

**PREDIKSI HARGA PENJUALAN MOBIL BEKAS DENGAN
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS
DAN LINEAR REGRESSION**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

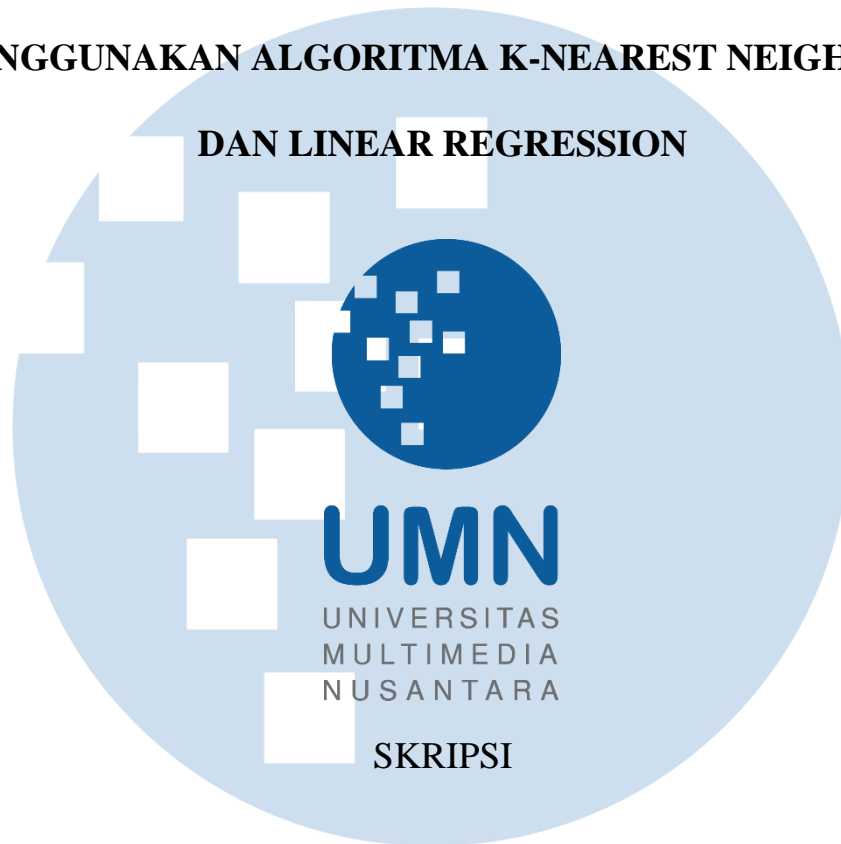
Vincentius Juan Antonio

0000043584

**PROGRAM STUDI SISEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2024

**PREDIKSI HARGA PENJUALAN MOBIL BEKAS DENGAN
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS
DAN LINEAR REGRESSION**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer (S.Kom)

Vincentius Juan Antonio

00000043584

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS TEKNIK DAN INFORMATIKA

UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2024

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Vincentius Juan Antonio

Nomor Induk Mahasiswa : 00000043584

Program studi : Sistem Informasi

Skripsi dengan judul:

PREDIKSI HARGA PENJUALAN MOBIL BEKAS DENGAN MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBORS DAN LINEAR REGRESSION merupakan hasil karya saya sendiri bukan plagiat dari karya ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan skripsi maupun dalam penulisan laporan skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas Akhir yang telah saya tempuh.

Tangerang, 23 Mei 2024

UMIN



Vincentius Juan Antonio

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PERSETUJUAN

Skripsi dengan judul

**PREDIKSI HARGA PENJUALAN MOBIL BEKAS DENGAN
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS DAN LINEAR
REGRESSION**

Oleh

Nama : Vincentius Juan Antonio
NIM : 00000043584
Program Studi : Sistem Informasi
Fakultas : Teknik Informatika

Telah disetujui untuk diajukan pada
Sidang Ujian Skripsi Universitas Multimedia Nusantara

Tangerang, 15 Mei 2024

Pembimbing



Ririn Ikana Desanti, S.Kom., M.Kom
313058001

Ketua Program Studi Sistem Informasi



Ririn Ikana Desanti, S.Kom., M.Kom

HALAMAN PENGESAHAN

HALAMAN PENGESAHAN

Skripsi dengan judul

**PREDIKSI HARGA PENJUALAN MOBIL BEKAS DENGAN
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS DAN LINEAR
REGRESSION**

