

**PENERAPAN METODE REGRESI LINIER UNTUK PREDIKSI
WAKTU PENCETAKAN KTP DISDUKCAPIL KECAMATAN
SAMARINDA UTARA**

SKRIPSI

Diajukan Oleh:

Akmal Mugni Fawwazrin

2011102441166



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR
JULI 2024**

**PENERAPAN METODE REGRESI LINIER UNTUK PREDIKSI
WAKTU PENCETAKAN KTP DISDUKCAPIL KECAMATAN
SAMARINDA UTARA**

SKRIPSI

Diajukan Sebagai Salah Satu Persyaratan Untuk Memperoleh Gelar Sarjana Komputer
Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur

Diajukan Oleh:

Akmal Mugni Fawwazrin

2011102441166



**PROGRAM STUDI S1 TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR**

JULI 2024

LEMBAR PERSETUJUAN

PENERAPAN METODE REGRESI LINIER UNTUK PREDIKSI WAKTU PENCETAKANKTP DISDUKCAPIL KECAMATAN SAMARINDA UTARA

SKRIPSI

Diajukan Oleh:

Akmal Mugni Fawwazrin

2011102441166

Disetujui untuk diujikan

Pada tanggal 22 Maret 2024

Pembimbing



Abdul Rahim, S.Kom., M.Cs

NIDN. 000904701

Mengetahui,

Koordinator Skripsi



Abdul Rahim, S.Kom., M.Cs

NIDN. 000904701

LEMBAR PENGESAHAN

**PENERAPAN METODE REGRESI LINIER UNTUK PREDIKSI
WAKTU PENCETAKAN KTP DISDUKCAPIL KECAMATAN
SAMARINDA UTARA**

SKRIPSI



Diajukan Oleh:

Akmal Mugni Fawwazrin

2011102441166

Diseminarkan dan Diujikan

Pada tanggal 15 Juli 2024

Penguji I	Penguji II
 <u>Fendy Yulianto, S.Kom., M.Kom</u> NIDN. 1102079402	 <u>Abdul Rahim, S.Kom., M.Cs</u> NIDN. 000904701

Mengetahui,

Ketua

Program Studi Teknik Informatika



Arbansyah, S.Kom., M.TI.
NIDN. 1118019203

PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan dibawah ini:

Nama : Akmal Mugni Fawwazrin
NIM : 2011102441166
Program Studi : Teknik Informatika
Judul Penelitian : Penerapan Metode Regresi Linier Untuk Prediksi Waktu Pencetakan KTP Disdukcopil Kecamatan Samarinda Utara

Menyatakan bahwa **skripsi** yang saya tulis ini benar-benar hasil karya saya sendiri, dan bukan merupakan hasil plagiasi/falsifikasi/fabrikasi baik sebagian atau seluruhnya.

Atas pernyataan ini, saya siap menanggung risiko dan sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam **skripsi** saya ini, atau klaim dari pihak lain terhadap keaslian saya ini.

Samarinda, 20 Juni 2024
Yang Membuat Pernyataan



Akmal Mugni Fawwazrin
NIM. 2011102441166

ABSTRAK

Dinas Kependudukan dan Pencatatan Sipil (Disdukcapil) memiliki peran penting dalam administrasi kependudukan, termasuk pencetakan Kartu Tanda Penduduk (KTP). Proses pencetakan KTP sering mengalami ketidakpastian waktu tunggu, menyebabkan ketidaknyamanan bagi pemohon. Penelitian ini bertujuan mengembangkan model prediksi waktu pencetakan KTP menggunakan Metode Regresi Linier, yang dipilih karena kesederhanaannya dalam menaksir parameter model berbasis runtun waktu. Model ini diharapkan mampu mengurangi kekhawatiran pemohon yang membutuhkan KTP untuk berbagai keperluan penting. Regresi Linier terbukti efektif dalam berbagai aplikasi prediksi sebelumnya, seperti produksi padi, harga saham, dan harga rumah. Untuk mengukur akurasi, digunakan *Mean Absolute Error* (MAE) karena memberikan ukuran yang mudah dipahami tentang seberapa baik kinerja model dengan menghitung rata-rata absolut dari kesalahan prediksi. Hasil penelitian menunjukkan bahwa model Regresi Linier yang diterapkan menghasilkan akurasi R^2 sebesar 98,38% serta *Mean Absolute Error* (MAE) sebesar 11,85957078. Dengan hasil ini, diharapkan model prediksi waktu pencetakan KTP dapat meningkatkan efisiensi pelayanan Disdukcapil, mengurangi ketidaknyamanan pemohon, dan memberikan kontribusi baru dalam penerapan Metode Regresi Linier di konteks administrasi kependudukan, khususnya dalam mengoptimalkan estimasi waktu pencetakan KTP.

Keywords: KTP, Prediksi, Regresi Linier, MAE

ABSTRACT

The Population and Civil Registration Office (Disdukcapil) plays a crucial role in population administration, including the issuance of National Identity Cards (KTP). The KTP printing process often experiences uncertain waiting times, causing discomfort for applicants. This study aims to develop a model to predict KTP printing times using the Linear Regression method, chosen for its simplicity in estimating model parameters based on time-series data. The model is expected to alleviate concerns for applicants who urgently need their KTPs for various important purposes. Linear Regression has proven effective in previous prediction applications, such as rice production, stock prices, and housing prices. To measure accuracy, the Mean Absolute Error (MAE) is used, providing an easily understandable measure of model performance by calculating the average absolute errors of predictions. The research results indicate that the applied Linear Regression model achieved an R^2 accuracy of 98.38% and a Mean Absolute Error (MAE) of 11,85957078. With these results, the KTP printing time prediction model is expected to enhance the efficiency of Disdukcapil's services, reduce applicant discomfort, and contribute new insights into the application of Linear Regression methods in population administration contexts, particularly in optimizing KTP printing time estimates.

Keywords: KTP, Prediction, Linear Regression, MAE

PRAKATA

Segala puji dan syukur penulis panjatkan kepada Allah SWT Sang Tuhan Yang Maha Esa dan Maha Segalanya, atas berkah rahmah dan curahan hidayat Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Penerapan Metode Regresi Linear Untuk Prediksi Waktu Pencetakan KTP Disdukcapil Kecamatan Samarinda Utara” ini tepat pada waktunya. Skripsi ini ditulis dalam rangka memenuhi syarat untuk mencapai gelar Sarjana Komputer pada Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur.

Penulis berupaya sebaik mungkin dalam menyelesaikan Skripsi ini yang menjadi sebuah karya yang dapat dipersembahkan. Dalam menyelesaikan studi dan penulisan skripsi ini, penulis banyak memperoleh bantuan baik pengajaran, bimbingan maupun arahan dari berbagai pihak baik secara langsung maupun tidak langsung. Untuk itu penulis menyampaikan penghargaan terbaik dan terima kasih yang tak terhingga kepada:

1. Kepada Allah SWT yang senantiasa memberikan jalan dan bantuan kepada hambanya yang lemah dan pemalas ini.
2. Kepada kedua orang tua penulis, Heliyana dan Surachman, saya persembahkan gelar sarjana ini kepada mereka berdua yang sudah memberikan banyak hal yang tidak bisa terbalaskan jasa budinya. Segala doa dan harapan mereka yang membuat saya dapat berdiri di titik ini dan semoga harapan mereka bisa terkabulkan dengan adanya saya dihidup mereka.
3. Kepada teman saya di prodi Teknik Informatika yang membantu selama masa kuliah
4. Kepada Bapak Ir. Sarjito sarjito, M.T., Ph.D selaku Dekan Fakultas Sains dan Teknologi Universitas Muhammadiyah Kalimantan Timur.
5. Kepada Bapak Arbansyah, S.Kom., M.TI selaku Kepala Program Studi Teknik Informatika yang senantiasa memberikan petunjuk dan nasehat kepada penulis.
6. Kepada Bapak Abdul Rahim, S.Kom., M.Cs selaku Dosen Pembimbing Skripsi yang senantiasa memberikan bantuan kepada penulis sehingga penulisan skripsi dapat selesai tepat waktu.
7. Kepada Bapak Taghfirul Azhima Yoga Siswa, S.Kom, M.Kom selaku Dosen Pembimbing Akademik yang sudah banyak membantu penulis terkait bimbingan selama masih menjadi mahasiswa hingga penulis dapat menyelesaikan studi di Jurusan Teknik Informatika Universitas Muhammadiyah Kalimantan Timur

8. Staff Dosen dan Tendik UMKT yang sudah banyak membantu penulis dalam menimba ilmu selama di bangku perkuliahan.

Serta kepada semua orang yang telah berjasa bagi hidup penulis dan tidak bisa disebutkan satu per satu. Pastinya tak henti-henti penulis sampaikan semoga amal baik semua pihak mendapat balasan yang berlipat ganda dari sang pencipta yang Maha Pengasih dan Penyayang Allah SWT. Amin.

Samarinda, 20 Juni 2024

Penyusun,

A handwritten signature in black ink, consisting of several overlapping loops and a long horizontal stroke extending to the right.

