algoritma K-Nearest Neighbors dan Decision Tree dalam melakukan prediksi.	Subjek atau studi kasus yang diteliti berbeda. Pada penelitian ini melakukan prediksi dan deteksi penyakit malaria.
“Analysis of Decision Tree and K-Nearest Neighbor Algorithm in the Classification of Breast Cancer” (Rajaguru & Sannasi Chakravarthy, 2019)	Penggunaan algoritma K-Nearest Neighbors dan Decision Tree.	Subjek atau studi kasus yang diteliti berbeda. Pada penelitian ini penerapan dilakukan pada klasifikasi kanker payudara.
“Model Prediksi Kemenangan Tim dalam Game League Of Legend Menggunakan Algoritma Decision Tree” (Sandag, 2021)	Penggunaan algoritma Decision Tree.	Subjek atau studi kasus yang diteliti berbeda. Pada penelitian ini penerapan dilakukan pada prediksi kemenangan tim Game League of Legend.
“Prediksi Harga Mobil Bekas dengan Machine Learning” (Kriswantara et al., 2021)	Subjek atau studi kasus yang diteliti sama.	Algoritma yang digunakan berbeda. Penelitian ini menggunakan <i>Deep Neural Network (DNN)</i> , Decision Tree, dan Random Forest.
“Estimasi Harga Jual Mobil Bekas Menggunakan Metode Regresi Linier Berganda” (Dewi et al., 2020)	Subjek atau studi kasus yang diteliti sama.	Algoritma yang digunakan berbeda. Penelitian ini menggunakan Regresi Linier Berganda.

2. Pencarian data yang akan dijadikan dataset dalam pengujian juga dilakukan oleh peneliti. Dataset yang digunakan peneliti merupakan data sekunder yang diperoleh dari Kaggle yaitu berupa data harga mobil bekas yang berasal dari OLX sesuai dengan keterangan yang terdapat pada Kaggle. Kemudian data tersebut diolah dan dilakukan penyaringan sesuai dengan batasan penelitian yang sudah ditetapkan sebelumnya. Penyaringan data tersebut dilakukan untuk memudahkan peneliti dalam menjaga fokus dalam penelitian. Penyaringan data tersebut dilakukan pada Microsoft Excel.
3. Kemudian peneliti melakukan pembuatan *flowchart* pembuatan model harga mobil bekas untuk dapat menghasilkan perbandingan akurasi kedua

algoritma. *Flowchart* atau diagram alir dari sebuah sistem atau model memberikan gambaran proses atau alur dari sistem atau model yang akan dibangun. Diagram alir juga dapat digunakan sebagai langkah-langkah dari pemodelan harga mobil bekas menggunakan algoritma K-Nearest Neighbors dan Decision Tree.



Gambar 4.1 *Flowchart* Keseluruhan Penelitian

Gambar 4.1 merupakan *flowchart* dari proses keseluruhan penelitian. *Import library* merupakan tahapan paling pertama. *Library* merupakan sebuah kumpulan paket yang terdapat pada Python. Paket-paket tersebut dibuat oleh

perusahaan tertentu dengan tujuan memudahkan kode program sehingga tinggal menggunakan *function* yang sudah terdapat pada paket di dalam *library* tersebut. Kemudian data yang sudah dikumpulkan dimasukkan kemudian memasuki tahap *pre-processing* data. Data yang tidak lengkap akan dibuang yang termasuk ke dalam data *cleansing*. Selanjutnya melakukan normalisasi data. Kemudian data dibagi menjadi kelompok *training* data, dan *test* data. Setelah itu data akan dinormalisasi dan pemodelan dengan K-Nearest Neighbors dan Decision Tree dilakukan. Setelah proses pemodelan, evaluasi model diperlukan untuk mengetahui apakah model sudah optimal. Apabila belum dilakukan penyempurnaan untuk memaksimalkan kinerja model. Terakhir kedua pemodelan dilakukan perbandingan dan dilakukan analisa performa hasil penelitian yang dilakukan.

