



**Komparasi Algoritma Regresi Linier Berganda dan *Decision Tree Regressor*
Untuk Memprediksi Harga Mobil Bekas**

TUGAS AKHIR

Wisnu Nugroho Setya Bakti

41518010114

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA**

2022



**Komparasi Algoritma Regresi Linier Berganda dan *Decision Tree Regressor*
Untuk Memprediksi Harga Mobil Bekas**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Wisnu Nugroho Setya Bakti

41518010114

MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2022

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41518010114

Nama : Wisnu Nugroho Setya Bakti

Judul Tugas Akhir : Komparasi Algoritma Regresi Linier Berganda dan Decision Tree Regressor Untuk Memprediksi Harga Mobil Bekas

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

UNIVERSITAS
MERCU BUANA

Jakarta, 16 Juni 2022



Wisnu Nugroho Setya Bakti

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Wisnu Nugroho Setya Bakti
NIM : 41518010114
Judul Tugas Akhir : Komparasi Algoritma Regresi Linier Berganda dan Decision Tree Regressor Untuk Memprediksi Harga Mobil Bekas

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana Hak Bebas Royalti Noneksklusif (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 16 Juni 2022



Wisnu Nugroho Setya Bakti

SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Wisnu Nugroho Setya Bakti
 NIM : 41518010114
 Judul Tugas Akhir : Komparasi Algoritma Regresi Linier Berganda dan *Decision Tree Regressor* Untuk Memprediksi Harga Mobil Bekas

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

No	Luaran	Jenis		Status	
1	Publikasi Ilmiah	Jurnal Nasional Tidak Terakreditasi		Diajukan	✓
		Jurnal Nasional Terakreditasi	✓		
		Jurnal Internasional Tidak Bereputasi		Diterima	
		Jurnal Internasional Bereputasi			
Disubmit/dipublikasikan di :	Nama Jurnal	: Jurnal Media Informatika Budidarma			
	ISSN	: 2548-8368			
	Link Jurnal	: https://ejournal.stmik-budidarma.ac.id/index.php/mib/index			
	Link File Jurnal Jika Sudah di Publish				

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 04 Juli 2022



Wisnu Nugroho Setya Bakti

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010114
Nama : Wisnu Nugroho Setya Bakti
Judul Tugas Akhir : Komparasi Algoritma Regresi Linier Berganda dan
Decision Tree Regressor Untuk Memprediksi Harga
Mobil Bekas

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 29 Juli 2022




UNIVERSITAS
MERCU BUANA
(Puji Rahayu, Dr, MT)

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010114
Nama : Wisnu Nugroho Setya Bakti
Judul Tugas Akhir : Komparasi Algoritma Regresi Linier Berganda dan Decision Tree Regressor Untuk Memprediksi Harga Mobil Bekas

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 29 Juli 2022



(Saruni Dwiasnati, ST, MM, M.Kom)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518010114
Nama : Wisnu Nugroho Setya Bakti
Judul Tugas Akhir : Komparasi Algoritma Regresi Linier Berganda dan Decision Tree Regressor Untuk Memprediksi Harga Mobil Bekas

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 29 Juli 2022


(Wawan Gunawan, S.Kom, MT)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41518010114
Nama : Wisnu Nugroho Setya Bakti
Judul Tugas Akhir : Komparasi Algoritma Regresi Linier Berganda dan *Decision Tree Regressor* Untuk Memprediksi Harga Mobil Bekas

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, Agustus 2022

Menyetujui,



UNIVERSITAS
MERCU BUANA

(Vina Ayumi S.Kom., M.Kom.)

Dosen Pembimbing

Mengetahui,



(Wawan Gunawan, S.Kom, MT)

Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)

Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kita panjatkan kehadiran Allah SWT atas berkah, rahmat dan hidayah-NYA yang senantiasa dilimpahkan kepada penulis, sehingga penulis bisa menyelesaikan tugas akhir penulis dengan judul “*Komparasi Algoritma Regresi Linier Berganda dan Decision Tree Regressor Untuk Memprediksi Harga Mobil Bekas*” sebagai syarat untuk menyelesaikan Program Sarjana (S1) pada Program Sarjana Fakultas Ilmu Komputer Jurusan Informatika Universitas Mercu Buana.