Akmal Mugni Fawwazrin

DAFTAR ISI

LEMBAR PERSETUJUAN	iii
LEMBAR PENGESAHAN	iv
PERNYATAAN KEASLIAN PENELITIAN	v
ABSTRAK.....	vi
PRAKATA.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Manfaat Penelitian	2
1.5 Batasan Masalah	3
BAB II METODE PENELITIAN.....	4
2.1 Objek Penelitian.....	4
2.2 Alat dan Bahan.....	4
2.2.1 Data Pencetakan KTP	4
2.2.2 Python	4
2.2.3 Google Colab	4
2.2.4 Regresi Linier.....	4
2.3 Tahapan Penelitian.....	5
2.3.1 Pengumpulan Data	5
2.3.2 Persiapan Data.....	6
2.3.3 Modelling Regresi Linier Berganda.....	7
2.3.4 Evaluasi Hasil	8
BAB III HASIL DAN PEMBAHASAN	9
3.1 Data Collection	9
3.2 Data Preprocessing.....	9
3.2.1 Pembersihan Data	9
3.2.2 Seleksi Data	10
3.2.3 Transformasi Data.....	10
3.3 Implementasi Model	12
3.3.1 Data Correlation.....	12

3.3.2	Split Data	13
3.3.3	Modelling Regresi Linier Berganda.....	14
3.3.4	Hasil	14
3.4	Perhitungan Regresi Linier Berganda	15
3.5	Evaluasi Hasil	20
BAB IV PENUTUP		23
4.1	Simpulan	23
4.2	Implikasi	23
DAFTAR RUJUKAN.....		25
LAMPIRAN.....		27

DAFTAR TABEL

	Halaman
Tabel 2.1 Atribut Data Pencetakan KTP	6
Tabel 2.2 Data Pencetakan KTP.....	6
Tabel 3.1 Label Encoding Keterangan	11
Tabel 3.2 Konversi Waktu	11
Tabel 3.3 Rasio Perbandingan.....	13
Tabel 3.4 Hasil Prediksi	15
Tabel 3.5 Data Perhitungan Regresi Linear	16
Tabel 3.6 Tabel Pembantu.....	16
Tabel 3.7 Data Testing	19
Tabel 3.8 Hasil Perhitungan	19
Tabel 3.9 Perbandingan Nilai Aktual dan Prediksi	20

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Tahapan Penelitian	5
Gambar 3.1 Data Collection.....	9
Gambar 3.2 Data Preprocessing	10
Gambar 3.3 Seleksi Data.....	10
Gambar 3.4 Transformasi Data	11
Gambar 3.5 Import Data.....	12
Gambar 3.6 Korelasi Data	12
Gambar 3.7 Hasil Korelasi	13
Gambar 3.8 Split Data.....	13
Gambar 3.9 Modelling Regresi Linier Berganda	14
Gambar 3.10 Konversi Hasil	14
Gambar 3.11 Menyimpan Hasil	15
Gambar 3.12 Grafik Perbandingan.....	21
Gambar 3.13 Hasil MAE Data Training.....	21
Gambar 3.14 R ² Score	211

DAFTAR LAMPIRAN

Lampiran 1 Surat Ijin Penelitian	27
Lampiran 2 Lembar Bimbingan	28
Lampiran 3 Kartu Kendali Bimbingan	29
Lampiran 4 Kode Pengujian Rasio Perbandingan.....	30
Lampiran 5 Kode Pengujian MAE.....	31
Lampiran 6 Riwayat Hidup.....	32

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dinas Kependudukan dan Pencatatan Sipil (Disdukcapil) memiliki peran penting dalam administrasi kependudukan suatu negara, Disdukcapil merupakan lembaga pemerintah yang bertanggung jawab atas pengelolaan dan pencatatan data penduduk di suatu wilayah tertentu (Indonesia, 2006). Tugas utama Disdukcapil meliputi pencatatan kelahiran, dokumen kependudukan, kartu keluarga (KK), akta nikah, dokumen pindah, dan pencetakan Kartu Tanda Penduduk (KTP) (Samosir et al., 2021). Salah satu tugas utama Disdukcapil adalah Pencetakan Kartu Tanda Penduduk (KTP)(Indonesia, 2006).

KTP adalah dokumen identitas resmi penduduk sebagai bukti diri yang diterbitkan oleh instansi pelaksana yang berlaku di seluruh wilayah Indonesia (Indonesia, 2006). Proses pencetakan KTP memiliki peran signifikan dalam memberikan identitas resmi bagi warga negara. Standar waktu pencetakan KTP adalah 15 menit setelah semua persyaratan terpenuhi (Disdukcapil Samarinda 2024, diakses 25 Juni 2024). Namun dalam beberapa kasus, waktu tunggu bagi pemohon KTP bisa menjadi masalah utama karena ketidakpastian dalam menentukan estimasi waktu pencetakan (Amus et al., 2023). Oleh karena itu, diperlukan sebuah model prediksi waktu pencetakan KTP.

Prediksi merupakan perkiraan terhadap suatu kondisi atau peristiwa di masa depan dengan memanfaatkan informasi berdasarkan data yang ada (Andriani et al., 2023). Salah satu Metode prediksi yang umum digunakan adalah Regresi Linier, sebuah Metode statistik yang digunakan untuk memprediksi berdasarkan hubungan antara variabel bebas dan variabel terikat (Marbun et al., 2018). Regresi Linier terbukti berhasil di berbagai bidang, seperti prediksi jumlah produksi padi dengan akurasi MAE 1,48950 (Suryanto, 2019), prediksi harga saham bank BRI dengan akurasi MAPE 13.751% (Putra et al., 2022), serta prediksi harga rumah dengan akurasi RMSE 259171.91 (Saiful, 2021).

Regresi Linier dipilih sebagai Metode prediksi pada penelitian ini karena kelebihanannya dalam menaksir parameter model yang sederhana dan data yang berbasis runtun waktu (Ayuni & Fitrihanah, 2019). Pada penelitian sebelumnya mengenai perbandingan model prediksi dalam prediksi harga saham, Regresi Linier terbukti lebih efektif dalam memprediksi dibandingkan dengan Random Forest serta Multilayer Perceptron karena menghasilkan nilai error lebih rendah dengan nilai RMSE sebesar 0.010, nilai MAPE sebesar 1,2%, nilai MAE sebesar 0.006 serta nilai koefisien determinasi (R^2) sebesar 99,8% (Fitri & Riana, 2022).

Penelitian ini juga melibatkan pemanfaatan Metode *Mean Absolute Error* (MAE) untuk mengukur akurasi kinerja model prediksi, MAE dipilih sebagai metrik evaluasi karena MAE merupakan metrik evaluasi yang mudah dihitung dan memberikan ukuran jelas terhadap kinerja sebuah model (Padhma M 2021, diakses 11 Juni 2024). Metode MAE menghitung nilai rata-rata absolute dari kesalahan model

prediksi tanpa menghiraukan tanda positif dan tanda negatif, semua kesalahan ditimbang pada skala yang sama karena nilai yang diambil adalah nilai mutlak, membuatnya mudah untuk dipahami (Nurfaidah & Abidin, 2024).

Meskipun Linear Regression telah diterapkan secara luas dalam penelitian-penelitian sebelumnya, metode ini belum pernah diimplementasikan untuk memprediksi waktu pencetakan KTP. Oleh karena itu, penulis berusaha mengisi celah ini dengan menerapkan Linear Regression sebagai alat prediksi untuk memprediksi waktu pencetakan KTP. Diharapkan bahwa penelitian dengan judul “Penerapan Metode Regresi Linier Untuk Prediksi Waktu Pencetakan KTP Disdukcapil Kecamatan Samarinda Utara” ini akan memberikan kontribusi baru dalam penerapan metode Linear Regression pada konteks administrasi kependudukan, khususnya dalam mengoptimalkan estimasi waktu pencetakan KTP.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas maka didapatkan rumusan masalah sebagai berikut:

- 1) Bagaimana mengembangkan model prediksi waktu pencetakan KTP di Disdukcapil Kecamatan Samarinda Utara menggunakan Metode Regresi Linier?
- 2) Berapa nilai akurasi MAE yang dihasilkan oleh model Regresi Linier dalam memprediksi waktu pencetakan KTP?

1.3 Tujuan Penelitian

Adapun tujuan penelitian sebagai berikut:

- 1) Mengembangkan model prediksi waktu pencetakan Kartu Tanda Penduduk (KTP) menggunakan Metode Regresi Linier.
- 2) Mengevaluasi nilai akurasi MAE yang dihasilkan oleh model Regresi Linier dalam memprediksi waktu pencetakan KTP.

1.4 Manfaat Penelitian

Adapun manfaat penelitian sebagai berikut:

- 1) Membantu mengoptimalkan proses administrasi kependudukan, khususnya dalam hal pencetakan Kartu Tanda Penduduk (KTP).
- 2) Model prediksi akan memberikan respon yang lebih cepat terhadap kebutuhan warga yang tidak sabar menunggu.
- 3) Penerapan Metode Regresi Linier dan evaluasi menggunakan MAE dapat dijadikan sebagai landasan untuk penelitian-penelitian serupa di masa depan.

1.5 Batasan Masalah

Untuk membatasi ruang lingkup penelitian, maka ditentukan batasan masalah sebagai berikut:

- 1) Data yang digunakan merupakan data waktu pencetakan KTP Disdukcapil Kecamatan Samarinda Utara yang diambil pada bulan Agustus 2023.
- 2) Mengolah serta analisis data waktu pencetakan KTP menggunakan Bahasa pemrograman Python.
- 3) Metode Regresi Linier yang digunakan dalam prediksi waktu pencetakan KTP pada penelitian ini adalah Regresi Linear Berganda.
- 4) Metode Mean Absolute Error (MAE) digunakan sebagai metrik evaluasi model prediksi waktu pencetakan KTP.

BAB II

METODE PENELITIAN

2.1 Objek Penelitian

Penelitian ini memiliki fokus khusus pada data pencetakan Kartu Tanda Penduduk (KTP) Disdukcapil Kecamatan Samarinda Utara pada bulan Agustus 2023 yang dikumpulkan selama pelaksanaan Program Kerja Lapangan (PKL). Penelitian ini bertujuan untuk mengembangkan sebuah model prediksi waktu pencetakan KTP dengan menerapkan Metode Regresi Linier Berganda.

2.2 Alat dan Bahan

Dalam proses penelitian ini, terdapat alat dan bahan yang digunakan untuk mendukung pelaksanaan kegiatan penelitian. Adapun alat dan bahan tersebut adalah sebagai berikut:

2.2.1 Data Pencetakan KTP

Pengumpulan data dilakukan dengan mengumpulkan data historis atau riwayat proses pencetakan KTP oleh Disdukcapil kecamatan samarinda utara yang berlokasi di Jalan Gunung Kapur, Lempake, Kec. Samarinda Utara, Kota Samarinda. Data ini mencakup pengumpulan data, verifikasi informasi, hingga proses pencetakan kartu KTP.