Oleh

Nama : Vincentius Juan Antonio
NIM : 0000043584
Program Studi : Sistem Informasi
Fakultas : Teknik Informatika

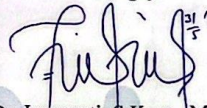
Telah diujikan pada hari Rabu, 22 Mei
Pukul 13.00 s.d 15.00 dan dinyatakan
LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang


Monika Evelin Johan, S.Kom., M.M.S.I.
0327059501

Penguji


Dr. Irmawati, S.Kom., M.M.S.I.
0805097703

Pembimbing


Ririn Ikana Desanti, S.Kom., M.Kom
313058001
Ketua Program Studi Sistem Informasi


Ririn Ikana Desanti, S.Kom., M.Kom

iv

Prediksi Harga Penjualan Mobil Bekas... Vincentius Juan Antonio, Universitas Multimedia
Nusantara

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : Vincentius Juan Antonio
NIM : 00000043584
Program Studi : Sistem Informasi
Jenjang : S1
Judul Karya Ilmiah : Prediksi Harga Penjualan Mobil Bekas Dengan Menggunakan Algoritma K-Nearest Neighbors dan Linear Regression

Menyatakan dengan sesungguhnya bahwa saya bersedia (**pilih salah satu**):

- Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial. Saya tidak akan mencabut kembali izin yang telah saya berikan dengan alasan apapun.
- Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: Dalam proses pengajuan penerbitan ke dalam jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*)*.

Tangerang, 15 Mei 2024

Yang menyatakan,



(Vincentius Juan Antonio)

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A

KATA PENGANTAR

Puji syukur peneliti panjatkan kepada Tuhan Yang Maha Esa karena berkat dan kasih karunia-Nya, peneliti dapat menyelesaikan skripsi dengan baik yang berjudul “Prediksi Harga Penjualan Mobil Bekas Dengan Menggunakan Algoritma K-Nearest Neighbors dan Linear Regression” dilakukan untuk memenuhi syarat mendapatkan gelar Sarjana Komputer pada jurusan Sistem Informasi, Teknik dan Informatika Universitas Multimedia Nusantara.

Peneliti juga menyadari bahwa adanya bantuan serta bimbingan dari berbagai pihak dari awal pembuatan skripsi ini sampai saat skripsi ini telah selesai dibuat. Oleh karena itu peneliti mengucapkan terima kasih kepada:

1. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng Niki Prastomo, S.T., M.Sc, selaku Dekan Fakultas Teknik Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom, selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Ririn Ikana Desanti, S.Kom., M.Kom, sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya tugas akhir ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan tugas akhir ini.
6. Rafael Stefan, Erland, Made, Devira, Darren, Miguel, yang telah memberikan dukungan dan semangat dalam menyelesaikan tugas akhir ini.
7. Teman-teman seperjuangan dari awal sampai penulisan tugas akhir Yeremia Steven, Rico Tanuwijaya, Marcellius Cahyadi, Gregorius Dinov dan Ferrozio Wijaya, yang telah memberikan masukan serta bantuan dalam penulisan tugas akhir ini.

Akhir kata, dalam menyelesaikan penelitian ini, peneliti memiliki harapan yang kuat bahwa hasil penelitian yang telah dikembangkan dalam bentuk skripsi ini dapat memberikan kontribusi berarti dan bermanfaat kepada para pembaca

Tangerang, 15 Mei 2024



(Vincentius Juan Antonio)



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

PREDIKSI HARGA PENJUALAN MOBIL BEKAS DENGAN MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS DAN LINEAR REGRESSION

Vincentius Juan Antonio

ABSTRAK

Dalam era ekonomi yang semakin baik dan kebutuhan masyarakat yang meningkat, kepemilikan mobil telah menjadi suatu kebutuhan yang tidak terpisahkan. Hal ini menciptakan peluang besar bagi bisnis-bisnis yang bergerak di sektor penjualan mobil, baik baru maupun bekas. Namun terdapat tantangan yang menjadi hambatan serius bagi pebisnis dalam mengambil keputusan harga yang tepat, yang pada akhirnya dapat mempengaruhi pertumbuhan dan profitabilitas bisnis.