Penulis menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak baik secara moral maupun spiritual. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Kedua orang tua penulis yang telah mendukung penulis serta kasih sayangnya sehingga penulis bisa menyelesaikan tugas akhir ini.
2. Ibu Vina Ayumi S,Kom.,M.Kom selaku dosen pembimbing penulis yang telah banyak membantu dan memberikan saran kepada penulis dalam menyelesaikan tugas akhir ini.
3. Bapak Emil R. Kaburuan Ph.D selaku Kepala Program Studi Informatika yang telah memberikan arahan dan bimbingan selama masa perkuliahan.
4. Bapak Wawan Gunawan S.Kom.,MT selaku Koordinator Tugas Akhir Prodi Informatika yang memberikan arahan selama periode tugas akhir.
5. Ibu Harni Kusniyati, M.Kom sebagai dosen pembimbing akademik yang telah membimbing penulis selama perkuliahan hingga sampai pada tugas akhir penulis.
6. *Last but not least, I wanna thank me, I wanna thank me for believing in me, I wanna thank me for doing all this hard work, I wanna thank me for having no days off, I wanna thank me for never quitting.*

Apa yang penulis sadari adalah bahwa penyusunan tugas akhir ini masih jauh dari kata sempurna dan masih terdapat kekurangan bahkan kelemahan dalam penyusunan tugas akhir ini, walaupun begitu, penulis berusaha melakukan yang terbaik untuk menyelesaikan tugas akhir ini.

Akhir kata, penulis berharap jika ada kritik atau saran apapun yang sifatnya membangun bagi saya dengan senang hati saya akan menerimanya.

Jakarta, Juni 2022

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	i
LEMBAR PERNYATAAN ORISINALITAS	ii
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR...	iii
SURAT PERNYATAAN LUARAN TUGAS AKHIR	iv
LEMBAR PERSETUJUAN PENGUJI	v
LEMBAR PENGESAHAN	viii
ABSTRAK	ix
ABSTRACT	x
KATA PENGANTAR	xi
DAFTAR ISI	xii
NASKAH JURNAL	1
KERTAS KERJA	8
BAB 1. LITERATUR REVIEW	9
BAB 2. SOURCE CODE	16
BAB 3. DATASET	20
BAB 4. TAHAPAN EKSPERIMEN	25
BAB 5. HASIL SEMUA EKSPERIMEN	30
DAFTAR PUSTAKA	32
LAMPIRAN DOKUMEN HAKI	34
LAMPIRAN KORESPONDENSI	35

NASKAH JURNAL

JURNAL MEDIA INFORMATIKA BUDIDARMA

Volume 5, Nomor X, Bulan 2021, Page 999-999

ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)

Available Online at <https://ejournal.stmik-budidarma.ac.id/index.php/mib>

DOI 10.30865/mib.v5i1.2293



Komparasi Algoritma Regresi Linear Berganda dan Decision Tree Regressor Untuk Memprediksi Harga Mobil Bekas

Wisnu Nugroho Setya Bakti¹, Vina Ayumi²

Fakultas Ilmu Komputer, Informatika, Universitas Mercu Buana, Jakarta, Indonesia

Email: ¹41518010114@student.mercubuana.ac.id, ²vina.ayumi@mercubuana.ac.id

Email Penulis Korespondensi: vina.ayumi@mercubuana.ac.id

Abstrak– Pandemi covid-19 sangat berdampak besar bagi kehidupan masyarakat. Tidak hanya bagi masyarakat, namun pandemi ini juga berdampak bagi pelaku usaha yang memiliki showroom mobil. Disituasi seperti ini tidak sedikit masyarakat yang mengalami penurunan pendapatan. Akibatnya, semakin banyak masyarakat yang membutuhkan dana tambahan untuk memenuhi kebutuhan sehari-harinya. Beberapa diantara mereka mengharuskan untuk menjual aset berharga mereka salah satunya adalah mobil. Oleh karena itu penulis membuat suatu model untuk membantu masyarakat yang pengetahuannya masih awam dalam menentukan harga mobil dan juga penelitian ini dapat membantu pelaku usaha yang ingin mendirikan usaha showroom mobil. Riset ini memanfaatkan metode Regresi Linear Berganda serta Decision Tree Regressor selaku pembandingnya untuk memprediksi harga mobil bekas. Rata-rata akurasi yang didapat pada algoritma Regresi Linier Berganda adalah 73.9% sedangkan rata-rata pada algoritma Decision Tree Regressor adalah 93.3%.

