

**PENGEMBANGAN SISTEM PREDIKSI HARGA MOBIL BEKAS OLX
MENGUNAKAN ALGORITHMMA *RANDOM FOREST***

SKRIPSI

**Disusun Untuk Memenuhi Sebagai Persyaratan
Mencapai Derajat Sarjana Informatika**



Disusun Oleh:

**Ervier Priande
1800018009**

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS AHMAD DAHLAN
YOGYAKARTA**

2024

**DEVELOPMENT OF OLX USED CAR PRICE PREDICTION SYSTEM USING
RANDOM FOREST ALGORITHM**

THESIS

**Arranged To As fulfil
as a requirement for Achieving a Bachelor's Degree in Informatics**



Arranged by :

**Ervier Priande
1800018009**

**INFORMATICS STUDY PROGRAM
INDUSTRIAL TECHNOLOGY FACULTY
AHMAD DAHLAN UNIVERSITY
YOGYAKARTA
2024**

LEMBAR PERSETUJUAN PEMBIMBING

SKRIPSI

**PENGEMBANGAN SISTEM PREDIKSI HARGA MOBIL BEKAS OLX
MENGUNAKAN ALGORITHMMA *RANDOM FOREST***



Telah disetujui oleh:

Pembimbing



Guntur Maulana Zamroni, B.Sc., M.Kom.

NIPM. 19840309 201810 111 1205917

LEMBAR PENGESAHAN

SKRIPSI

**PENGEMBANGAN SISTEM PREDIKSI HARGA MOBIL BEKAS OLX
MENGUNAKAN ALGORITHMMA RANDOM FOREST**

Dipersiapkan dan disusun oleh:

ERVIER PRIANDE

1800018009

Telah dipertahankan di depan Dewan Penguji
pada 13 Mei 2024
dan dinyatakan telah memenuhi syarat

Susunan Dewan Penguji

Ketua : Guntur Maulana Zamroni, B.Sc., M.Kom

Penguji 1 : Lisna Zahrotun, S.T., M.Cs.

Penguji 2 : Dr. Ardiansyah, S.T., M.Cs.

R/6'24
13/6 2024



Yogyakarta, Juni 2024
Dekan Fakultas Teknologi Industri
Universitas Ahmad Dahlan

Prof. Dr. Ir. Siti Jamilatun, M.T.
NIPM. 19660812 199601 011 0784324

LEMBAR PERNYATAAN KEASLIAN

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Ervier Priande

NIM : 1800018009

Prodi : Informatika

Judul TA/Skripsi : PENGEMBANGAN SISTEM PREDIKSI HARGA MOBIL BEKAS OLX
MENGUNAKAN ALGORITHMMA *RANDOM FOREST*

Dengan ini saya menyatakan bahwa Laporan Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar Ahli Madya/Kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, Juni 2024

Mengetahui,
Dosen Pembimbing



Guntur Maulana Zamroni, B.Sc., M.Kom
NIPM. 19840309 201810 111 1205917

Yang menyatakan,



Ervier Priande
NIM. 1800018009

PERNYATAAN TIDAK PLAGIAT

Saya yang bertanda tangan di bawah ini:

Nama : Ervier Priande
NIM : 1800018009
Email : ervier1800018009@webmail.uad.ac.id
Program Studi : S1 Informatika
Fakultas : Teknologi Industri
Judul Tesis : Pengembangan Sistem Prediksi Harga Mobil Bekas Olx
Menggunakan Algoritma *Random Forest*

Dengan ini menyatakan bahwa:

1. Hasil karya yang saya serahkan ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar kesarjanaan baik di Universitas Ahmad Dahlan maupun di institusi pendidikan lainnya.
2. Hasil karya saya ini bukan saduran/terjemahan melainkan merupakan gagasan, rumusan, dan hasil pelaksanaan penelitian/implementasi saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing akademik dan narasumber penelitian.
3. Hasil karya saya ini merupakan hasil revisi terakhir setelah diujikan yang telah diketahui dan disetujui oleh pembimbing.
4. Dalam karya saya ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali yang digunakan sebagai acuan dalam naskah dengan menyebutkan nama pengarang dan dicantumkan dalam daftar pustaka.