4. Pembuatan rancangan pengujian yang baik supaya dapat membangun model dengan jelas.
5. Pembuatan model harga mobil bekas.
6. Pengujian algoritma K-Nearest Neighbors dan Decision Tree dalam memperkirakan harga mobil bekas.
7. Terakhir, melakukan analisis performa dari kedua algoritma.

4.2 Rancangan Pengujian

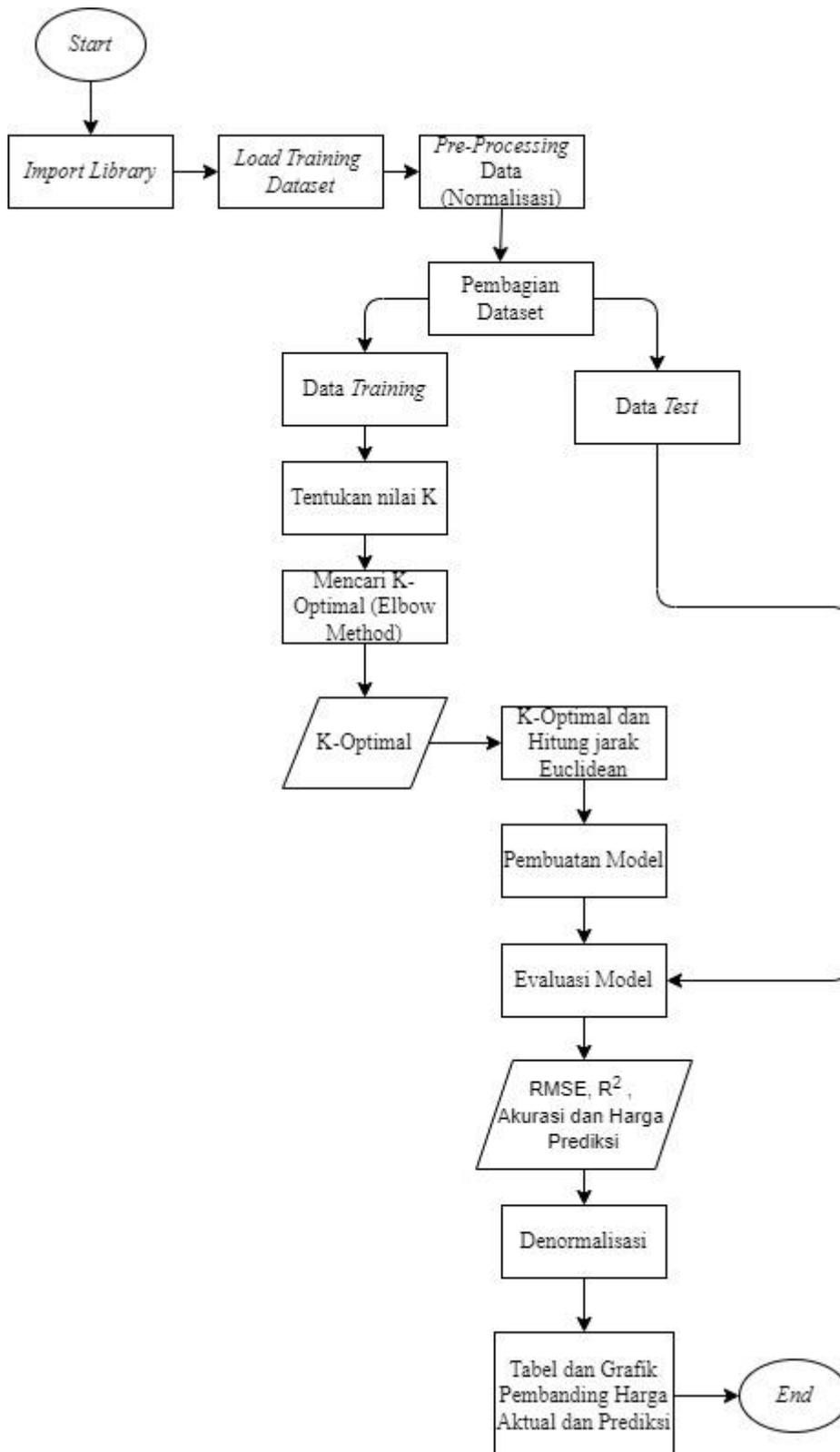
Perancangan pengujian dilakukan untuk memberikan gambaran model yang akan dibangun dengan jelas. Dengan perancangan yang baik dalam membangun model dapat menghasilkan akhir yang sesuai dengan keinginan peneliti atas model yang akan dibangun.

