

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DAN
REGRESI LINEAR DALAM PREDIKSI HARGA MOBIL BEKAS**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

MUHAMMAD FARID ALGHOZI

19.11.2824

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DAN
REGRESI LINEAR DALAM PREDIKSI HARGA MOBIL BEKAS**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

MUHAMMAD FARID ALGHOZI

19.11.2824

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA**

YOGYAKARTA

2023

HALAMAN PERSETUJUAN

SKRIPSI

IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DAN
REGRESI LINEAR DALAM PREDIKSI HARGA MOBIL BEKAS

yang disusun dan diajukan oleh

Muhammad Farid Alghozi

19.11.2824

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 4 Juli 2023

Dosen Pembimbing,



Rumini, M.Kom
NIK. 190302246

HALAMAN PENGESAHAN

SKRIPSI

IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DAN
REGRESI LINEAR DALAM PREDIKSI HARGA MOBIL BEKAS

yang disusun dan diajukan oleh

Muhammad Farid Alghozi

19.11.2824

Telah dipertahankan di depan Dewan Penguji
pada tanggal 4 Juli 2023

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dina Maulina, M.Kom
NIK. 190302250

Wiwi Widayani, M.Kom
NIK. 190302272

Rumini, M.Kom
NIK. 190302246



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 4 Juli 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.
NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Muhammad Farid Alghozi
NIM : 19.11.2824

Menyatakan bahwa Skripsi dengan judul berikut:

Implementasi Algoritma K-Nearest Neighbor dan Regresi Linear Dalam Prediksi Harga Mobil Bekas

Dosen Pembimbing : Rumini, M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 4 Juli 2023

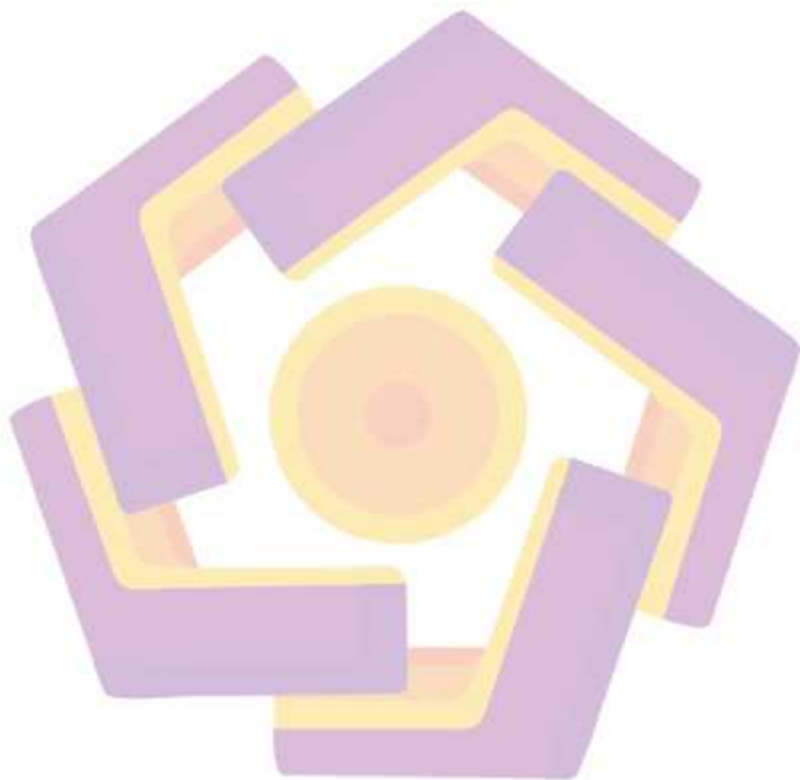
Yang Menyatakan,



Muhammad Farid Alghozi

HALAMAN PERSEMBAHAN

Skripsi ini penulis persembahkan kepada orang tua, kakak, dan adik yang tersayang serta teman-teman yang tercinta.



KATA PENGANTAR

Segala puji bagi Allah SWT serta sholawat dan salam tercurah kepada Rasulullah Muhammad SAW. Penulis menyelesaikan skripsi ini guna memenuhi persyaratan menyelesaikan program studi Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Penulis menyadari bahwa skripsi ini tidak akan terwujud tanpa adanya bantuan dari berbagai pihak. Untuk itu penulis menghaturkan terima kasih kepada yang terhormat:

1. Bapak Hanif Al Fatta, S.Kom., M.Kom. selaku dekan Fakultas Ilmu Komputer.
2. Ibu Windha Mega Pradya D, M.Kom. selaku Kaprodi Informatika.
3. Ibu Rumini, M.Kom. selaku dosen pembimbing.
4. Bapak dan Ibu dosen program studi Informatika.
5. Teman-teman mahasiswa program studi Informatika.

Penulis berharap semoga budi baiknya dibalas yang lebih oleh Allah SWT. Demikian pula semoga skripsi ini bermanfaat bagi pembacanya dalam memahami konsep dasar data mining dan penerapan algoritma data mining.