2.2.2 Python

Python adalah bahasa pemrograman yang banyak digunakan oleh pengembang untuk pengembangan aplikasi web, pengembangan perangkat lunak, *Data Science*, dan *Machine Learning* (AWS 2024, diakses 10 Maret 2024). Karena gratis dan mudah dipelajari, bahasa ini dapat digunakan di banyak platform, dan cocok dengan semua jenis sistem dan meningkatkan kecepatan pengembangan.

2.2.3 Google Colab

Google Colaboratory adalah dokumen yang dapat dieksekusi dan digunakan untuk menyimpan, menulis, serta berbagi program yang telah dibuat melalui Google Drive (Andre Oliver 2022, diakses 10 Maret 2024). *Software* ini pada dasarnya mirip dengan Jupyter Notebook gratis berbentuk *cloud* yang dapat digunakan dengan browser seperti Firefox dan Chrome. Penggunaanya dapat menjalankan kode Python tanpa perlu melakukan proses instalasi dan konfigurasi lainnya.

2.2.4 Regresi Linier

Regresi linier merupakan suatu Metode yang berfungsi untuk memeriksa korelasi antara variabel independen dan variabel dependen (Mardiani et al., 2023). Dalam konteks regresi linear sederhana, kita mempertimbangkan hubungan antara satu variabel independen (X) dan satu variabel dependen (Y), tujuannya adalah untuk membangun model dapat memprediksi nilai Y berdasarkan nilai X (Ayuni & Fitriyah, 2019). Persamaan dasar Metode Regresi Linier adalah sebagai berikut:

$$Y = a + bX \quad (2.1)$$

Dimana:

Y = Variabel Terikat

a = Konstanta (*intercept*)

b = Koefisien Regresi

X = Variabel Bebas

Jika terdapat masalah prediksi yang melibatkan dua atau lebih variabel bebas, Regresi Linier Berganda digunakan untuk mengatasi hal tersebut (Susanti & Sussolaikah, 2022). Persamaan Metode Regresi Linier Berganda adalah sebagai berikut:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (2.2)$$

Dimana:

Y = Variabel Terikat

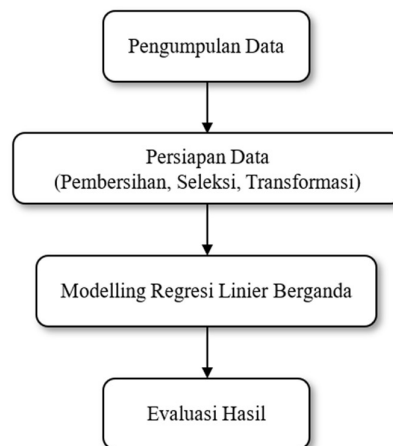
a = Konstanta (*intercept*)

b = Koefisien Regresi

X = Variabel Bebas

2.3 Tahapan Penelitian

Penelitian ini akan melibatkan 4 tahapan, dengan rincian tahapan pada gambar 2.1 sebagai berikut:



Gambar 2.1 Tahapan Penelitian

2.3.1 Pengumpulan Data

Data yang digunakan pada penelitian ini merupakan data pencetakan Kartu Tanda Penduduk (KTP) pada bulan Agustus 2023 sebanyak 350 data pencetakan yang dikumpulkan selama pelaksanaan

Program Kerja Lapangan (PKL) di Disdukcapil Kecamatan Lempake. Data tersebut memiliki beberapa atribut seperti KK, NIK, Nama, Umur, Jenis Kelamin, RT, Kelurahan, Keterangan, Tanggal Pengajuan, Waktu Pengajuan, serta Waktu Selesai sebagai label. Keterangan atribut dapat dilihat pada Tabel 2.1.

Tabel 2.1 Atribut Data Pencetakan KTP

NO	Atribut	Keterangan
1	KK	Nomor Kartu Keluarga
2	NIK	Nomor Induk Keluarga
3	Nama	Nama Pemohon
4	Umur	Umur Pemohon
5	Jenis Kelamin	Jenis Kelamin
6	RT	RT Pemohon
7	Kelurahan	Kelurahan Pemohon
8	Keterangan	Tujuan Pencetakan KTP
9	Tanggal Pengajuan	Tanggal Pengajuan
10	Waktu Pengajuan	Waktu Pengajuan
11	Waktu Selesai	Label/Waktu Selesai Pencetakan

Adapun contoh data yang digunakan dalam penelitian ini disajikan dalam bentuk Tabel 2.2 sebagai berikut.

Tabel 2.2 Data Pencetakan KTP

No	KK	NIK	Nama	Umur	Jenis Kelamin	RT	Kelurahan	Keterangan	Tanggal Pengajuan	Waktu Pengajuan	Waktu Selesai
1	6411012...	6472032...	THOMAS ...	42	L	12	SPJB	Pindahan	08/08/2023	08:05	08:10
2	6472052...	6472051...	SYAHRUL ...	20	L	6	SPJU	PRR	08/08/2023	08:32	08:48
3	6472052...	6472052...	ROHMAN...	17	L	15	TM	PRR	08/08/2023	09:19	09:31
4	6472051...	6472050...	MUHAMMAD...	17	L	29	SPJS	PRR	08/08/2023	09:47	10:03
5	6472052...	6472052...	LENDI ...	21	L	35	SPJT	Hilang	08/08/2023	10:10	10:16
...
346	6472050...	6472055...	PHEMA...	17	P	48	SPJT	PRR	08/30/23	09:03	09:27
347	6472053...	1210196...	MIKA...	17	P	29	TM	PRR	08/30/23	09:08	09:31
348	6472051...	6472052...	WINARKO	40	L	1	LPK	Hilang	08/30/23	09:13	09:21
349	6472052...	6472056...	RINA...	37	P	42	LPK	Hilang	08/30/23	09:19	09:26
350	3508182...	3508184...	NINGSIH	34	P	1	SPJU	Pindahan	08/30/23	09:26	09:40

2.3.2 Persiapan Data

Pada tahapan ini data yang telah didapat akan melewati beberapa persiapan seperti pembersihan data, seleksi data serta transformasi data untuk membersihkan data dari bentuk awalnya yang tidak terstruktur menjadi struktur yang lebih terorganisir untuk memudahkan proses prediksi. Berikut adalah langkah-langkahnya:

1) Pembersihan Data

Pembersihan data bertujuan untuk membersihkan data yang berpotensi mengurangi akurasi model seperti *Missing Value* (nilai yang hilang pada dataset) dengan cara menghapus baris/kolom data yang tidak memiliki nilai.

2) Seleksi Data

Pada tahapan ini akan dilakukan penghapusan atribut-atribut yang tidak relevan pada data seperti KK, NIK, Nama serta Tanggal Pengajuan sehingga data yang digunakan menjadi lebih efisien dan menghasilkan model yang optimal dalam melakukan prediksi.

3) Transformasi Data

Langkah ini dilakukan dengan tujuan mengubah jenis data kategorik menjadi bentuk numerik menggunakan teknik *Label Encoding*. Proses ini sangat penting karena pemodelan regresi linear hanya dapat menangani data numerik. Sebagai contoh, data mengalami transformasi dengan menggantikan setiap nilai dalam suatu kolom dengan angka berurutan seperti 0, 1, 2 dan seterusnya.

2.3.3 Modelling Regresi Linier Berganda

Karena proses prediksi melibatkan lebih dari 1 atribut maka Metode Regresi Linear Berganda digunakan untuk membangun model prediksi waktu pencetakan KTP dalam penelitian ini. Adapun tahapan dalam membuat model sebagai berikut:

1) *Import* Data

Import data merupakan langkah awal dalam proses pemodelan, berfungsi untuk menampilkan dataset yang telah dikumpulkan sebelumnya kedalam bentuk DataFrame.

2) Korelasi Data

Proses ini bertujuan untuk mencari hubungan antara atribut dengan label. Tujuannya adalah untuk mengidentifikasi atribut mana yang berpengaruh lebih signifikan dalam model prediksi ini. Korelasi ditentukan dengan memeriksa nilai positif terbesar dari atribut yang berkaitan dengan label data.

3) Split Data

Penelitian ini memanfaatkan teknik Split Data untuk mendapatkan hasil akurasi yang optimal dari proses dataset. Pada tahapan ini dilakukan pengujian terhadap beberapa rasio perbandingan untuk mendapatkan rasio perbandingan dataset yang terbaik.

4) Prediksi Regresi Linier Berganda

Proses ini akan menerapkan Metode Regresi Linier Berganda untuk memprediksi waktu pencetakan KTP berdasarkan atribut- atribut yang ada pada data.

5) Hasil

Hasil prediksi Regresi Linier Berganda nantinya akan berupa menit yang kemudian akan dikonversi ke dalam bentuk jam.