Masalah utama yang dihadapi oleh bisnis penjualan mobil bekas adalah kesulitan dalam menentukan harga yang sesuai dengan kondisi pasar. Harga yang ditetapkan sering kali tidak mencerminkan nilai aktual dari mobil bekas tersebut, yang dapat menyebabkan kerugian bagi pebisnis. Selain itu, penentuan harga yang tidak tepat juga dapat mempengaruhi citra merek dan kepercayaan konsumen. Oleh karena itu, diperlukan solusi yang efektif untuk membantu dalam menentukan harga penjualan mobil bekas yang lebih akurat dan kompetitif.

Penelitian ini mengusulkan penggunaan teknologi machine learning, dengan algoritma K-Nearest Neighbor (KNN) dan Linear Regression, untuk membuat prediksi harga penjualan mobil bekas yang lebih akurat. Metode ini dipilih karena kemampuannya dalam menangani berbagai jenis variabel, termasuk kategorikal dan numerik, yang umumnya ditemukan dalam dataset penjualan mobil bekas. Hasil penelitian menunjukkan bahwa penggunaan algoritma Linear Regression menghasilkan nilai MAE 5.4, MSE 760279536.33 dan Root Mean Squared Error (RMSE) sebesar 7.9, sedangkan algoritma K-Nearest Neighbor menghasilkan nilai MAE 10.5, MSE 3125496584.79 dan RMSE sebesar 16.2. Kedua algoritma ini menunjukkan potensi dalam meningkatkan akurasi prediksi harga penjualan mobil bekas, sehingga dapat membantu pebisnis dalam pengambilan keputusan yang lebih tepat dan efisien.

Kata kunci: *K-Nearest Neighbor, Linear Regression, Machine Learning, Mobil bekas.*

PREDICTION OF USED CAR SALES PRICES USING THE K-NEAREST NEIGHBORS AND LINEAR REGRESSION ALGORITHM

(Vincentius Juan Antonio)

ABSTRACT

In the era of a better economy and the increasing needs of society, car ownership has become an inseparable necessity. This creates great opportunities for businesses engaged in the car sales sector, both new and used. However, there are challenges that become serious obstacles for businesses in making the right pricing decisions, which in turn can affect business growth and profitability.

The main problem faced by used car sales businesses is the difficulty in determining prices that are in line with market conditions. The price set often does not reflect the actual value of the used car, which can lead to losses for the business. In addition, inappropriate pricing can also affect brand image and consumer trust. Therefore, an effective solution is needed to assist in determining more accurate and competitive used car sales prices.

This research proposes the use of machine learning technology, with K-Nearest Neighbor (KNN) and Linear Regression algorithms, to make more accurate used car sales price predictions. This method was chosen due to its ability to handle various types of variables, including categorical and numerical, which are commonly found in used car sales datasets. The results showed that the use of Linear Regression algorithm resulted in MAE value of 5.4, MSE 760279536.33 and Root Mean Squared Error (RMSE) of 7.9, while K-Nearest Neighbor algorithm resulted in MAE value of 10.5, MSE 3125496584.79 and RMSE of 16.2. Both algorithms show potential in improving the accuracy of used car sales price predictions, so that they can help business people make more informed and efficient decisions.