Kata Kunci: Regresi, Regresi Linier Berganda, Decision Tree Regressor, Harga Mobil Bekas, Machine Learning

Abstract– The COVID-19 pandemic has had a huge impact on people's lives. Not only for the community, but this pandemic also has an impact on business actors who have car showrooms. In a situation like this, not a few people experience a decrease in income. As a result, more and more people need additional funds to meet their daily needs. Some of them require to sell their valuable assets, one of which is a car. Therefore, the author makes a model to help people whose knowledge is still lay in determining the price of a car and also this research can help business people who want to set up a car showroom business. This research uses the Multiple Linear Regression method and the Decision Tree Regressor as a comparison to predict the price of a used car. The average accuracy obtained in the Multiple Linear Regression algorithm is 73.9% while the average for the Decision Tree Regressor algorithm is 93.3%.

Keywords: Regression, Multiple Linear Regression, Decision Tree Regressor, Price Car Used, Machine Learning

1. PENDAHULUAN

Pandemic covid-19 sangat berdampak besar bagi kehidupan seseorang. Disituasi pandemic covid-19 ini tidak sedikit masyarakat yang mengalami penurunan pendapatan bahkan ada beberapa masyarakat yang terpaksa harus kehilangan pekerjaannya dan bahkan pula ada beberapa perusahaan yang harus gulung tikar. Karena hal itu, banyak sekali masyarakat yang membutuhkan dana tambahan untuk memenuhi kebutuhan sehari-harinya. Beberapa diantara mereka terpaksa menjual aset-aset mereka untuk memenuhi kebutuhan tersebut, salah satu aset yang mereka jual adalah mobil. Mereka terpaksa memangkas harga mobil yang dijual dari harga pasar untuk mendapatkan dana tambahan dengan cepat. Hal ini pun menyebabkan harga pasaran mobil bekas menitik tajam. Oleh karena itu penulis akan membuat model untuk memprediksi harga mobil bekas yang diharapkan model yang penulis buat dapat memulihkan harga pasaran kendaraan beroda empat ini dan juga penulis berharap model yang penulis buat ini akan menjadi acuan kepada masyarakat dalam menentukan harga pada mobil bekas.

Penelitian ini akan dimulai dari pengumpulan data pada suatu *web showroom* mobil atau iklan penjualan mobil. Parameter yang dibutuhkan untuk melengkapi data yang diperlukan yaitu jenis mobil, tahun pembuatan, transmisi, jarak yang ditempuh dan juga *engine size* pada mobil. Lalu data yang telah terkumpul akan diolah untuk membuat model prediksi dengan menggunakan algoritma Regresi Linier Berganda dan algoritma *Decision Tree Regressor* sebagai pembandingnya.

Regresi linier berganda ialah suatu metode analisis yang berupaya menarangkan suatu ikatan antara dua variabel ataupun lebih, khususnya variabel sebab akibat ataupun dapat disebut pula analisis regresi. Analisis ini sudah dikembangkan guna mempelajari suatu pola serta mengukur hubungan antar variabel [14].

Decision Tree Regressor ialah suatu algoritma yang gunanya untuk pengambilan keputusan. Algoritma ini akan mencari penyelesaian permasalahan memanfaatkan node yang silih berhubungan sehingga akan menciptakan struktur yang menyamai suatu pohon. Algoritma ini menggambarkan model prediksi pada sesuatu keputusan yang menggunakan struktur hirarki ataupun pohon. Tiap pohon akan memiliki cabang, yang nantinya cabang tersebut mewakili suatu atribut yang harus dipenuhi guna menuju cabang berikutnya sampai berakhir ataupun sampai terdapat cabang lagi [3].

Adapun beberapa penelitian yang relevan dalam memprediksi harga mobil bekas, yaitu salah satunya adalah penelitian yang dilaksanakan oleh Bambang Kriswantara dan teman temannya pada tahun 2021 dalam jurnal 'Prediksi Harga Mobil Bekas Dengan Machine Learning'. Penelitian ini menghasilkan *Mean Absolute Error(MAE)* sebesar 501232 serta menghasilkan tingkat akurasi sebesar 88% [12].

First Author, Copyright © 2021, MIB, Page 1

Submitted: dd/mm/yyyy; Accepted: dd/mm/yyyy; Published: dd/mm/yyyy

JURNAL MEDIA INFORMATIKA BUDIDARMA

Volume 5, Nomor X, Bulan 2021, Page 999-999

ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)

Available Online at <https://ejournal.stmik-budidarma.ac.id/index.php/mib>

DOI 10.30865/mib.v5i1.2293



Kemudian penelitian selanjutnya yang dilaksanakan oleh Nur Nafi'yah pada tahun 2015 dalam jurnal 'Penerapan Regresi Linear dalam Memprediksi Harga Jual Mobil Bekas'. Riset ini menggunakan metode Regresi Linear guna memprediksi harga jual mobil bekas Toyota Avanza, Toyota Kijang Innova dan Daihatsu Xenia berdasarkan kategori dan tahun pembuatan, harga beli baru serta kisaran harga jual mobil bekas. Dengan menggunakan metode tersebut menghasilkan tingkat akurasi sebesar 80% [6].