Penulis juga menyadari bahwa skripsi ini tak luput dari kesalahan dan kekurangan, maka dari itu Penulis menerima kritik dan saran dengan hati yang terbuka.

Yogyakarta, 4 Juli 2023

Penulis

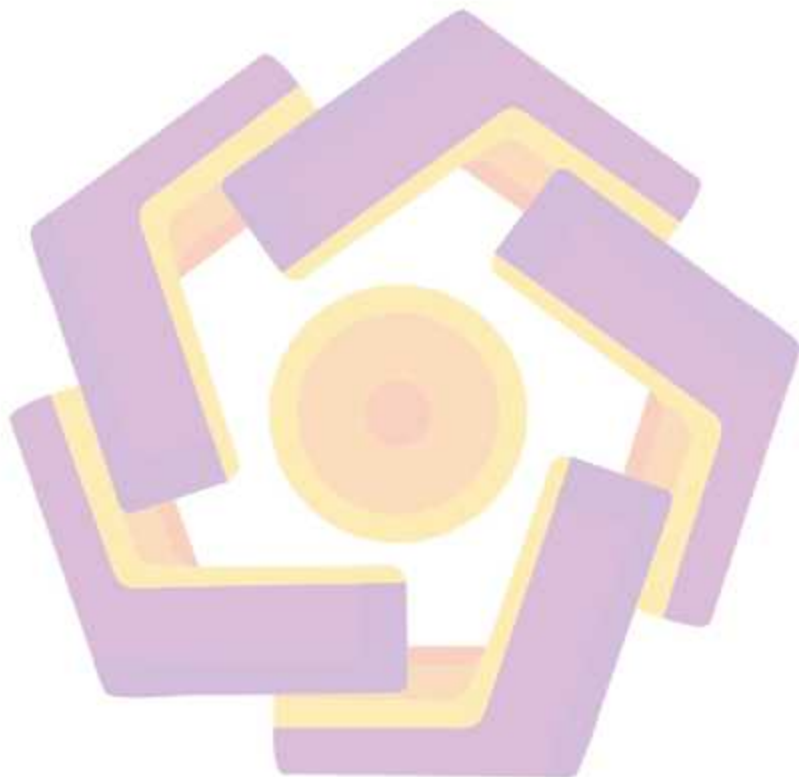
DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
INTISARI	xii
ABSTRACT	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	10
2.2.1 Data Mining	10
2.2.2 K-Nearest Neighbor	11
2.2.3 Regresi Linear	12
2.2.4 Root Means Square Error	13
2.2.5 Mobil Bekas	14
2.2.6 Jupyter Notebook	14

BAB III METODE PENELITIAN	15
3.1 Objek Penelitian	15
3.2 Alur Penelitian	15
3.3 Alat dan Bahan	16
3.3.1 Hardware	16
3.3.2 Software	16
3.4 Dataset	17
3.4.1 Preprocessing Data	17
3.4.2 Pemilihan Variabel	21
3.4.3 Modeling Data	27
BAB IV HASIL DAN PEMBAHASAN	29
4.1 Penerapan Algoritma K-Nearest Neighbor	29
4.2 Penerapan Algoritma Regresi Linear	31
4.2 Perbandingan Hasil antara Algoritma K-Nearest Neighbor dan Regresi Linear	31
4.3 Pengujian secara Manual dengan Data Spesifik	32
BAB V PENUTUP	34
5.1 Kesimpulan	34
5.2 Saran	34
REFERENSI	35
LAMPIRAN	39

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	6
Tabel 2.1 Keaslian Penelitian	23
Tabel 4.1 Perbandingan Hasil Akurasi K-Nearest Neighbor dan Regresi Linear	31
Tabel 4.2 Pengujian secara Manual dengan Data Spesifik	31