2.3.4 Evaluasi Hasil

Pada penelitian ini Metode *Mean Absolute Error* (MAE) digunakan untuk menghitung rata-rata perbedaan absolut antara nilai prediksi dan nilai (Rusdy et al., 2022). MAE merupakan salah satu Metode pengujian yang digunakan untuk mengevaluasi hasil akurasi model prediksi (Alita et al., 2021). Nilai MAE menunjukkan rata-rata kesalahan absolut antara nilai sebenarnya dan hasil prediksi (Li et al., 2021). Persamaan MAE dijelaskan sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.3)$$

Dimana:

\hat{y}_i = Nilai Hasil Prediksi

y_i = Nilai Sebenarnya

n = Jumlah Data

Selain itu, Koefisien determinasi atau R² juga digunakan untuk mengetahui seberapa pengaruh variable bebas (X) terhadap variabel terikat (Y), Semakin mendekati nilai 1 maka variabel bebas (X) dapat memiliki hubungan linier yang kuat terhadap variabel terikat (Y) (Yuliara, 2016). Persamaan untuk menghitung nilai R² adalah sebagai berikut:

$$r^2 = \frac{(b_1 \Sigma X_1 Y) + (b_2 \Sigma X_2 Y) + \dots + (b_n \Sigma X_n Y)}{\Sigma Y^2} \quad (2.4)$$

Dimana:

Y = Variabel Terikat

b = Koefisien Regresi

X = Variabel Bebas

BAB III

HASIL DAN PEMBAHASAN

3.1 Data Collection

Data *Collection* merupakan langkah yang digunakan untuk menggali informasi dari data yang nantinya akan digunakan. Proses ini penting untuk memahami data secara lebih mendalam sebelum dilakukan tahap Data *Preprocessing*.

```
import pandas as pd
data = pd.read_csv('Pencetakan_KTP.csv')
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 350 entries, 0 to 349
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype
---  -
0   KK                     350 non-null   int64
1   NIK                    350 non-null   int64
2   Nama                   350 non-null   object
3   Umur                   350 non-null   int64
4   Jenis_Kelamin         350 non-null   object
5   RT                     349 non-null   float64
6   Kelurahan              349 non-null   object
7   Keterangan             346 non-null   object
8   Tanggal_Pengajuan     350 non-null   object
9   Waktu_Pengajuan       350 non-null   object
10  Waktu_Selesai         350 non-null   object
dtypes: float64(1), int64(3), object(7)
memory usage: 30.2+ KB
```

Gambar 3.1 Data Collection

Gambar 3.1 memberikan informasi tentang setiap kolom dalam dataset, termasuk jumlah nilai non-null, dan tipe data dari setiap kolom. Berdasarkan output tersebut, dataset terdiri dari total 11 kolom, 4 kolom dengan tipe data numerik dan 7 kolom dengan tipe data objek, dengan jumlah data sebanyak 350 data. Beberapa kolom memiliki nilai non-null yang kurang dari total baris dataset, menandakan adanya *Missing Value* dalam dataset tersebut.

3.2 Data Preprocessing

Pada bagian ini, akan dijelaskan hasil dari setiap langkah persiapan data yang telah dilakukan, termasuk pembersihan data, seleksi data, dan transformasi data. Proses ini sangat penting untuk memastikan bahwa data yang digunakan dalam pemodelan prediktif adalah data yang bersih, relevan, dan dalam format yang sesuai. Adapun penjelasan setiap langkah sebagai berikut:

3.2.1 Pembersihan Data

Pembersihan data adalah langkah awal yang sangat penting untuk memastikan bahwa dataset yang digunakan tidak mengandung nilai hilang yang dapat mempengaruhi akurasi model. Pada tahap ini, hal yang akan dilakukan dapat dilihat pada Gambar 3.2.

```
data.isnull().sum()
KK                0
NIK               0
Nama              0
Umur              0
Jenis_Kelamin    0
RT                1
Kelurahan        1
Keterangan       4
Tanggal_Pengajuan 0
Waktu_Pengajuan  0
Waktu_Selesai    0
dtype: int64

data = data.dropna()
```

Gambar 3.2 Data Preprocessing

Perintah `'data.isnull().sum()'` pada Gambar 3.2 digunakan untuk menghitung jumlah nilai *Missing Value* di setiap kolom DataFrame. Kemudian, dengan menggunakan perintah `'data.dropna()'`, baris-baris yang mengandung *Missing Value* dihapus, sehingga menghasilkan DataFrame baru yang telah dibersihkan dari *Missing Value*. Data yang awalnya berjumlah 350 data berubah menjadi 344 data setelah dilakukan pembersihan data

3.2.2 Seleksi Data

Seleksi data dilakukan untuk menghapus atribut yang tidak relevan sehingga hanya data yang signifikan dan relevan yang digunakan untuk membangun model prediksi. Hal ini dilakukan Regresi Linier Berganda memerlukan variabel independen yang bersifat numerik atau kategorikal yang dapat diukur.

```
data = data.drop(columns=['KK', 'NIK', 'Nama', 'Jenis_Kelamin', 'Kelurahan', 'Tanggal_Pengajuan'])
```

Gambar 3.3 Seleksi Data

Perintah `'data.drop()'` pada Gambar 3.3 digunakan untuk menghapus kolom-kolom tertentu dari DataFrame. Dalam hal ini, kolom yang dihapus adalah 'KK', 'NIK', 'Nama', dan 'Tanggal_Pengajuan'. Penggunaan parameter `columns = ['KK', 'NIK', 'Nama', 'Jenis_Kelamin', 'Kelurahan', 'Tanggal_Pengajuan']` menginstruksikan Python untuk menghapus kolom-kolom tersebut dari DataFrame.

3.2.3 Transformasi Data

Transformasi data dilakukan untuk mengubah data kategorik menjadi bentuk numerik yang dapat diproses oleh model regresi linear. Proses ini menggunakan teknik *Label Encoding* dan melibatkan beberapa tahapan yang ditunjukkan pada Gambar 3.4.


```

from sklearn.preprocessing import LabelEncoder

le = LabelEncoder()
data['Keterangan'] = le.fit_transform(data['Keterangan'])

data['Waktu_Pengajuan'] = data['Waktu_Pengajuan'].apply(lambda x: int(x.split(':')[0]) * 60 + int(x.split(':')[1]) - 480)
data['Waktu_Selesai'] = data['Waktu_Selesai'].apply(lambda x: int(x.split(':')[0]) * 60 + int(x.split(':')[1]) - 480)

```

Gambar 3.4 Transformasi Data

LabelEncoder dari *library scikit-learn* untuk mengkonversi nilai-nilai dalam kolom 'Keterangan' menjadi bilangan bulat yang unik. Ini dilakukan dengan memanggil Metode '*fit_transform()*' dari objek *LabelEncoder* pada setiap kolom tersebut. Adapun nilai kolom yang berubah dapat dilihat pada Tabel 3.1 berikut:

Tabel 3.1 Label Encoding Keterangan

Keterangan	Label Encoding
Hilang	0
PRR / Baru Rekam	1
Pemekaran	2
Pindahan	3
Rusak	4
Ubah Data	5

Selain itu, format waktu pada kolom 'Waktu_Pengajuan' dan 'Waktu_Selesai' di ubah dari jam menjadi menit menggunakan fungsi '*lambda*' yang memecah string waktu dan mengonversi jam dan menit ke dalam satuan menit. Proses ini penting dalam mempersiapkan data, terutama ketika menggunakan Metode *Machine Learning* yang memerlukan input dalam bentuk numerik. Tabel 3.2 adalah contoh konversi yang dilakukan.

Tabel 3.2 Konversi Waktu

Format Jam	Satuan Menit
8:05	5
8:23	32
9:19	79
9:47	107
10:10	130

Fungsi '*apply*' digunakan untuk menerapkan fungsi '*lambda*' pada setiap elemen dalam kolom. Kemudian fungsi '*lambda*' akan memisahkan waktu ke dalam jam dan menit menggunakan '*split(':')*', mengkonversi jam menjadi menit dengan mengalikannya dengan 60, lalu menambahkan menitnya. Setelah itu hasilnya akan dikurangkan dengan 480 karena waktu yang diambil hanya dari pukul 8 pagi.

Hasil akhirnya adalah waktu dalam menit yang menggantikan nilai asli dalam kolom 'Waktu_Pengajuan' dan 'Waktu_Selesai'.

3.3 Implementasi Model

Langkah awal yang dilakukan dalam implementasi model adalah mengimpor data yang telah dikumpulkan sebelumnya dan menampilkannya dalam bentuk DataFrame. Data ini mencakup berbagai atribut yang relevan untuk prediksi waktu pencetakan KTP. Dengan menggunakan *Library* seperti *pandas*, data di *import* dan ditampilkan sehingga memudahkan dalam pengelolaan dan analisis lebih lanjut.

```
import pandas as pd
data = pd.read_csv('Pencetakan_KTP.csv')
```

Gambar 3.5 Import Data

Dengan menjalankan perintah '*pd.read_csv('')*' pada Gambar 3.5, data yang ada di file CSV akan dibaca dan dimuat ke dalam variabel data, sehingga dapat diolah dan dianalisis lebih lanjut menggunakan berbagai fungsi yang disediakan oleh *pandas*. Proses ini merupakan langkah awal yang penting dalam analisis data, memungkinkan data tersebut diakses dan diolah dengan mudah dalam penggunaan Python.

3.3.1 Data Correlation

Langkah selanjutnya adalah melakukan analisis korelasi untuk mengidentifikasi hubungan antara atribut dengan label. Tujuan dari analisis ini adalah untuk menemukan atribut yang memiliki pengaruh signifikan terhadap waktu pencetakan. Dengan menggunakan Metode korelasi, atribut-atribut dalam dataset dianalisis untuk menemukan nilai korelasi positif maupun negatif terbesar terhadap label.

```
korelasi = data.corr()
korelasi
```

Gambar 3.6 Korelasi Data

Fungsi '*data.corr()*' pada Gambar 3.6 digunakan untuk menghitung matriks korelasi dari semua pasangan kolom dalam DataFrame. Nilai korelasi berkisar antara -1 dan 1, dimana nilai 1 menunjukkan korelasi positif sempurna, nilai -1 menunjukkan korelasi negatif sempurna, dan nilai 0 menunjukkan tidak ada korelasi linear. Dengan menggunakan matriks korelasi ini, dapat diidentifikasi atribut-atribut mana yang memiliki hubungan paling kuat dengan label, sehingga atribut-atribut yang signifikan dapat dipilih untuk digunakan dalam model prediksi. Gambar 3.7 menunjukkan nilai korelasi pada tiap atribut.