Penelitian selanjutnya yang dilaksanakan oleh Suria Sandi Winarto dan Sutojo dalam jurnal 'Menentukan Harga Mobil Bekas dengan Menggunakan Metode *Fuzzy Mamdani* dan metode Jaringan Syaraf Tiruan. Dari pengujian ini didapatkan hasil perhitungan rata-rata persentase kesalahan absolut (Mean Percentage Absolute Error) dari metode Fuzzy Mamdani menghasilkan nilai kesalahan sebesar 2.089902% serta tingkatan akurasi dari metode ini yakni 97.910098%. Sementara itu dengan memakai metode Jaringan Syaraf Tiruan menghasilkan nilai MAPE sebesar 2.06086% serta tingkatan akurasi pada metode ini menciptakan nilai sebesar 97.93914% [10].

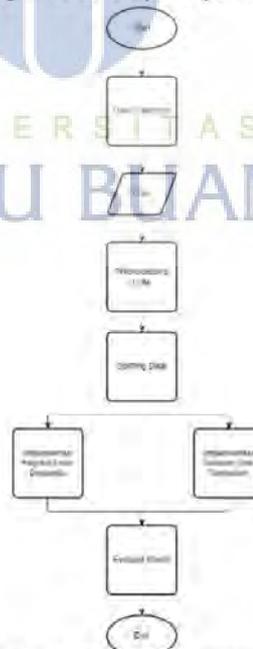
Selanjutnya adapun riset yang dilaksanakan oleh Fachrudin Pakaja dan Agus Naba pada tahun 2015 dalam jurnal 'Peramalan Penjualan Mobil Menggunakan Jaringan Syaraf Tiruan dan *Certainty Factor*'. Pada penelitian ini menggunakan data penjualan mobil Honda tahun 2015 dengan variabel input penjualan daerah 30000 unit, penjualan dealer 25000 unit dan penjualan tunai 25000. Dengan menggunakan nilai *Certainty Factor (CF)* sebesar 0.5 menghasilkan prediksi penjualan sebanyak 29579 unit dengan nilai kesalahan sebesar 4.205% [7].

Kemudian penelitian selanjutnya yang dilaksanakan oleh Nabarun Pal dan kawan kawan dalam jurnalnya yang berjudul 'How Much is my car worth? A methodology for predicting used cars' prices using random forest' yang diterbitkan pada tahun 2019. Pada penelitian ini menggunakan dataset dari kaggle. Menghasilkan 83.62% untuk data ujinya dan 95% untuk data yang dilatih [8].

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tujuan dibuatnya penelitian ini ialah untuk membuat sebuah model yang mana untuk memprediksi harga mobil bekas. Berikut dibawah ini adalah gambar dari alu proses penelitian tersebut.



Gambar 1. Flowchart proses penelitian



2.1.1. Data Collection

Tahap pertama pada penelitian ini adalah *data collection*. Tahap ini diperlukan untuk mengumpulkan data dari *website showroom* mobil dan situs iklan mobil seperti *carro.id* dan *carsome*. Penulis menggunakan metode *web scraping* untuk mengumpulkan data tersebut. Data yang diterapkan pada riset ini ialah data spesifikasi mobil tahun pembuatan 2010 sampai dengan tahun pembuatan 2021, total data yang telah terkumpul yaitu sekitar 4999 data.

Tahun	Brand	Model	EngineSize	Jarak	Transmisi	Harga
2019	Toyota	Land Cruiser	4.5	17000	Automatic	2395000000
2018	Toyota	Land Cruiser	4.5	25000.0	Automatic	2275000000
2018	Toyota	Land Cruiser	4.5	25000.0	Automatic	2025000000
2015	Toyota	Land Cruiser	4.5	67200.0	Automatic	1595000000
2015	Toyota	Land Cruiser	4.5	67200.0	Automatic	1595000000

Tabel 1. Contoh dataset yang telah dikumpulkan

2.1.2. Preprocessing Data

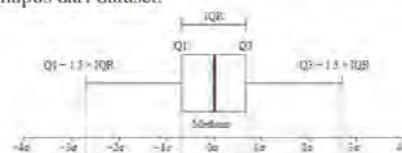
2.1.2.1. Data Cleaning

Pada proses *data cleaning*, penulis memeriksa dataset apakah terdapat data yang kosong atau tidak. Hasilnya terdapat 40 data yang kosong pada variabel jarak. Kemudian penelitian menghitung rata-rata dari keseluruhan data pada variabel jarak yang nantinya hasil rata-rata tersebut akan diisi kedalam variabel yang kosong.