4.2.1 Flowchart K-Nearest Neighbors

Diagram alir atau *flowchart* dari K-Nearest Neighbors ditunjukkan dalam gambar 4.2 di bawah ini. Langkah pertama yang dilakukan dalam proses pemodelan harga mobil bekas menggunakan K-Nearest Neighbors dengan bahasa Python adalah melakukan *import library* yang dibutuhkan kemudian melakukan pemuatan *training* dataset. Data akan melalui tahap *pre-Processing* (*Cleansing* dan

Normalisasi). Setelah itu data akan dibagi menjadi data *training* dan data test. Kemudian untuk menentukan nilai K-optimal menggunakan metode Elbow. Dengan K-optimal dihitung jarak Euclidean. Dengan K-optimal dan perhitungan jarak Euclidean dilakukan prediksi data. Hasil yang dikeluarkan berupa hasil prediksi harga dan akurasi. Setelah melakukan prediksi, evaluasi model dibutuhkan untuk mengetahui apakah model sudah dibuat dengan baik atau tidak. Evaluasi model menggunakan RMSE dan R^2 . Terakhir, denormalisasi dan pembuatan grafik pembandingan harga actual dengan harga prediksi.

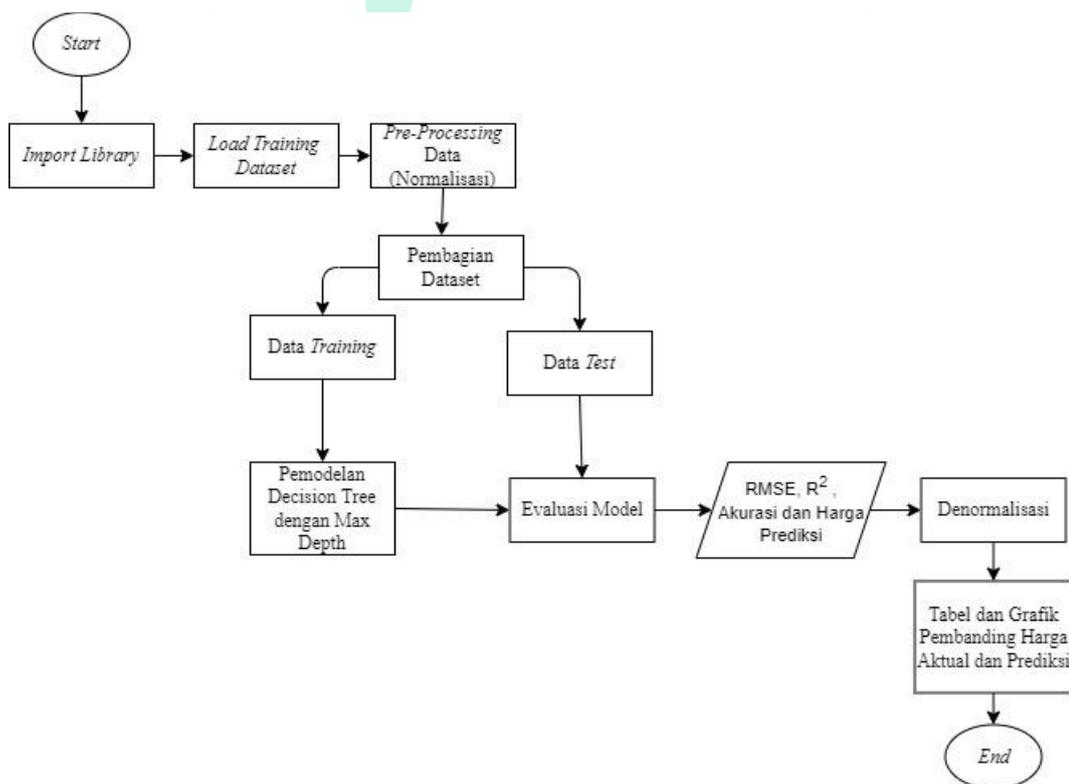




Gambar 4.2 Flowchart K-NN

4.2.2 Flowchart Decision Tree

Diagram alir atau *flowchart* dari pemodelan yang dilakukan algoritma Decision Tree ditunjukkan dalam gambar 4.3 di bawah ini. Langkah pertama yang dilakukan dalam proses pemodelan harga mobil bekas menggunakan Decision Tree dengan bahasa Python adalah melakukan *import library* yang dibutuhkan kemudian melakukan pemuatan *training dataset*. Data akan melalui tahap *pre-Processing* (*Cleansing* dan Normalisasi). Setelah itu data akan dibagi menjadi data *training* dan data test. Kemudian dalam data training dilakukan pemodelan algoritma Decision Tree. Dilakukan prediksi data dengan hasil yang dikeluarkan berupa hasil prediksi harga dan akurasi. Setelah melakukan prediksi, evaluasi model dibutuhkan untuk mengetahui apakah model sudah dibuat dengan baik atau tidak. Evaluasi model menggunakan RMSE dan R^2 . Terakhir, denormalisasi dan pembuatan grafik pembandingan harga actual dengan harga prediksi.