	Umur	RT	Keterangan	Waktu_Pengajuan	Waktu_Selesai
Umur	1.000000	0.043878	0.367494	-0.064426	-0.121579
RT	0.043878	1.000000	0.049949	-0.062125	-0.061679
Keterangan	0.367494	0.049949	1.000000	-0.155001	-0.203164
Waktu_Pengajuan	-0.064426	-0.062125	-0.155001	1.000000	0.989696
Waktu_Selesai	-0.121579	-0.061679	-0.203164	0.989696	1.000000

Gambar 3.7 Hasil Korelasi

Dalam pengembangan model prediksi waktu pencetakan KTP ini, tiga atribut yang dipilih sebagai yang paling berpengaruh terhadap label 'Waktu_Selesai', yaitu 'Umur', 'Keterangan', dan 'Waktu_Pengajuan'. Pengambilan keputusan ini didasarkan pada hasil analisis korelasi, di mana ketiga atribut ini memiliki korelasi yang cukup signifikan dengan 'Waktu_Selesai'. 'Umur' dan 'Keterangan' memiliki korelasi negatif yang lemah, sementara 'Waktu_Pengajuan' memiliki korelasi positif yang lebih kuat dengan 'Waktu_Selesai'.

3.3.2 Split Data

Setelah mengidentifikasi atribut yang signifikan, langkah selanjutnya adalah membagi dataset menjadi dua bagian, yaitu data *Training* dan data *Testing*. Untuk menentukan rasio yang optimal antara data *Training* dan data *Testing*, dilakukan pengujian terhadap beberapa rasio perbandingan untuk melihat rasio mana yang menghasilkan akurasi tertinggi.

Tabel 3.3 Rasio Perbandingan

Rasio Perbandingan	Akurasi Data Training	Akurasi Data Testing
70:30	0.985058	0.980566
80:20	0.984837	0.979068
90:10	0.985317	0.966859

Berdasarkan hasil pengujian yang ditampilkan pada tabel 3.3, rasio 70:30 menunjukkan akurasi tertinggi, dengan nilai akurasi sebesar 0,98 untuk data *Training* serta data *Testing*. Dari hasil ini, dapat disimpulkan bahwa rasio 70:30 memberikan akurasi terbaik untuk pembagian data *Training* dan *Testing*. Oleh karena itu, rasio ini digunakan dalam proses pelatihan model.

```
from sklearn.model_selection import train_test_split
xtrain, xtest, ytrain, ytest = train_test_split(x, y, train_size=0.7, random_state = 42)
```

Gambar 3.8 Split Data

Fungsi `'train_test_split'` dari *Library scikit-learn* pada Gambar 3.8 adalah untuk membagi dataset menjadi dua bagian. Variabel *x* berisi fitur atau atribut yang akan digunakan untuk prediksi, sedangkan

y berisi label atau target yang akan diprediksi. Dengan parameter `'train_size=0.7'`, sebanyak 70% dari data akan dialokasikan untuk data training (x_{train} dan y_{train}) dan 10% sisanya untuk data *Testing* (x_{test} dan y_{test}). Penggunaan `random_state=42` memastikan bahwa pembagian data dilakukan secara acak namun konsisten.

3.3.3 Modelling Regresi Linier Berganda

Dengan data yang telah dibagi, Metode Regresi Linier Berganda diterapkan untuk membangun model prediksi. Model dilatih menggunakan data training yang telah disiapkan sebelumnya. Metode Regresi Linier Berganda digunakan untuk memprediksi waktu pencetakan KTP berdasarkan atribut-atribut yang telah dipilih melalui analisis korelasi.

```
from sklearn.linear_model import LinearRegression
reg = LinearRegression().fit(xtrain, ytrain)

y_pred = reg.predict(xtest)
```

Gambar 3.9 Modelling Regresi Linier Berganda

Model *LinearRegression* dari *Library scikit-learn* pada gambar 3.9 digunakan untuk memprediksi nilai target berdasarkan data yang digunakan. Proses ini dimulai dengan melatih model Regresi Linier Berganda menggunakan data *Training* yang terdiri dari fitur (x) dan target (y) melalui Metode `'fit()`'. Setelah model dilatih, model tersebut kemudian digunakan untuk memprediksi nilai-nilai target pada data uji (x_{test}) dengan menggunakan Metode `'predict()`'. Hasil prediksi ini disimpan dalam `'y_pred'`, yang berisi nilai-nilai target yang diprediksi berdasarkan model yang telah dilatih sebelumnya.

3.3.4 Hasil

Hasil prediksi model berbentuk satuan menit, oleh karena itu dibuat sebuah fungsi untuk mengkonversi hasil prediksi menjadi format jam. Fungsi ini akan mengubah satuan menit menjadi format waktu yang lebih terstruktur. Fungsi yang dibuat dapat dilihat pada gambar 3.10.

```
def convert_integer_to_time(minutes):
    jam = minutes // 60
    menit = minutes % 60
    return f"jam:02d}:{menit:02d}"
```

Gambar 3.10 Konversi Hasil

Fungsi ini menghitung jumlah jam dengan membagi total menit dengan 60 menggunakan operator (`//`), yang menghasilkan jumlah jam. Selanjutnya, menggunakan operator (`%`) untuk menghitung sisa waktu dalam bentuk menit. Hasilnya diformat menggunakan *f-string* (`f"jam:02d}:{menit:02d}"`), yang akan menampilkan waktu dalam format jam dan menit.

```

data_prediksi = pd.DataFrame({
    'Selesai': y_pred.astype(int) - xtest['Waktu_Pengajuan'].astype(int),
    'Waktu_Selesai': y_pred.astype(int)
})

data_prediksi['Selesai'] = data_prediksi['Selesai'].astype(str) + ' menit'
data_prediksi['Waktu_Selesai'] = data_prediksi['Waktu_Selesai'].apply(convert_integer_to_time)

```

Gambar 3.11 Menyimpan Hasil

Gambar 3.11 merupakan langkah terakhir yang dilakukan untuk membuat sebuah DataFrame baru ‘*data_prediksi*’ menggunakan library pandas. DataFrame ini memiliki satu kolom awal yang disebut ‘Waktu_Selesai’, berisi nilai prediksi ‘*y_pred*’ yang sudah diubah tipe datanya menjadi integer. Selanjutnya, kolom ‘Waktu_Selesai’ ditambahkan dengan menerapkan fungsi *convert_integer_to_time* ke setiap nilai dalam kolom. Hasil akhirnya berisi prediksi ‘Waktu_Selesai’ dalam format waktu yang lebih terstruktur dan mudah dipahami.

Tabel 3.4 Hasil Prediksi

No	Selesai	Waktu_Selesai
1	14 menit	09:36
2	12 menit	10:38
3	30 menit	11:45
4	36 menit	15:06
5	14 menit	08:31
...
100	14 menit	11:21
101	27 menit	10:08
102	5 menit	11:22
103	6 menit	10:30
104	22 menit	13:26

Tabel 3.4 menampilkan hasil prediksi berdasarkan model Regresi Linier Berganda, dimana kolom ‘Selesai’ adalah lama waktu pencetakan KTP dalam bentuk menit dan kolom ‘Waktu_Selesai’ adalah hasil prediksi yang telah diubah menjadi format jam.

3.4 Perhitungan Regresi Linier Berganda

Pada penelitian ini Regresi Linear yang digunakan adalah Regresi Linear Berganda. Regresi Linear Berganda merupakan teknik statistik yang digunakan untuk memodelkan hubungan antara satu variabel dependen *Y* dengan dua atau lebih variabel independen *X1, X2, X3, ..., Xn*. Perhitungan regresi linear berganda melibatkan beberapa langkah utama, yaitu membentuk matriks, menghitung determinan, dan menentukan nilai koefisien. Adapun data yang digunakan dapat dilihat pada tabel 3.5.

Tabel 3.5 Data Perhitungan Regresi Linear

No	Umur	Keterangan	Waktu_Pengajuan	Waktu_Selesai
1	56	3	423	435
2	23	5	352	360
3	18	3	156	169
4	39	3	22	32
5	17	1	403	462
...
236	18	1	17	49
237	78	0	125	138
238	17	1	91	109
239	17	1	342	370
240	23	5	75	84
Total	7034	546	40082	44873

Data yang digunakan adalah data bersih yang telah diolah pada tahap *Preporcessing*, dengan total data sebanyak 240 data *Training*. Untuk mendapatkan nilai yang nantinya akan digunakan dalam perhitungan matriks A, maka di buat tabel pembantu dengan melakukan perhitungan terhadap variabel independen X1 (Umur), X2 (Keterangan), X3 (Waktu_Pengajuan), dan variabel dependen Y (Waktu_Selesai). Tabel 3.6 menunjukan perhitungan yang dilakukan terhadap variabel independen serta variabel dependen.

Tabel 3.6 Tabel Pembantu

No	X1*Y	X2*Y	X3*Y	X1*X2	X1*X3	X2*X3	X1^2	X2^2	X3^2
1	24360	1305	184005	168	23688	1269	3136	9	178929
2	8280	1800	126720	115	8096	1760	529	25	123904
3	3042	507	26364	54	2808	468	324	9	24336
4	1248	96	704	117	858	66	1521	9	484
5	7854	462	186186	17	6851	403	289	1	162409
...
236	882	49	833	18	306	17	324	1	289
237	10764	0	17250	0	9750	0	6084	0	15625
238	1853	109	9919	17	1547	91	289	1	8281
239	6290	370	126540	17	5814	342	289	1	116964
240	1932	420	6300	115	1725	375	529	25	5625
Total	1274971	88863	12489149	17931	1164511	81100	255624	1816	11498286

Langkah selanjutnya adalah menghitung nilai Determinan A , A_0 , A_1 , A_2 dan A_3 dengan menggunakan matriks A . Karena terdapat 3 persamaan dengan 3 variabel yang tidak diketahui nilainya, yaitu a , b_1 , b_2 serta b_3 , persamaan tersebut dapat dinyatakan dalam persamaan matriks sebagai berikut:

$$\begin{bmatrix} n & \sum X_1 & \sum X_2 & \sum X_3 \\ \sum X_1 & \sum X_1^2 & \sum X_1 X_2 & \sum X_1 X_3 \\ \sum X_2 & \sum X_1 X_2 & \sum X_2^2 & \sum X_2 X_3 \\ \sum X_3 & \sum X_1 X_3 & \sum X_2 X_3 & \sum X_3^2 \end{bmatrix} \begin{bmatrix} a \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum X_1 Y \\ \sum X_2 Y \\ \sum X_3 Y \end{bmatrix} \quad (3.1)$$