Tahun	0
Brand	0
Model	0
EngineSize	0
Jarak	40
Transmisi	0
Harga	0

Tabel 2. Variabel yang terdapat *Missing Values*

Selanjutnya penulis memeriksa dataset apakah terdapat data *outliers* atau tidak. Penulis menggunakan visualisasi *boxplot* untuk melihat data *outliers* dari setiap variabel. Lalu untuk menganganinya penulis menggunakan metode IQR, yang dimana metode ini akan membuat sebuah jarak batas minimal dan batas maksimal dari suatu variabel. Kemudian, jika terdapat sebuah data yang diluar dari batas minimal atau batas maksimal yang telah ditentukan maka data tersebut akan dihapus dari dataset.



Gambar 2. Ilustrasi *Interquartile Range*

2.1.2.2. Data Transformation

Pada proses *data transformation*, penulis mengubah variabel yang memiliki tipe data *string* menjadi tipe data *integer*. Perubahan tipe data variabel dilakukan agar ketika proses penghitungan variabel tersebut dapat diproses. Penerapan tahap ini dilakukan pada variabel Brand, Model dan Transmisi.

Tahun	Brand	Model	EngineSize	Jarak	Transmisi	Harga
2018	1	12	1.5	40000	0	400000000
2018	1	12	1.5	40000	0	400000000
2018	2	8	2.5	11200	0	399000000
2016	1	12	1.5	71400	0	399000000
2018	2	37	2.0	20000	0	399000000

Tabel 3. Perubahan tipe data yang memiliki variabel *string* menjadi *integer*

Selanjutnya penulis melakukan normalisasi data pada penelitian ini. Normalisasi data dilakukan agar nilai pada suatu variabel memiliki rentang nilai yang sama, yang dimana nilai tersebut tidak ada yang terlalu besar maupun tidak ada nilai yang terlalu kecil, oleh karena itu membuat analisis menjadi lebih mudah. Metode ini dilakukan untuk menghindari hasil



yang tidak seimbang yang dikarenakan nilai pada suatu variabel yang lebih besar daripada variabel yang lain. Penerapan normalisasi data dilakukan untuk seluruh variabel pada dataset.

Tahun	Brand	Model	EngineSize	Jarak	Transmisi	Harga
0.636364	1.0	0.0	0.0	0.549955	0	0.213483
0.636364	1.0	0.0	0.0	0.320453	1	0.243127
0.636364	1.0	0.0	0.0	0.254125	0	0.283427
0.636364	1.0	0.0	0.0	0.105539	1	0.255453
0.636364	1.0	0.0	0.0	0.288900	0	0.290637

Tabel 4. Tampilan *dataset* setelah melalui proses normalisasi data

2.1.3. Splitting Data

Setelah preprocessing pada dataset selesai. Tahap selanjutnya ialah memisahkan dataset menjadi data training untuk melatih dataset dan data test untuk pengujian dataset. Untuk menerapkan hal tersebut penulis menggunakan metode split validation. Split validation merupakan sebuah metode membagi data training dan data testing dengan membaginya secara acak. Peneliti membagi 5 skenario eksperimen pada pembagian data. Berikut dibawah ini adalah scenario tersebut.

Skenario	Data Training	Data Testing
1	90%	10%
2	80%	20%
3	70%	30%
4	60%	40%
5	50%	50%

Tabel 5. Distribusi *Data Training* dan *Data Testing*

2.1.4. Implementasi Algoritma Regresi Linear Berganda dan Decision Tree Regressor

Setelah data telah diproses dan mendapatkan data yang bersih, lalu tahap selanjutnya adalah data akan diimplementasikan menggunakan algoritma Regresi Linier Berganda dan Decision Tree Regressor. Untuk implementasinya peneliti membagi menjadi dua variabel yaitu variabel independen dan variabel dependen. Variabel yang digunakan pada variabel independen adalah Tahun, Brand, Jarak, EngineSize Model dan Transmisi, lalu variabel yang digunakan untuk variabel dependen adalah variabel Harga.