Gambar 4.3 Flowchart Decision Tree

4.2.3 Data

Dengan data yang diperoleh dari Kaggle dilakukan penyaringan untuk melakukan pemodelan mobil bekas sesuai dengan batasan yang ditentukan. Batasan tersebut dibuat untuk memudahkan peneliti dan juga menjaga fokus terhadap inti penelitian. Penyaringan data dilakukan untuk pemilihan fitur yang digunakan untuk prediksi harga. Data dengan ketentuan diantara lain sebagai berikut ini.

1. Kapasitas mesin yang digunakan adalah 1000-1500cc,
2. Model mobil yang digunakan adalah Avanza, Ertiga, Grand livina, dan Xenia,
3. Transmisi mobil yang digunakan adalah manual dan *automatic*,
4. *Input* atau fitur yang digunakan dalam pengujian adalah umur mobil, model, jarak tempuh, kapasitas mesin, dan jenis transmisi mobil,
5. *Output* dalam pengujian adalah harga mobil bekas, dan
6. Hasil akhir adalah akurasi dari pemodelan harga mobil bekas.

Data model dan jenis transmisi diubah menjadi integer dalam proses penyaringan data yang digunakan sesuai dengan batasan yang sudah dijelaskan di atas. Perubahan tersebut ditunjukkan pada tabel 4.2.

Tabel 4.2 Kode Data

No	Kolom	Keterangan	Kode Integer
1	Model	Avanza	101
2	Model	Ertiga	102
3	Model	Grand livina	103
4	Model	Xenia	104
5	Transmisi	Manual	1
6	Transmisi	<i>Automatic</i>	2

Perubahan data dari non-numerik menjadi numerik dilakukan saat proses pre-processing data. Dengan pengubahan data ini memudahkan peneliti untuk melakukan pengujian. Seluruh data yang digunakan dan masuk ke dalam dataset adalah data numerik.