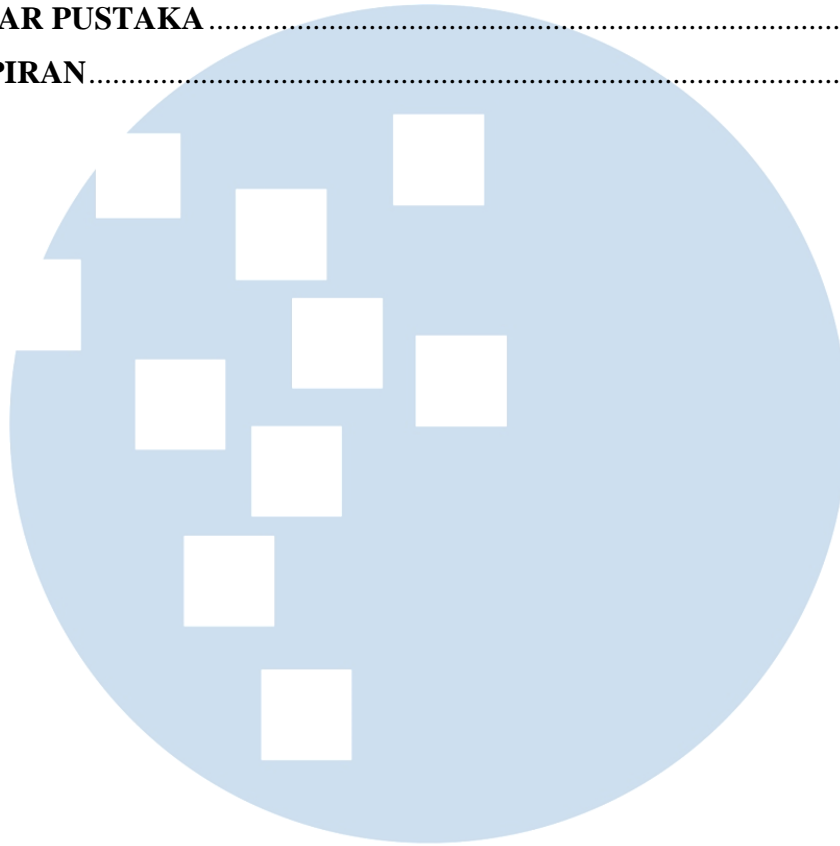
Keywords: *K-Nearest Neighbor, Linear Regression, Machine Learning, Used Car*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH MAHASISWA	v
KATA PENGANTAR	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	5
1.4 Tujuan dan Manfaat Penelitian	5
1.4.1 Tujuan Penelitian	5
1.4.2 Manfaat Penelitian	5
1.5 Sistematika Penulisan	5
BAB II LANDASAN TEORI	7
2.1 Penelitian Terdahulu	7
2.2 Teori yang digunakan	13
2.2.1 Prediksi	13
2.2.2 Showroom	14
2.2.3 Mobil Bekas	14
2.3 Algoritma	15
2.3.1 K-Nearest Neighbor	15
2.3.2 Linear Regression	16

2.4	Tools dan Teknik	17
2.4.1	Python	17
2.4.2	Jupyter	18
2.4.3	Scraping	18
2.4.4	Octoparse	19
2.4.5	Streamlit	19
2.4.6	Cross-Industry Standard Process for Data Mining (CRISP-DM) 19	
BAB III METODOLOGI PENELITIAN		21
3.1	Gambaran Umum Objek Penelitian	21
3.2	Metode Penelitian	23
3.2.1	Alur Penelitian	24
3.3	Teknik Pengumpulan Data	29
3.4	Teknik Analisis Data	30
BAB IV ANALISIS DAN HASIL PENELITIAN		33
4.1	Menentukan Topik Penelitian	33
4.2	Pengumpulan Data	33
4.3	Business understanding	33
4.4	Data Understanding	34
4.2.1	Collect data	35
4.2.2	Analisa Data	39
4.2.3	Pemilihan Variabel	45
4.5	Data Preperation	46
4.3.1	Import dataset	46
4.3.2	Clean Data	48
4.3.3	Label Encoder	51
4.3.4	Data Split	54
4.6	Evaluation	57
4.7	Hasil dan Pembahasan	61
4.8	Deployment	62
BAB V SIMPULAN DAN SARAN		67
5.1	Simpulan	67

5.2 Saran	67
DAFTAR PUSTAKA	69
LAMPIRAN	72



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

Table 2. 1 Penelitian Terdahulu	7
Table 3. 1 Perbandingan Algoritma	28
Table 3. 2 Perbandingan Program	31
Table 4. 1 Hasil perbandingan antar model	58
Table 4. 2 Perbandingan Linear Regression dan KNN	61



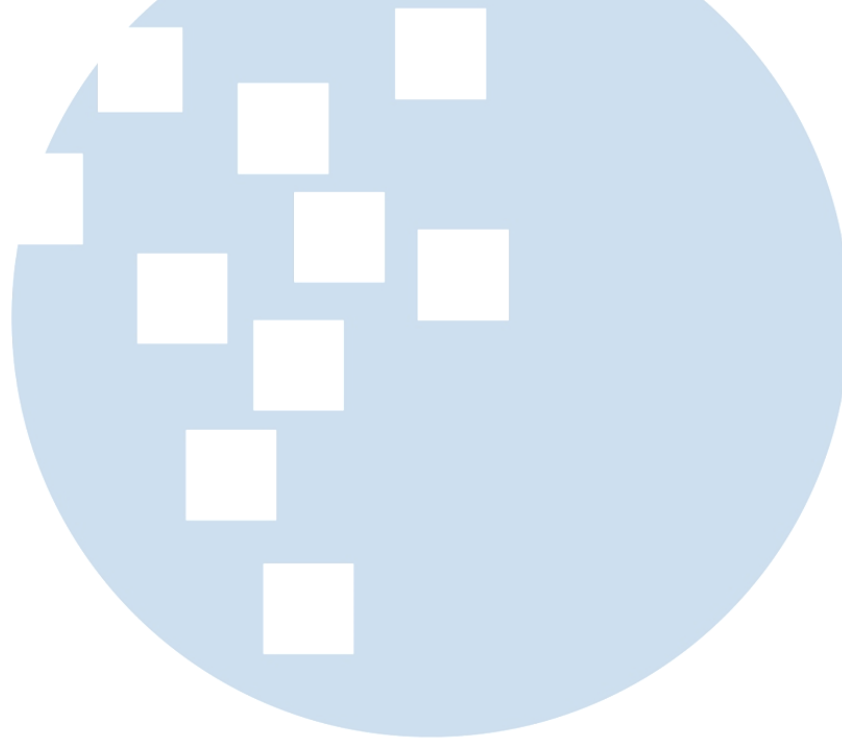
UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR GAMBAR