2.1.4.1. Regresi Linear Berganda

Merupakan suatu metode analisis yang berupaya menarangkan suatu ikatan antara dua variabel ataupun lebih, khususnya variabel sebab akibat ataupun dapat disebut pula analisis regresi. Analisis ini sudah dikembangkan guna mempelajari suatu pola serta mengukur hubungan antar variabel. Proses yang dilakukan pada algoritma Regresi Linier Berganda adalah menghitung data yang sudah disediakan untuk mencari nilai konstanta dan juga mencari nilai koefisien pada setiap variabel sehingga akan menghasilkan nilai persamaan Regresi Linier Berganda.

2.1.4.2. Decision Tree Regressor

Proses yang dilakukan pada algoritma Decision Tree Regressor adalah algoritma ini akan mencari solusi permasalahan menggunakan node yang saling berhubungan sehingga akan menghasilkan struktur yang menyerupai sebuah pohon. Algoritma ini menggambarkan model prediksi pada sesuatu keputusan yang menggunakan struktur hirarki ataupun pohon. Tiap pohon akan memiliki cabang, yang nantinya cabang tersebut mewakili suatu atribut yang harus dipenuhi guna menuju cabang berikutnya sampai berakhir ataupun sampai terdapat cabang lagi.

2.1.5. Evaluasi Model

Tahap terakhir yaitu evaluasi model pada penelitian untuk mengukur keakuratan pada model dengan menggunakan parameter, MSE (*Mean Squared Error*), MAPE (*Mean Absolute Percentage Error*) dan *R2 Score*.

R2 Score

$$R^2 = 1 - \frac{SSE}{SST} \quad (1)$$

Evaluasi *R2 Score* digunakan untuk mengukur akurasi yang mewakili proporsi sebuah varians pada variabel dependen yang dapat diprediksi oleh variabel independen dan juga berfungsi untuk mengukur sejauh mana korelasi antar variabel satu dengan variabel yang lain.

MAPE

$$MAPE = \sum_{i=1}^n \left[\frac{|y_i - \hat{y}_i|}{\hat{y}_i} \right] \times 100\% \quad (2)$$

Mean Absolute Percentage, merupakan sebuah evaluasi pada suatu model yang digunakan untuk memperkirakan kesalahan. Untuk pengukuran pada MAPE ini menggunakan persentase sehingga ketika melakukan evaluasi pada model membuatnya lebih mudah untuk dipahami. MAPE akan



berfungsi dengan baik jika ada nilai yang ekstrem pada data. MAPE juga berfungsi untuk mengukur kerugian dalam analisis regresi dan evaluasi pada suatu model. Berikut dibawah ini merupakan tabel interpretasi pada MAPE.

Range MAPE	Arti Nilai
< 10%	Model Prediksi Sangat Baik
10 - 20%	Model Prediksi Baik
20 - 50 %	Model Prediksi Layak
>50%	Model Prediksi Buruk

Tabel 6. Representasi Persentase nilai MAPE

MSE

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2 \quad (3)$$

Evaluasi digunakan untuk mengukur rata-rata kuadrat pada kesalahan suatu algoritma. Jika nilai rata-rata tersebut semakin besar maka semakin besar kesalahan pada nilai rata-rata tersebut.

3. HASIL DAN PEMBAHASAN

Berikut dibawah ini merupakan eksperimen yang dilaksanakan oleh peneliti

Decision Tree Regressor

Experiment	Data Test	Data Train	R2	MSE	MAPE
1	10%	90%	93.2%	0.0022	9.2%
2	20%	80%	93.3%	0.0021	9.1%
3	30%	70%	92.7%	0.0022	9.6%
4	40%	60%	92%	0.0026	9.8%
5	50%	50%	91.9%	0.0026	10.1%

Tabel 7. Hasil Perhitungan dari Model *Decision Tree Regressor*

Pada eksperimen yang telah dilakukan di algoritma *Decision Tree Regressor*, dapat dilihat pada tabel diatas bahwa pada pembagian data test sebesar 20% dan data train sebesar 80% menghasilkan nilai *R2 Score* sebesar 93.3% dengan nilai MSE sebesar 0.0021 dan nilai MAPE sebesar 9.1%. Dari hasil percobaan yang dilaksanakan pada algoritma *Decision Tree Regressor* pembagian distribusi data sangat penting untuk mencapai nilai akurasi yang baik. Maka dari itu perlu percobaan dengan beberapa variasi distribusi data untuk mencapai nilai akurasi yang baik.