Pernyataan ini saya buat dengan sesungguhnya. Apabila di kemudian hari terbukti ada penyimpangan dan ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya saya ini, serta sanksi lain yang sesuai dengan ketentuan yang berlaku di Universitas Ahmad Dahlan.

Yogyakarta, Juni 2024

Yang Menyatakan



Ervier Priande

PERNYATAAN PERSETUJUAN AKSES

Saya yang bertanda tangan di bawah ini:

Nama : Ervier Priande
NIM : 1800018009
Email : ervier1800018009@webmail.uad.ac.id
Program Studi : S1 Informatika
Fakultas : Teknologi Industri
Judul Tesis : Pengembangan Sistem Prediksi Harga Mobil Bekas Olx
Menggunakan Algorithma *Random Forest*

Dengan ini Saya menyerahkan hak sepenuhnya kepada Perpustakaan Universitas Ahmad Dahlan untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tesis elektronik sebagai berikut (beri tanda pada kotak):



Saya (~~tidak mengizinkan~~ **mengizinkan**)* karya tersebut diunggah ke dalam aplikasi Repository Perpustakaan Universitas Ahmad Dahlan. Demikian pernyataan ini Saya buat dengan sebenarnya.

Yogyakarta, Juni 2024

Yang Menyatakan



Ervier Priande

Mengetahui,

Dosen Pembimbing Skripsi



Guntur Maulana Zamroni, B.Sc., M.Kom.
NIPM. 19840309 201810 111 1205917

MOTTO HIDUP

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Hadits qudsy

أَنَا اللَّهُ لَا إِلَهَ إِلَّا أَنَا مَنْ لَمْ يَشْكُرْ عَلَيَّ نِعْمَاتِي وَلَمْ يَصْنِرْ عَلَيَّ بَلَائِي وَلَمْ يَرْضَ بِقَضَائِي
فَلْيَتَّخِذْ رَبًّا سِوَايَ

Artinya: Aku Allah, tiada Tuhan melainkan Aku; siapa tidak bersyukur atas nikmat-nikmat pemberian-Ku, tidak bersabar atas ujian-Ku dan ridla terhadap kepastian qadla-Ku, maka carilah Tuhan selain Aku.



HALAMAN PERSEMBAHAN

Dengan rasa syukur yang sangat mendalam , dengan telah diselesaikanya Skripsi ini oleh penulis Mempersembahkan kepada :

1. Tuhan yang maha Esa ALLAH S.W.T
2. Kedua orang tua Roswer Dahnes Sastra dan Ernita Sari yang tidak ada henti – hentinya mendoakan dan mensupport penulis baik secara moril dan materil.
3. Tante saya Restu Aie yang selalu medoakan dan menyemangati
4. keluarga besar saya yang berada dikampung halaman yang telah memberikan doa kepada saya.
5. Bapak Bambang Robi'in, S.T, M,T selaku dosen pembimbing akademik, terimakasih atas pengarahan serta bimbingan kepada saya selama perkuliahan.
6. Bapak Guntur Maulana Zamroni Bsc, M.kom selaku dosen pembimbing skripsi, terimakasih telah meluangkan waktunya untuk membimbing saya dan memberikan arahan dalam penulisan skripsi ini
7. Ibu Lisna Zahrotun S.T, M.cs dan Bapak Dr. Ardiansyah, S.T, M.cs terimakasih telah menyempatkan waktu menguji saya dan memberikan pengarahan penulisan dalam Skripsi ini.
8. Teman-teman seperjuangan saya , satu angkatan saya infromatika 2018 'Skuy A' yang telah memberikan motivasi dan semangat belajar.
9. Sahabat saya Andri Al-Hafid , Ari Tuharea, Muhammad Basyir dan Ridhan Aufa Mahfud yang telah menemani saya baik suka maupun duka selama berada dimasa kuliah.
10. Untuk teman teman saya yang dikampung yang telah mendoakan , teriamakasi atas semua rasa pertemanan yang pernah ada.

KATA PENGANTAR

Puji syukur saya panjatkan kehadirat Allah SWT. Karena hidayah dan inayahnya dapat memberikan nikmat dan karunianya, sehingga saya dapat menyelesaikan skripsi yang berjudul “Pengembangan Sistem Prediksi Harga Mobil Bekas Olx Menggunakan Algoritma Random Forest.” Dalam penulisan ini terdapat kekurangan dalam penyajiannya, sehingga kritik dan saran yang membangun dapat menyempurnakan saya ini.

Pada penulisan ini banyak yang membantu dalam penulisan berupa motivasi maupun dukungan yang telah memberikan waktu, tenaga, doa, dan perhatiannya untuk kelancaran dalam penyusunan skripsi. Dalam penulisan laporan ini penulis banyak menerima masukan dan dukungan dari berbagai pihak, oleh karena itu penulis ingin mengucapkan banyak terima kasih kepada:

1. Prof. Dr. Muchlas, M.T. sebagai Rektor Universitas Ahmad Dahlan.
2. Prof. Dr. Ir. Siti Jamilatun, M.T. sebagai Dekan Fakultas Teknologi Industri Universitas Ahmad Dahlan.
3. Bapak Dr. Murinto, S.Si., M.Kom. sebagai ketua program studi informatika
4. Bapak Guntur Maulana Zamroni, B.Sc., M.Kom. sebagai dosen pembimbing Skripsi yang telah meluangkan waktu ditengah kesibukan beliau, memberikan saran serta pengarahan kepada penulis dalam proses penulisan skripsi ini.
5. Ibu Lisna Zahrotun, S.T., M.cs. dan bapak Dr. Ardiansyah, S.T., M.cs. selaku dosen penguji yang telah meluangkan waktunya untuk menguji dan memberi arahan dalam penulisan skripsi ini.
6. Bapak Bambang Robi'in, S.T., M.T. selaku Dosen Pembimbing Akademik yang telah memberikan bimbingan dan bantuan selama penulis menempuh studi di Fakultas Teknologi Industri Universitas Ahmad Dahlan.
7. Bapak dan ibu dosen prodi infomatika Fakultas Teknologi Industri Universitas Ahmad Dahlan yang telah memberikan ilmu pengetahuan kepada penulis
8. kedua orang tua penulis, Roswer Dahnes Sastra dan Ernita Sari, Terimakasih atas segala kasih sayang yang diberikan dalam membersarkan dan membimbing penulis selama ini sehingga penulis dapat terus berjuang dalam meraih mimpi dan cita cita .
9. Teman-teman penulis, Skuy A Prodi Infomatika Angkatan 2018 yang telah memberikan semangat serta berbagi ilmu kepada penulis untuk terus berkembang.

10. Sahabat penulis, Andri Al-Hafid, Bukhari Fitra Tuharea, Muhammad Basir, dan Ridhan Afa Mahfud , yang telah berjuang bersama dan berbagi suka duka selama perkuliahan. Semoga persahabatan kita terus berlanjut hingga nanti kita tua.
11. Semua pihak yang tidak dapat disebutkan satu persatu yang telah banyak membantu memberikan pemikiran dan pandangan demi kelancaran dan keberhasilan penyusunan skripsi ini.
12. Terakhir, terima kasih untuk Ervier Priande, karena telah mampu berjuang dan bertahan hingga sejauh ini. Bertahan dari berbagai tekanan diluar keadaan dan tak pernah memutuskan untuk menyerah walaupun sesulit apapun keadaannya. Sehingga bisa memenuhi tanggung jawab sebagai anak dan memberikan bukti kepada orang tua penulis bahwa penulis masih ingin belajar dan berkembang seberat apapun keadaannya.

Semoga skripsi ini dapat memberikan manfaat dan kontribusi bagi perkembangan ilmu pengetahuan dan teknologi. Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, oleh karena itu kritik dan saran yang membangun sangat penulis harapkan guna perbaikan di masa yang akan datang. Akhir kata, semoga Allah SWT senantiasa memberikan rahmat dan hidayah-Nya kepada kita semua. Amin