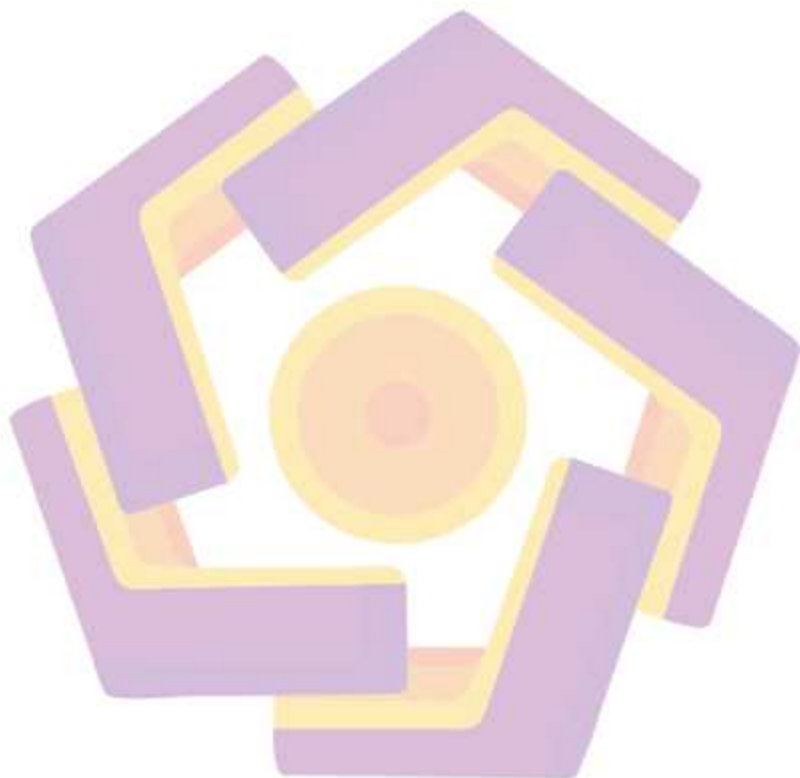
DAFTAR GAMBAR

Gambar 3.1 Alur Penelitian	15
Gambar 3.2 Sampel Dataset	17
Gambar 3.3 Kolom dengan jumlah data yang kosong atau null	18
Gambar 3.4 Fungsi untuk penghapusan kolom dan data yang duplikat	18
Gambar 3.5 Fungsi untuk deteksi dan hapus data outlier pada kolom Harga	19
Gambar 3.6 Fungsi untuk mengisi data yang kosong atau null	20
Gambar 3.7 Fungsi untuk mengubah tipe data pada kolom	20
Gambar 3.8 Tipe kolom setelah diubah	21
Gambar 3.9 Fungsi untuk penghapusan sejumlah data pada kolom Model	21
Gambar 3.10 Matriks Tingkat Korelasi Variabel Bertipe Numerikal	22
Gambar 3.11 Fungsi untuk mengubah kolom bertipe kategorikal menjadi numerikal	27
Gambar 3.12 Fungsi untuk penskalaan pada kolom Jarak tempuh dan Harga	28
Gambar 3.13 Fungsi untuk pembagian data menjadi data latih dan data uji	28
Gambar 4.1 Nilai Akurasi K-Nearest Neighbor dengan nilai $K = 5$	29
Gambar 4.2 Grafik Metode Elbow	30
Gambar 4.3 Nilai Akurasi K-Nearest Neighbor dengan nilai $K = 1$	30
Gambar 4.4 Nilai Akurasi Regresi Linear	31

DAFTAR LAMPIRAN

Lampiran 1. Source Code

39



INTISARI

Mobil merupakan salah satu alat transportasi yang banyak digunakan oleh masyarakat. Jenis-jenis mobil barupun terus bermunculan seiring berkembangnya zaman. Masyarakat akan menjual mobil lamanya untuk membeli mobil jenis terbaru. Tetapi ada sebagian masyarakat yang lebih memilih untuk membeli mobil keluaran lama dengan kondisi bekas dikarenakan harganya yang lebih murah. Dalam menentukan pilihan saat memilih mobil bekas yang tepat dengan harga yang sesuai keinginan, maka calon pembeli harus memiliki informasi yang lebih akurat dan efisien. Beberapa faktor perlu dipertimbangkan dalam memilih mobil bekas, seperti tahun, jarak tempuh, kapasitas mesin, merek, model, varian, tipe bahan bakar, transmisi, dan tipe bodi. Berdasarkan hal tersebut maka penelitian ini bertujuan untuk membantu calon pembeli dengan cara melakukan penerapan Algoritma K-Nearest Neighbor dan Regresi Linear dalam memprediksi harga mobil bekas. Dataset yang digunakan bersumber dari situs Kaggle yaitu dataset mobil bekas. Penelitian ini menghasilkan tingkat akurasi sebesar 90% dengan nilai RMSE 56.56 untuk algoritma K-Neighbor sedangkan algoritma Regresi Linear menghasilkan tingkat akurasi sebesar 83% dengan nilai RMSE 74.73. Melalui hasil tersebut dapat disimpulkan bahwa dalam penelitian ini penerapan algoritma K-Nearest Neighbor memiliki tingkat akurasi yang lebih tinggi dibandingkan algoritma Regresi Linear.

Kata kunci: prediksi harga mobil bekas, algoritma K-Nearest Neighbor, algoritma Regresi Linear.

ABSTRACT

The car is a means of transportation that is widely used by the community. New types of cars keep popping up with the times. People will sell their old cars to buy the newest types of cars. But there are some people who prefer to buy old cars with used conditions because the prices are cheaper. In making choices when choosing the right used car at the price you want, the prospective buyer must have information that is more accurate and efficient. Several factors need to be considered in choosing a used car, such as year, mileage, engine capacity, make, model, variant, fuel type, transmission, and body type. Based on this, this study aims to help prospective buyers by applying the K-Nearest Neighbor Algorithm and Linear Regression in predicting used car prices. The dataset used is sourced from the Kaggle site, namely the used car dataset. This study produces an accuracy rate of 90% with an RMSE value of 56.56 for the K-Neighbor algorithm while the Linear Regression algorithm produces an accuracy rate of 83% with an RMSE value of 74.73. Through these results it can be concluded that in this study the application of the K-Nearest Neighbor algorithm has a higher level of accuracy than the Linear Regression algorithm.

Keyword: used car price prediction, K-Nearest Neighbor algorithm, Linear Regression algorithm.