Regresi Linier Berganda

Experiment	Data Test	Data Train	R2	MSE	MAPE
1	10%	90%	73.9%	0.0088	19.5%
2	20%	80%	73.1%	0.0084	19.7%
3	30%	70%	73%	0.0083	19.9%
4	40%	60%	73.4%	0.0088	19.5%
5	50%	50%	73.5%	0.0088	19.5%

Tabel 8. Hasil Perhitungan dari Model Regresi Linear Berganda

Sedangkan dengan menggunakan algoritma Regresi Linier Berganda dengan distribusi data test sebesar 10% dan data train sebesar 90% menghasilkan nilai *R2 Score* sebesar 73.9%, dengan nilai MSE 0.0088 dan nilai MAPE sebesar 19.5%. Dari hasil percobaan yang telah dilakukan pada algoritma Regresi Linier Berganda pembagian distribusi data sangat penting untuk dilakukan, karena hal tersebut akan menentukan nilai akurasinya. Maka dari itu perlu percobaan dengan beberapa variasi distribusi data untuk mencapai nilai akurasi yang baik.

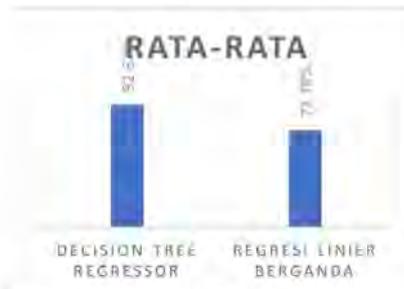
JURNAL MEDIA INFORMATIKA BUDIDARMA

Volume 5, Nomor X, Bulan 2021, Page 999-999

ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)

Available Online at <https://ejurnal.stmik-budidarma.ac.id/index.php/mib>

DOI 10.30865/mib.v5i1.2293



Dapat dilihat bahwa nilai rata-rata akurasi pada model *decision tree regressor* adalah 92,62% sedangkan nilai rata-rata akurasi regresi linier berganda adalah 73,3%. dapat disimpulkan bahwa Decision Tree Regressor lebih baik daripada Regresi Linier Berganda.

4. KESIMPULAN

Berdasarkan penelitian yang dilaksanakan dapat diperoleh tingkat akurasi menggunakan algoritma Decision Tree Regressor dengan distribusi data *test* 20% dan data *train* 80% menghasilkan 93.3% pada R2 score, kemudian nilai MSE menghasilkan nilai sebesar 0.0021 dan nilai MAPE sebesar 9.1%, sedangkan pada penelitian yang menggunakan algoritma Regresi Linier Berganda dengan distribusi data *test* sebesar 10% dan data *train* sebesar 90% menghasilkan nilai R2 Score sebesar 73.9%, dengan nilai MSE 0.0088 dan nilai MAPE sebesar 19.5%..

Dapat disimpulkan bahwa hasil penelitian Decision Tree Regressor lebih baik daripada Regresi Linier Berganda dikarenakan tingkat akurasinya yang lebih tinggi. Namun jika dilihat melalui interpretasi nilai MAPE algoritma Regresi Linier Berganda dikategorikan 'baik' digunakan untuk dimasa yang akan datang walaupun akurasinya hanya sebesar 73.9%. Sedangkan pada Decision Tree Regressor selain tingkat akurasinya yang tinggi algoritma ini dikategorikan 'sangat baik' untuk digunakan di masa yang akan datang jika merujuk dari nilai MAPE.

UNIVERSITAS

MERCUBUANA

REFERENCES

- [1] E. Dewi, S. Mulyani, F. Muliady, D. Ramadhan, A. Ariyanto, and D. Ramdani, "Estimasi Harga Jual Mobil Bekas Menggunakan Metode Regresi Linier Berganda," *e-Jurnal JUSITI (Jurnal Sist. Inf. dan Teknol. Informasi)*, vol. 9, no. 1, pp. 1–8, 2020, doi: 10.36774/jusiti.v9i1.649.
- [2] A. Ahmad, "Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning."
- [3] D. Sartika and D. I. Sensuse, "Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian," *J. Tek. Inform. Dan Sist. Inf.*, vol. 1, no. 2, pp. 151–161, 2017, [Online]. Available: <https://doi.org/10.35957/jatisi.v3i2.78>.
- [4] J. Gajera, Prashant Gondaliya, Akshay Kavathiya, "Old Car Price Prediction With Machine Learning," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 3, no. 3, pp. 284–290, 2021.
- [5] M. Mehdi Karakoç, G. Çelik, and A. Varol, "Car Price Prediction Using An Artificial Neural Network," *East. Anatol. J. Sci.*, vol. VI, no. II, pp. 44–48, 2020.
- [6] N. Nafi'iyah, "Penerapan regresi linear dalam memprediksi harga jual mobil bekas," pp. 1–5, 2015.
- [7] F. Pakaja and A. Naba, "Peramalan Penjualan Mobil Menggunakan Jaringan Syaraf Tiruan dan Certainty Factor," *Neural Networks*, vol. 6, no. 1, pp. 23–28, 2015.
- [8] N. Pal, P. Arora, P. Kohli, D. Sundararaman, and S. S. Palakurthy, "How Much is my car worth? A methodology for predicting used cars' prices using random forest," *Adv. Intell. Syst. Comput.*, vol. 886, pp. 413–422, 2019, doi: 10.1007/978-3-030-03402-3_28.
- [9] K. Samruddhi and R. Ashok Kumar, "Used Car Price Prediction using K-Nearest Neighbor Based Model," *Int. J. Innov. Res. Appl. Sci. Eng.*, vol. 4, no. 3, pp. 686–689, 2020, doi: 10.29027/ijirase.v4.i3.2020.686-689.
- [10] S. S. Winarto and T. Sutojo, "Menentukan Harga Mobil Bekas Dengan Menggunakan Metode Fuzzy Mamdani Dan Metode Jaringan Syaraf Tiruan," *Techno.COM*, vol. 11, no. 3, pp. 134–141, 2012.
- [11] Yennimar, Kelvin, Suwandi, and Amir, "Comparison Analysis of SVM Algorithm with Linear Regression in