Yogyakarta, 20 Juni 2024
Penulis



Ervier Priande

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING.....	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR PERNYATAAN KEASLIAN.....	iv
PERNYATAAN TIDAK PLAGIAT.....	v
PERNYATAAN PERSETUJUAN AKSES.....	vi
KATA PENGANTAR.....	viii
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	ix
DAFTAR KODE PROGRAM.....	x
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN.....	xiii
ABSTRAK.....	xiv
BAB 1 PENDAHULUAN.....	1
1.1. Latar belakang.....	1
1.2. Batasan masalah.....	4
1.3. Rumusan masalah.....	5
1.4. Tujuan penelitian.....	5
1.5. Manfaat penelitian.....	5
BAB II KAJIAN TEORITIS.....	6
2.1 Kajian Penelitian Terdahulu.....	6
2.2 Landasan teori.....	9
A. <i>Web mining dan Web scraping</i>	9
B. Classification and Rgression Tree (CART).....	11
C. <i>Random forest</i>	12
D. <i>Machine learning</i>	13

E. Python	14
F. Flask.....	15
G. Pengujian Akurasi	16
H. Web Development.....	17
BAB III METODOLOGI PENELITIAN	19
3.1. Metode Pengumpulan Data.....	19
3.2. Kebutuhan Hardware dan Software.....	19
3.3. Analisis Kebutuhan Data	20
3.4. Tahapan Penelitian	22
BAB IV HASIL DAN PEMBAHASAN.....	26
4.1 Implementasi.....	26
4.1.1 <i>Web scraping</i>	26
4.1.2 Preprocessing	26
4.1.3 Analisa Deskriptif.....	28
4.1.4 Label Encoding.....	30
4.2 Pengujian.....	34
4.2.1. Menentukan Variabel Dependen dan Independen	34
4.2.2. Data Training dan Testing.....	35
4.2.3. Model <i>Random forest</i>	36
4.2.4. Evaluasi model.....	38
4.3 Perbandingan Data Prediksi dan Data Aktual	41
4.3.1 Hasil Tampilan <i>Website</i>	43
BAB V PENUTUP	48
5.1 Kesimpulan	48
5.2 Saran	49
DAFTAR PUSTAKA	50
LAMPIRAN 1	54

LAMPIRAN 2	60
LAMPIRAN 3	61
LAMPIRAN 4	63
LAMPIRAN 5	67

DAFTAR GAMBAR

Gambar 2.1 Classification and Regression.....	11
Gambar 2.2 <i>Random forest</i>	13
Gambar 3.1 Tahapan penelitian	22
Gambar 4.1 Hasil <i>Web scraping</i>	26
Gambar 4.3 Persentase Mobil Berdasarkan Jenis Bahan Bakar	29
Gambar 4.4 Persentase Mobil Berdasarkan Jenis Bahan bakar	30
Gambar 4.5 Halaman From Prediksi	44
Gambar 4.6 Perbandingan Antara Prediksi Data Training Dan Testing Aktual.....	41
Gambar 4.7 Form Prediksi.....	43
Gambar 4.8 Pemilihan Merek Mobil.....	44
Gambar 4.9 Pemilihan Nama Mobil.....	44
Gambar 4.10 Pemilihan Model Mobil.....	45
Gambar 4.11 Pemilihan Transmisi Mobil.....	45
Gambar 4.12 Harga Prediksi Mobil.....	56

DAFTAR KODE PROGRAM

Kode program 4.1 Proses Labeling Kilometer (Km)	28
Kode program 4.2 Proses label encoding	31
Kode program 4.3 Proses Penentuan Variabel Dependen dan Independen	35
Kode program 4.4 Proses Pembagian Data.....	35
Kode program 4.5 Proses Algorithma Random forest	36
Kode program 4.6 Pengujian Hasil Prediksi	37
Kode program 4.7 Proses Pembuatan Model	37
Kode program 4.8 akurasi prediksi menggunakan Mean Absolute Error (MAE).....	39
Kode program 4.9 kode program Mean Percentage Error (MAPE)	40
Kode program 4.10 kode program mengenai Absolute Error (AE).....	38
Kode program 4.11 kode program mengenai form prediksi	43

DAFTAR TABEL

Tabel 2.1 Penelitian terdahulu.....	7
Tabel 3.1 Definisi Operasional Variabel (DOV).....	20
Tabel 4.1 Kategori kilometer (Km)	27
Tabel 4.2 Label Merek.....	31
Tabel 4.3 Nama Mobil	31
Tabel 4.4 Model Mobil	32
Tabel 4.5 Jenis Bahan Bakar.....	32
Tabel 4.6 Jenis Transmisi.....	33
Tabel 4.7 Hasil Data Training dan Data Testing	35
Tabel 4.8 Hasil Akurasi dan Jumlah tree	37
Tabel 4.9 Hasil perbandingan harga aktuan dengan hasil prediksi	42