Dimana:

n = Jumlah Data

a = Konstanta (intercept)

b_i = Koefisien Regresi

Matriks pertama adalah matriks A , Matriks kedua adalah nilai-nilai yang akan dicari dan Matriks ketiga adalah nilai H yang merupakan hasil dari perkalian antara nilai-nilai Y . Dengan menggunakan hasil dari tabel pembantu, Maka perhitungan determinan dari matriks A adalah sebagai berikut:

$$A = \begin{bmatrix} 240 & 7034 & 546 & 40082 \\ 7034 & 255624 & 17931 & 1164511 \\ 546 & 17931 & 1816 & 81100 \\ 40082 & 1164511 & 81100 & 11498286 \end{bmatrix} = 27316156751950700$$

Untuk menghitung determinannya, kita bisa menggunakan Metode ekspansi kofaktor atau Metode lain yang sesuai. Setelah mendapatkan hasil determinan matriks A , langkah selanjutnya adalah menghitung matriks A_0 , A_1 , A_2 , serta A_3 . Berikut adalah perhitungan dari setiap matriks:

- 1) Matriks A_0 dihitung dengan menggantikan kolom pertama dengan nilai H

$$A_0 = \begin{bmatrix} 44873 & 7034 & 546 & 40082 \\ 1274971 & 255624 & 17931 & 1164511 \\ 88863 & 17931 & 1816 & 81100 \\ 12489149 & 1164511 & 81100 & 11498286 \end{bmatrix} = 983270794940830000$$

- 2) Matriks A_1 dihitung dengan menggantikan kolom kedua dengan nilai H

$$A_1 = \begin{bmatrix} 240 & 44873 & 546 & 40082 \\ 7034 & 1274971 & 17931 & 1164511 \\ 546 & 88863 & 1816 & 81100 \\ 40082 & 12489149 & 81100 & 11498286 \end{bmatrix} = -12814013437205300$$

- 3) Matriks A_2 dihitung dengan menggantikan kolom ketiga dengan nilai H

$$A_2 = \begin{bmatrix} 240 & 7034 & 44873 & 40082 \\ 7034 & 255624 & 1274971 & 1164511 \\ 546 & 17931 & 88863 & 81100 \\ 40082 & 1164511 & 12489149 & 11498286 \end{bmatrix} = -91013916371081000$$

4) Matriks A_3 dihitung dengan menggantikan kolom keempat dengan nilai H

$$A_3 = \begin{bmatrix} 240 & 7034 & 546 & 44873 \\ 7034 & 255624 & 17931 & 1274971 \\ 546 & 17931 & 1816 & 88863 \\ 40082 & 1164511 & 81100 & 12489149 \end{bmatrix} = 28182233421495900$$

Setelah mendapatkan semua nilai determinan matriks A , maka diperoleh persamaan untuk menghitung nilai a serta b sebagai berikut:

$$a = \frac{Det A_0}{Det A} \quad b_1 = \frac{Det A_1}{Det A} \quad b_2 = \frac{Det A_2}{Det A} \quad b_3 = \frac{Det A_3}{Det A} \quad (3.2)$$

1) Menghitung nilai a

$$a = \frac{Det A_0}{Det A} = \frac{983270794940830000}{27316156751950700} = 35,9959$$

2) Menghitung nilai b_1

$$b_1 = \frac{Det A_1}{Det A} = \frac{-12814013437205300}{27316156751950700} = -0,4691$$

3) Menghitung nilai b_2

$$b_2 = \frac{Det A_2}{Det A} = \frac{-91013916371081000}{27316156751950700} = -3,3319$$

4) Menghitung nilai b_3

$$b_3 = \frac{Det A_3}{Det A} = \frac{28182233421495900}{27316156751950700} = 1,0317$$

Setelah mendapatkan nilai a dan b , langkah selanjutnya adalah membuat model persamaan regresi linear berganda menggunakan persamaan 2.2. Berikut adalah model persamaan regresi linear berganda:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3$$

$$Y = 35,9959 - 0,4691 * X_1 - 3,3319 * X_2 + 1,0317 * X_3$$

Dengan menggunakan persamaan diatas, akan dilakukan perhitungan linear regresi berganda menggunakan data *Testing* yang diambil setelah proses *Split Data*. Data *Testing* yang digunakan dapat dilihat pada Tabel 3.7.

Tabel 3.7 Data Testing

No	Umur	Keterangan	Waktu_Pengajuan
1	30	3	82
2	39	3	146
3	18	1	195
4	18	1	390
5	25	3	17
...
100	37	3	187
101	18	1	101
102	58	3	197
103	63	5	156
104	27	3	304

Dengan perbandingan, 70:30 didapatkan data *Testing* sebanyak 104 data. Data ini akan dihitung menggunakan persamaan yang telah didapat sebelumnya untuk mendapatkan hasil nilai prediksi. Perhitungan dari setiap data dapat dilihat pada Tabel 3.8 berikut:

Tabel 3.8 Hasil Perhitungan

No	Perhitungan Regresi Linear	Hasil Prediksi
1	$Y = 35,9959 - 0,4691(30) - 3,3319(3) + 1,0317(82)$	96,5271875
2	$Y = 35,9959 - 0,4691(39) - 3,3319(3) + 1,0317(146)$	158,3344480
3	$Y = 35,9959 - 0,4691(18) - 3,3319(1) + 1,0317(195)$	225,4028709
4	$Y = 35,9959 - 0,4691(18) - 3,3319(1) + 1,0317(390)$	426,5854738
5	$Y = 35,9959 - 0,4691(25) - 3,3319(3) + 1,0317(17)$	31,8118206
...
100	$Y = 35,9959 - 0,4691(37) - 3,3319(3) + 1,0317(187)$	201,5725802
101	$Y = 35,9959 - 0,4691(18) - 3,3319(1) + 1,0317(101)$	128,4225392
102	$Y = 35,9959 - 0,4691(58) - 3,3319(3) + 1,0317(197)$	202,0385335
103	$Y = 35,9959 - 0,4691(63) - 3,3319(5) + 1,0317(156)$	150,7293584
104	$Y = 35,9959 - 0,4691(27) - 3,3319(3) + 1,0317(304)$	326,9731436

3.5 Evaluasi Hasil

Evaluasi model adalah tahap terakhir yang dilakukan untuk mengukur kinerja model prediksi. Mean Absolute Error (MAE) adalah Metode pengujian yang digunakan untuk mengevaluasi akurasi model prediksi ini. Nilai MAE didapat dengan menghitung rata-rata kesalahan absolut antara nilai aktual dan hasil prediksi. Tabel 3.9 menunjukkan perbandingan antara nilai aktual dan nilai prediksi, serta kesalahan absolut untuk setiap hasil prediksi.

Tabel 3.9 Perbandingan Nilai Aktual dan Prediksi

No	Nilai Aktual	Nilai Prediksi	Absolute Error
1	89	96,5271875	7,5271875
2	162	158,3344480	3,6655520
3	314	225,4028709	88,5971291
4	422	426,5854738	4,5854738
5	35	31,8118206	3,1881794
...
100	203	201,5725802	1,4274198
101	122	128,4225392	6,4225392
102	201	202,0385335	1,0385335
103	165	150,7293584	14,2706416
104	313	326,9731436	13,9731436
Total	1233,3953612

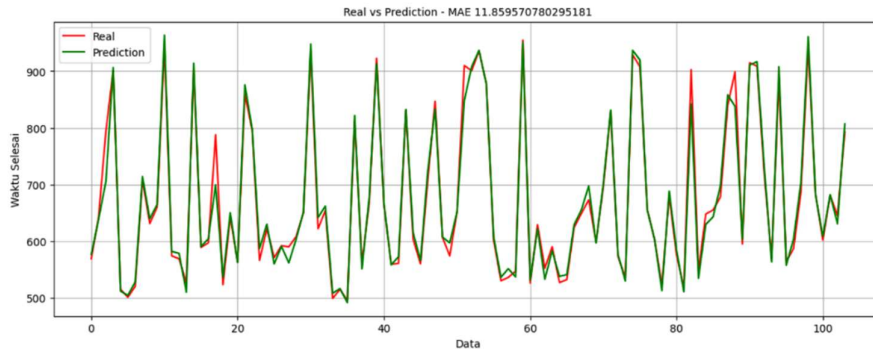
Nilai Absolute error dihitung sebagai perbedaan absolut antara nilai aktual dan nilai prediksi untuk setiap observasi. Dalam hal ini, total absolute error adalah 1233,3953612. Setelah itu, kita membagi total absolute error ini dengan jumlah data prediksi untuk mendapatkan nilai rata-rata dari kesalahan absolut. Karena ada 104 data prediksi dalam dataset ini, kita dapat menghitung MAE dengan persamaan berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

$$MAE = \frac{1233,3953612}{104} = 11,85957078$$

Nilai MAE sebesar 11,85957078 menunjukkan bahwa, secara rata-rata prediksi model Regresi Linier Berganda meleset sekitar 12 menit dari nilai aktualnya. Nilai MAE yang lebih rendah mengindikasikan bahwa model memiliki kinerja yang lebih baik karena kesalahan prediksinya lebih kecil. Dalam konteks penelitian ini, nilai MAE membantu mengevaluasi sejauh mana model Regresi

Linier Berganda dapat diandalkan dalam memprediksi waktu pencetakan KTP berdasarkan atribut yang dipilih.