First Author, Copyright © 2021, MIB, Page 6

Submitted: dd/mm/yyyy; Accepted: dd/mm/yyyy; Published: dd/mm/yyyy

JURNAL MEDIA INFORMATIKA BUDIDARMA

Volume 5, Nomor X, Bulan 2021, Page 999-999

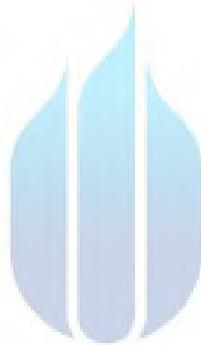
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)

Available Online at <https://ejournal.stmik-budidarma.ac.id/index.php/mib>

DOI 10.30865/mib.v5i1.2293



- Predicting used Car Prices," vol. 5, no. 36, pp. 2720–2728, 2022.
- [12] B. Kriswantara, Kurniawati, and F. H. Pardede, "PREDIKSI HARGA MOBIL BEKAS DENGAN MACHINE LEARNING," *J. Ilm. Indones.*, vol. 6, no. 5, p. 6, 2021, doi: <http://dx.doi.org/10.36418/syntax-literat.v6i5.2716>.
- [13] S. Glen, "MAPE (Mean Absolute Percentage Error)." <https://www.statisticshowto.com/mean-absolute-percentage-error-mape/>.
- [14] Agus Tri Basuki and Nano Prawoto, *Analisis Regresi Dalam Penelitian Ekonomi dan Bisnis*. Raja Grafindo Persada, 2015.
- [15] S. Pudaruth, "Predicting the Price of Used Cars using Machine Learning Techniques (Information & Computation Technology, 2014)," *Int. J. Inf. Comput. Technol.*, vol. 4, no. 7, pp. 753–764, 2014.
- [16] K. Noor and S. Jan, "Vehicle Price Prediction System using Machine Learning Techniques," *Int. J. Comput. Appl.*, vol. 167, no. 9, pp. 27–31, 2017, doi: 10.5120/ijca2017914373.
- [17] D. . Das Adhikary, R. Sahu, and S. Pragyna Panda, "Prediction of Used Car Prices Using Machine Learning," *Biol. Inspired Tech. Many Criteria Decis. Mak.*, vol. 271, pp. 131–140, 2022, doi: https://doi.org/10.1007/978-981-16-8739-6_11.



UNIVERSITAS
MERCU BUANA

First Author, Copyright © 2021, MIB, Page 7

Submitted: dd/mm/yyyy; Accepted: dd/mm/yyyy; Published: dd/mm/yyyy

Universitas Mercu Buana

<http://digilib.mercubuana.ac.id/>

KERTAS KERJA

Ringkasan

Pada kertas kerja ini merupakan sebuah material kelengkapan dari artikel jurnal yang berjudul “Komparasi Algoritma Regresi Linier Berganda dan Decision Tree Regressor Untuk Memprediksi Harga Mobil Bekas”. Lalu pada kertas kerja ini berisi semua materi dari hasil penelitian Tugas Akhir penulis yang mungkin tidak disertakan dalam artikel jurnal ilmiah. Adapun pembahasan pada kertas kerja ini yaitu : literature review, dataset yang digunakan, source code dan hasil eksperimen secara keseluruhan.