DAFTAR LAMPIRAN

LAMPIRAN 1. Data Product.....	50
LAMPIRAN 2. Label Encoding Merek.....	55
LAMPIRAN 3. Label Encoding Nama.....	56
LAMPIRAN 4. Label Encoding Model.....	58
LAMPIRAN 5. Data Hasil Perbandingan.....	62

PENGEMBANGAN SISTEM PREDIKSI HARGA MOBIL BEKAS OLX MENGUNAKAN ALGORITMA RANDOM FOREST

ERVIER PRIANDE
ABSTRAK

Mobil merupakan kendaraan beroda empat yang populer karena kenyamanannya dalam bepergian dan perlindungan dari cuaca ekstrem. Dalam pasar mobil bekas yang berkembang, prediksi harga mobil bekas secara manual kerap tidak akurat karena keterbatasan data, waktu, kesalahan manusia, dan ketidakmampuan beradaptasi dengan perubahan pasar. Penelitian ini bertujuan membantu pengusaha mobil bekas memprediksi harga dengan lebih akurat dan efisien menggunakan algoritma *Random Forest*

Penelitian ini terdiri dari beberapa tahapan. Tahapan pertama yaitu mengumpulkan data produk kendaraan dari *olx.com* menggunakan *webscrapping Python* dilanjutkan dengan preprocessing, analisis deskriptif, label encoding, menentukan variabel independen dan dependen dan pembagian data set untuk pengujian. Selanjutnya dilakukan konfigurasi model untuk melakukan beberapa pengujian dengan parameter yang berbeda. Setelah model pengujian prediksi terbentuk dilakukan penyimpanan model ke dalam format *pickle* serta melakukan perbandingan harga prediksi dengan harga sebenarnya.

Hasil analisis menunjukkan bahwa merek mobil Toyota mendominasi dengan jumlah 580 unit, diikuti oleh Honda (348) dan Suzuki (153). Sebagian besar mobil menggunakan bahan bakar bensin (1684) daripada solar (167). Transmisi otomatis lebih umum (1462) dibandingkan dengan transmisi manual (389). Hasil prediksi harga mobil bekas menggunakan Algoritma *Random forest* menghasilkan tingkat akurasi sebesar 94,79% dan *Mean Absolute Percentage Error* (MAPE) sekitar 5.21% dan *Mean Absolute Error* (MAE) 9729279.74 dengan penggunaan 250 pohon ($n_estimator$), dengan perbedaan rata-rata antara harga aktual dan prediksi sekitar 0,10%.

Kata kunci : prediksi harga, *random forest*, *webscrapping*.

DEVELOPMENT OF OLX USED CAR PRICE PREDICTION SYSTEM USING RANDOM FOREST ALGORITHM

**ERVIER PRIANDE
ABSTRACT**

Cars are four-wheeled vehicles that are popular because of their comfort in traveling and protection from extreme weather. In the growing used car market, manual predictions of used car prices are often inaccurate due to limited data, time, human error, and the inability to adapt to market changes. This research aims to help used car entrepreneurs predict prices more accurately and efficiently using the Random Forest algorithm

This research consists of several stages. The first stage is collecting vehicle product data from olx.com using Python webscraping followed by preprocessing, descriptive analysis, label encoding, determining independent and dependent variables and dividing the data set for testing. Next, the model configuration is carried out to carry out several tests with different parameters. After the prediction testing model is formed, save the model in pickle format and compare the predicted prices with the actual prices.

The analysis results show that the Toyota car brand dominates with 580 units, followed by Honda (348) and Suzuki (153). Most cars use petrol (1684) rather than diesel (167). Automatic transmissions are more common (1462) compared to manual transmissions (389). The results of predicting used car prices using the Random Forest algorithm produce an accuracy rate of 94.79% and a Mean Absolute Percentage Error (MAPE) of around 5.21% and a Mean Absolute Error (MAE) of 9729279.74 with the use of 250 trees (n_estimator), with an average difference between actual and predicted prices are around 0.10%.

Keywords: price prediction, random forest, webscraping,