Gambar 1. 1 Data Penggunaan Kendaraan	1
Gambar 3. 1 Visualisasi Dataset Mobil Bekas	21
Gambar 3. 2 Alur Penelitian.....	24
Gambar 3. 3 Diagram data understanding.....	26
Gambar 3. 4 Diagram Preperation	27
Gambar 4. 1 Halaman Kendaraan Cintamobil	35
Gambar 4. 2 Tampilan halaman awal Octoparse	35
Gambar 4. 3 URL cintamobil.....	36
Gambar 4. 4 Tampilan setelah masuk URL	37
Gambar 4. 5 Data Preview	37
Gambar 4. 6 Data digunakan.....	38
Gambar 4. 7 Visualisasi banyak jenis kendaraan.....	39
Gambar 4. 8 Visualisasi warna pada jenis kendaraan	40
Gambar 4. 9 Visualisasi tahun Kendaraan	41
Gambar 4. 10 Visualisasi rata- rata kilometer.....	42
Gambar 4. 11 Visualisasi rata - rata harga kendaraan.....	43
Gambar 4. 12 Visualisasi jumlah transmisi kendaraan	43
Gambar 4. 13 Visualisasi Engine Size pada Model Kendaraan.....	44
Gambar 4. 14 Visualisasi Jumlah kendaraan berdasarkan model	45
Gambar 4. 15 Hasil dataframe	46
Gambar 4. 16 Deskriptif dataframe.....	47
Gambar 4. 17 Nilai unik.....	48
Gambar 4. 18 Data drop	48
Gambar 4. 19 Melihat Data NULL	49
Gambar 4. 20 Menghapus nilai NULL	49
Gambar 4. 21 Data Duplikat	50
Gambar 4. 22 Drop duplikat.....	51
Gambar 4. 23 Tampilan Info.....	51
Gambar 4. 24 Encoder.....	52
Gambar 4. 25 Sample	53
Gambar 4. 26 Persiapan data.....	54
Gambar 4. 27 Split Data.....	55
Gambar 4. 28 Model Regresi Linear.....	56
Gambar 4. 29 Model K-Nearest Neighbor.....	56
Gambar 4. 30 Metrik Evaluasi Regresi Linear.....	57
Gambar 4. 31 Metrik Evaluasi KNN.....	57
Gambar 4. 32 Input Data.....	59
Gambar 4. 33 Hasil Enkripsi.....	59
Gambar 4. 34 Prediksi Harga LR.....	60
Gambar 4. 35 Prediksi harga KNN	60
Gambar 4. 36 Tampilan Interface	63

Gambar 4. 37 Button Prediksi.....	64
Gambar 4. 38 Input Interface	65
Gambar 4. 39 Hasil Linear Regression	65
Gambar 4. 40 Hasil K-Nearest Neighbor.....	65