Gambar 3.12 Grafik Perbandingan

Perbandingan antara nilai aktual dan nilai prediksi juga dapat dilihat pada Gambar 3.12 di atas, dimana garis warna merah menunjukkan data aktual sedangkan garis warna hijau menunjukkan data hasil prediksi. Hal ini memberikan pandangan langsung tentang seberapa dekat atau jauh hasil nilai prediksi dari nilai sebenarnya, memungkinkan evaluasi yang lebih mendalam terhadap keakuratan model yang digunakan.

```
from sklearn.metrics import mean_absolute_error
mae=mean_absolute_error(ytrain,y_pred)
print(mae)
11.48745823630302
```

Gambar 3.13 Hasil MAE Data *Training*

Gambar 3.13 menunjukkan bahwa pengujian nilai MAE juga dilakukan pada data *Training* yang menghasilkan nilai MAE sebesar 11,48745823, menunjukkan bahwa model Regresi Linier Berganda memiliki kinerja yang cukup konsisten. Perbedaan kecil antara nilai MAE pada data *Training* dan *Testing* menandakan model tidak mengalami overfitting yang signifikan, sehingga dapat memprediksi data baru dengan kesalahan prediksi tidak jauh berbeda dari data pelatihan. Evaluasi ini menunjukkan bahwa model memiliki kinerja yang baik dan dapat diandalkan dalam memprediksi waktu pencetakan KTP berdasarkan atribut yang dipilih.

```
reg.score(x, y)
```

0.9838423659127112

Gambar 3.14 R² Score

Adapun hasil akurasi pada Gambar 3.14 yang dihasilkan dari model yang sudah diuji dengan data prediksi dalam bentuk R² yaitu sebesar 0.9838. Hasil tersebut mengindikasikan bahwa hubungan linear

yang kuat antara atribut Umur, Keterangan, serta Waktu Pengajuan dengan label Waktu Selesai. Semakin nilai R^2 mendekati satu, maka semakin baik prediksi yang dihasilkan oleh model prediksi Regresi Linier Berganda.

BAB IV

PENUTUP

4.1 Simpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan pada penelitian ini, diperoleh kesimpulan bahwa metode regresi linier berganda dapat digunakan untuk memprediksi waktu pencetakan KTP secara efektif. Proses penelitian dimulai dari pengumpulan data sebanyak 350 data pencetakan KTP. Selanjutnya, dilakukan preprocessing data yang mencakup pemilihan atribut yang tepat, yaitu Umur, Keterangan, dan Waktu Pengajuan, serta penentuan rasio terbaik untuk pembagian data *Training* dan *Testing*, yaitu 70:30.

Model regresi linier berganda menghasilkan nilai R^2 sebesar 0,9838, yang mengindikasikan adanya hubungan linear yang sangat kuat antara atribut Umur, Keterangan, dan Waktu Pengajuan dengan label Waktu_Seleasai. Dengan nilai R^2 yang tinggi tersebut, model ini mampu memprediksi waktu pencetakan KTP dengan akurasi sebesar 98,38%. Hal ini menunjukkan bahwa model regresi linier berganda yang dikembangkan dalam penelitian ini dapat diandalkan untuk memprediksi waktu pencetakan KTP berdasarkan atribut yang telah dipilih, sehingga dapat meningkatkan efisiensi dan efektivitas dalam proses pencetakan KTP.

Selain itu, model regresi linier yang digunakan juga menghasilkan nilai Mean Absolute Error (MAE) sebesar 11,85957078. Nilai MAE ini menunjukkan bahwa rata-rata kesalahan prediksi waktu pencetakan KTP oleh model regresi linier adalah sekitar 12 menit. Hal ini memperkuat kesimpulan bahwa Metode regresi linier tidak hanya memiliki akurasi yang tinggi, tetapi juga memiliki tingkat kesalahan prediksi yang rendah, sehingga dapat diandalkan untuk memprediksi waktu pencetakan KTP secara efektif.

4.2 Implikasi

Berdasarkan kesimpulan yang telah diuraikan, berikut adalah implikasi penelitian yang dapat diambil:

1. Dengan kemampuan memprediksi waktu pencetakan KTP secara akurat, Disdukcapil dapat mengoptimalkan alur kerja dan memperkirakan kebutuhan sumber daya secara lebih efisien. Hal ini akan membantu dalam mengurangi waktu tunggu bagi pemohon dan meningkatkan kepuasan masyarakat terhadap pelayanan publik.
2. Hasil penelitian ini dapat diintegrasikan ke dalam sistem informasi yang digunakan oleh Disdukcapil untuk mengelola proses pencetakan KTP. Aplikasi berbasis web atau mobile dapat dikembangkan untuk memberikan estimasi waktu penyelesaian kepada pemohon secara real-time, meningkatkan transparansi dan komunikasi antara instansi dan masyarakat.
3. Keberhasilan penggunaan Metode regresi linier dalam penelitian ini dapat menjadi dasar untuk mengeksplorasi dan mengembangkan Metode prediksi lain yang mungkin lebih kompleks atau

dapat menangani lebih banyak variabel. Penelitian lebih lanjut dapat melibatkan teknik machine learning lainnya seperti regresi non-linier, pohon keputusan, atau Metode ensemble untuk meningkatkan akurasi prediksi.

Dengan demikian, penelitian ini tidak hanya memberikan kontribusi teoritis dalam bidang analisis data dan prediksi, tetapi juga memiliki dampak praktis yang signifikan dalam meningkatkan kualitas pelayanan publik.





DAFTAR RUJUKAN

- Alita, D., Putra, A. D., & Darwis, D. (2021). Analysis of classic assumption test and multiple linear regression coefficient test for employee structural office recommendation. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 15(3), 295.
<https://doi.org/10.22146/ijccs.65586>
- Amus, M. N., Yohanes, S., Udju, H. R., Kholipah, S. A., Nusa, U., Kupang, C., & Cirebon, P. S. C. I. (2023). PELAYANAN PUBLIK DALAM KONTES KEPENDUDUKAN CATATAN SIPIL KABUPATEN MANGGARAI * Correspondence INFO ARTIKEL Diterima : 28-05-2023 Direvisi : 09-06-2023 Disetujui : 09-09-2023 Pendahuluan Sejak diberlakukan otonomi daerah , pelayanan publik telah menjadi to. *Jurnal Indonesia Sosial Teknologi*, 4(6), 655–663.
- Andre Oliver. (2022). *Mengenal Google Colab: Mulai dari Definisi, Cara Menggunakan, hingga Manfaatnya*. Glints. <https://glints.com/id/lowongan/google-colab-adalah/>
- Andriani, W., Gunawan, & Prayoga, A. E. (2023). Prediksi Nilai Emas Menggunakan Algoritma Regresi Linear. *Jurnal Ilmiah Informatika Komputer*, 28(1), 27–35.
<https://doi.org/10.35760/ik.2023.v28i1.8096>
- AWS. (2024). *What is Python ? What are the benefits of Python ? How is Python used ?*
https://aws.amazon.com/what-is/python/?nc1=h_ls
- Ayuni, G. N., & Fitriyah, D. (2019). Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ. *Jurnal Telematika*, 14(2), 79–86.
<https://journal.ithb.ac.id/telematika/article/view/321>
- Disdukcapil Samarinda. (2024). *Dinas Kependudukan & Pencatatan Sipil Kota Samarinda*.
https://disdukcapil.samarindakota.go.id/pelayanan_ktp
- Fitri, E., & Riana, D. (2022). Analisa Perbandingan Model Prediction Dalam Prediksi Harga Saham Menggunakan Metode Linear Regression, Random Forest Regression Dan Multilayer Perceptron. *METHOMIKA Jurnal Manajemen Informatika Dan Komputerisasi Akuntansi*, 6(1), 69–78. <https://doi.org/10.46880/jmika.vol6no1.pp69-78>
- Indonesia, P. (2006). *Undang-undang (UU) Nomor 23 Tahun 2006 tentang Administrasi Kependudukan* (Issue December, pp. 1–6).
- Li, W., Wu, H., Zhu, N., Jiang, Y., Tan, J., & Guo, Y. (2021). Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU). *Information Processing in Agriculture*, 8(1), 185–193. <https://doi.org/10.1016/j.inpa.2020.02.002>
- Marbun, M., Sihotang, H. T., & Nababan, M. A. (2018). Perancangan Sistem Peramalan Jumlah Wisatawan Asing. *Jurnal Mantik Penusa*, 2(1), 41–49.
- Mardiani, E., Rahmansyah, N., Ningsih, S., & ... (2023). Komparasi Metode Knn, Naive Bayes, Decision Tree, Ensemble, Linear Regression Terhadap Analisis Performa Pelajar Sma. *Innovative: Journal Of ...*, 3(2), 13880–13892. <http://j->

- innovative.org/index.php/Innovative/article/view/1949%0Ahttp://j-innovative.org/index.php/Innovative/article/download/1949/1468
- Nurfaidah, H., & Abidin, W. (2024). Penerapan Metode Single Moving Average Dalam Peramalan Curah Hujan Kota Makassar. *Jurnal MSA (Matematika Dan Statistika Serta Aplikasinya)*, 11(2), 134–139. <https://doi.org/10.24252/msa.v11i2.45815>
- Padhma M. (2021). *Pros of the MAE Evaluation Metric Cons of the MAE evaluation metric*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/10/evaluation-metric-for-regression-models/>
- Putra, J. S., Ramadhani, R. D., & Burhanuddin, A. (2022). Prediksi Harga Saham Bank Bri Menggunakan Algoritma Linear Regression Sebagai Strategi Jual Beli Saham. *Journal of Dinda : Data Science, Information Technology, and Data Analytics*, 2(1), 1–10. <https://doi.org/10.20895/dinda.v2i1.273>
- Rusdy, A. M. A., Purnawansyah, P., & Herman, H. (2022). Penerapan Metode Regresi Linear Pada Prediksi Penawaran dan Permintaan Obat Studi Kasus Aplikasi Point Of Sales. *Buletin Sistem Informasi Dan Teknologi Islam*, 3(2), 121–126. <https://doi.org/10.33096/busiti.v3i2.1130>
- Saiful, A. (2021). Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(1), 41–50. <https://doi.org/10.35957/jatisi.v8i1.701>
- Samosir, F. R. S., Damanik, I. S., Suhendro, D., Solikhun, & Susiani. (2021). Analisis Penilaian Kualitas Jenis Pelayanan Tebaik dengan Metode Analytic Network Process (ANP) di Kantor Dinas Kependudukan dan Pencatatan Sipil Kota Pematangsiantar. *BIOS : Jurnal Teknologi Informasi Dan Rekayasa Komputer*, 2(1), 39–50. <https://doi.org/10.37148/bios.v2i1.21>
- Suryanto, A. A. (2019). Penerapan Metode Mean Absolute Error (Mea) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi. *Saintekbu*, 11(1), 78–83. <https://doi.org/10.32764/saintekbu.v11i1.298>
- Susanti, P., & Sussolaikah, K. (2022). Penerapan Metode Regresi Linear Untuk Memprediksi Harga Jual Mobil Bekas Yaris Dan Jazz Pada Wilayah Dki Jakarta. *Network Engineering Research Operation*, 7(2), 133. <https://doi.org/10.21107/nero.v7i2.314>
- Yuliara, I. M. (2016). Regresi linier berganda. In *Book Section* (pp. 1–6). <http://www.mendeley.com/research/regresi-linier-berganda-1/>

LAMPIRAN

Lampiran 1 Surat Ijin Penelitian

 <p>UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR Berprestasi Berwawasan Berkeadilan</p>	UMKT Program Studi Teknik Informatika Fakultas Sains dan Teknologi	Telp. 0541-748511 Fax.0541-766832 Website http://informatika.umkt.ac.id email: informatika@umkt.ac.id	  
---	---	---	---

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Nomor : 003-034/FST.1/A.7/C/2024
Lampiran : -
Perihal : **Permohonan Pengambilan Data**

Kepada Yth.
Kepala Camat Samarinda Utara
di -
Tempat

Assalamu'alaikum Warrahmatullahi Wabarrakatuh

Puji Syukur kepada Allah Subhanahu wa ta'ala yang senantiasa melimpahkan Rahmat-Nya kepada kita sekalian. Aamiin.

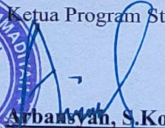
Sehubungan untuk memenuhi Tugas Akhir/Skripsi Tahun Akademik 2023/2024, maka dengan ini kami mengajukan permohonan untuk melakukan pengambilan data di Kecamatan Samarinda Utara. Adapun data yang diminta yaitu data pencetakan KTP, dengan nama mahasiswa sebagai berikut:


Nama : Akmal Mugni Fawwazrin
NIM : 2011102441166
Program Studi : Teknik Informatika

Demikian surat permohonan ini dibuat. Atas perhatian dan kerjasamanya kami mengucapkan terima kasih.

Wassalamu'alaikum Warrahmatullahi Wabarrakatuh

Samarinda, 23 Dzulhijjah 1445 H
30 Juni 2024 M

Ketua Program Studi S1 Teknik Informatika

Arban Yan, S.Kom., M.TI
NIDN. 1118019203



Kampus 1 : Jl. Ir. H. Juanda, No.15, Samarinda
Kampus 2 : Jl. Pelita, Pesona Mahakam, Samarinda

Lampiran 2 Lembar Bimbingan

Dokumen Revisi dan Tindak Lanjut Peserta Ujian Sidang Hasil

Nama : Akmal Mugni Fawwazrin

NIM : 2011102441166

No	Nama Dosen	Permintaan Revisi	Tindak lanjut	Halaman
1	Fendy Yulianto	Memperbaiki judul	Menyesuaikan judul dengan komentar yang diberikan	Cover
2	Fendy Yulianto	Penulisan paragraf latar belakang tidak konsisten, setiap paragraf minimal 6 baris, maksimal 8 baris	Memperbaiki setiap paragraf latar belakang agar lebih konsisten	1
3	Fendy Yulianto	Setiap awal kalimat paragraf baru, harus nyambung dengan paragraf sebelumnya	Memperbaiki awal kalimat pada paragraf baru agar nyambung dengan paragraf sebelumnya	1
4	Fendy Yulianto	Setiap akhir kalimat pada latar belakang harus memiliki referensi, kecuali paragraf terakhir	Menambahkan referensi pada setiap akhir kalimat pada latar belakang	1
5	Fendy Yulianto	Berikan kalimat pengantar pada bawah sub bab, minimal 2 baris	Menambahkan kalimat pengantar pada setiap sub bab	4
6	Fendy Yulianto	Memasukan penjelasan MAE kedalam bagian evaluasi model	Memasukan penjelasan MAE kedalam bagian evaluasi model	7
7	Fendy Yulianto	Jelaskan berapa data yang digunakan dan berikan sample datanya	Menambahkan jumlah data yang digunakan dan memberikan gambar sampel data yang digunakan	6
8	Fendy Yulianto	Bagaimana cara menghilangkan <i>missing value</i>	Memberikan penjelasan cara menghilangkan <i>missing value</i>	6
9	Fendy Yulianto	Memperhatikan penulisannya terkait typo dan lainnya	Memperbaiki penulisan terkait typo dan lainnya	Seluruh Halaman
10	Abdul Rahim	Menambahkan studi kasus pada judul	Menambahkan studi kasus pada judul	Cover
11	Abdul Rahim	Menjelaskan sumber data yang diambil	Menjelaskan lokasi dimana data diambil	5
12	Abdul Rahim	Jika hasil dan pembahasan digabung, buat pembahasan pada setiap sub bab	Memberikan pembahasan pada setiap gambar, tabel, serta hasil pada setiap sub bab	9-22

Menyetujui

Penguji 1



Fendy Yulianto, S.Kom., M.Kom
NIDN. 1102079402

Penguji 2



Abdul Rahim, S.Kom., Mcs
NIDN. 000904701

Lampiran 3 Kartu Kendali Bimbingan

KARTU KENDALI BIMBINGAN LAPORAN KARYA ILMIAH

Nama Mahasiswa : Akmal Mugni Fawwazrin
NIM : 2011102441166
Nama Dosen Pembimbing : Abdul Rahim, S.Kom., M.Cs
Judul Penelitian : PENERAPAN METODE REGRESI LINIER UNTUK PREDIKSI
WAKTU PENCETAKAN KTP DISDUKCAPIL KECAMATAN
SAMARINDA UTARA

No	Tanggal	Uraian Pembimbing	Paraf Dosen
1	17 Jan, 2024	Konsultasi terkait data yang akan digunakan	
2	7 Feb, 2024	Konsultasi terkait judul skripsi akan digunakan	
3	19 Feb, 2024	Konsultasi tentang model penelitian yang telah dibuat	
4	21 Mar, 2024	Bimbingan terkait pengumpulan proposal skripsi di website Simpel	
5	26 Apr, 2024	Konsultasi terkait implementasi metode yang digunakan dalam penelitian	
6	17 Mei, 2024	Konsultasi tahapan penelitian pada BAB III	
7	29 Mei, 2024	Memberikan arahan tentang tahapan penelitian	
8	30 Mei, 2024	Konsultasi Bab III dengan catatan menjelaskan secara mendetail tentang tahap penelitian pada bab 2	
9	4 Jun, 2024	Konsultasi Bab III dan IV serta penulisan jurnal	

Dosen Pembimbing



Abdul Rahim, S.Kom., M.Cs
NIDN. 0009047901

Mengetahui,
Ketua Program Studi



Arban Syah, S.Kom., M.TI
NIDN. 1118019203

Lampiran 4 Kode Pengujian Rasio Perbandingan

```
#fungsi untuk melakukan pembagian rasio
def train_and_evaluate(x, y, test_size):
    xtrain, xtest, ytrain, ytest = train_test_split(x, y,
test_size=test_size, random_state=42)
    reg = LinearRegression().fit(x, y)
    r2_train = reg.score(xtrain, ytrain)
    r2_test = reg.score(xtest, ytest)
    return r2_train, r2_test

# Menghitung nilai R^2 untuk rasio yang berbeda
r2_train_70, r2_test_30 = train_and_evaluate(x, y, 0.30)
r2_train_80, r2_test_20 = train_and_evaluate(x, y, 0.20)
r2_train_90, r2_test_10 = train_and_evaluate(x, y, 0.10)

# Membuat data rasio
ratios = ['70:30', '80:20', '90:10']
r2_train_scores = [r2_train_70, r2_train_80, r2_train_90]
r2_test_scores = [r2_test_30, r2_test_20, r2_test_10]

#Menampilkan hasil perbandingan rasio
akurasi = pd.DataFrame({'Perbandingan': ratios, 'Data Train':
r2_train_scores, 'Data Test': r2_test_scores})
akurasi
```

Lampiran 5 Kode Pengujian MAE

```
import matplotlib.pyplot as plt
from sklearn.metrics import mean_absolute_error

#Menghitung Nilai MAE
mae=mean_absolute_error(ytest,y_pred)

#Membuat Plot
fig = plt.figure(figsize=(14,5))
plt.title(f'Real vs Prediction - MAE {mae}', fontsize=10)
plt.plot(range(len(xtest)),ytest,color='red')
plt.plot(range(len(xtest)),y_pred,color='green')
plt.xlabel('Data', fontsize=10)
plt.ylabel('Waktu Selesai', fontsize=10)
plt.legend(labels=['Real', 'Prediction'], fontsize=10)
plt.grid(True)
plt.show()
```

Lampiran 6 Riwayat Hidup

RIWAYAT HIDUP



Penulis dilahirkan di Samarinda pada tanggal 14 September 2002 menjadi anak ketiga dari tiga bersaudara dari pasangan Abdul Muis dan Heliyana. Pada saat ini penulis tinggal di Jalan Juanda 1 Anggrek Biru No. 06, RT. 22, Kel. Air Putih, Kec. Samarinda Ulu. Pendidikan sekolah di TK Alam Permai tahun 2008, SD Negeri 025 Samarinda lulus tahun 2014, SMP Negeri 5 Samarinda lulus tahun 2017, Lalu melanjutkan pendidikan di SMA Negeri 3 Samarinda lulus pada tahun 2020. Penulis saat ini sedang menempuh pendidikan kuliah di Universitas Muhammadiyah Kalimantan Timur Kota Samarinda, jurusan yang diambil adalah Teknik Informatika dan belajar di gedung Fakultas Sains dan Teknologi, dan saat ini juga penulis sedang menjalankan penelitian skripsi.