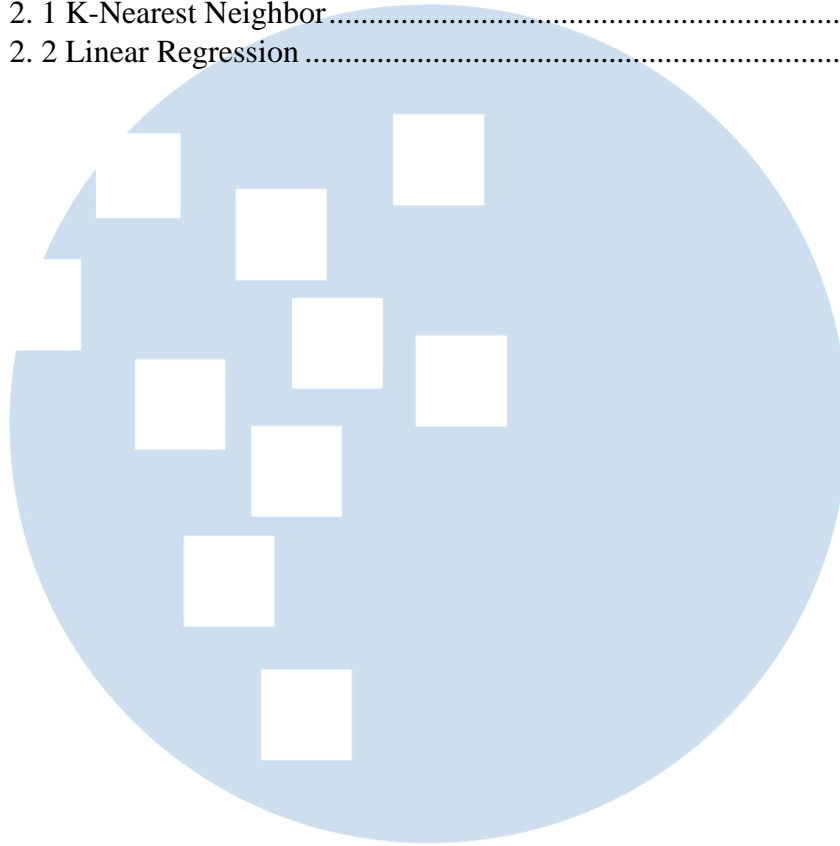


UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR RUMUS

Rumus 2. 1 K-Nearest Neighbor.....	15
Rumus 2. 2 Linear Regression	16

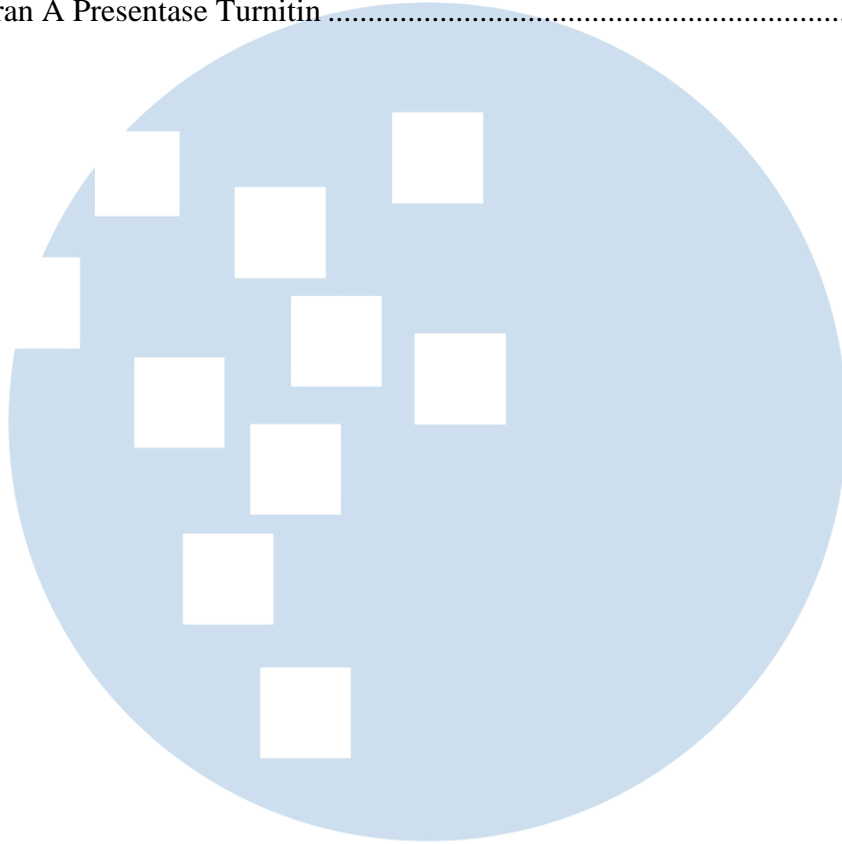


UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Presentase Turnitin 72



UMMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA