

Penerapan Regresi Linear Ganda Untuk Memprediksi Hasil Nilai Kuesioner Mahasiswa Dengan Menggunakan Python

Muhammad Sholeh¹, Rr. Yuliana Rachmawati², Eko Nur Cahyo³

^{1,2,3}Program Studi Informatika, Fakultas Teknologi Industri
Institut Sains & Teknologi AKPRIND Yogyakarta

e-mail: ¹muhash@akprind.ac.id, ²yuliana@akprind.ac.id,
³eko@akprind.ac.id

Intisari

Salah satu aktivitas yang rutin diselenggarakan pada proses akademik adalah pengisian kuesioner yang dilakukan mahasiswa. Salah satu item dalam kuesioner adalah umpan balik untuk dosen dalam melaksanakan proses pembelajaran. Mahasiswa mengisi form yang sudah disediakan. Dan hasil kuesioner ini akan menghasilkan nilai kinerja dosen dari sisi kuesioner mahasiswa dengan rentang angka 0-4. Penelitian yang ini melakukan pemrosesan analisis prediksi dengan menggunakan regresi linear. *Datasheet* diolah sesuai dengan data-data yang diperlukan. Data yang digunakan merupakan data hasil kuesioner dari data tahun 2016-2020. *Datasheet* terdiri dari 7 kolom dan 800 baris. Dari *datasheet* akan dilakukan pemisahan data. Dari data yang ada, sebanyak 80% akan digunakan untuk *data training* dan *data test* sebanyak 20%. Metode penelitian menggunakan metode *waterfall*. Penelitian diawali dengan analisis sistem dengan melakukan pencarian dan pengumpulan *datasheet*, desain sistem dilakukan dengan pemisahan data-data yang tidak diperlukan serta koding dan implementasi. Implementasi dalam prediksi hasil kuesioner akan menggunakan Python dengan menggunakan berbagai *library*, diantaranya Pandas yang digunakan analisis data, Matplotlib digunakan visualisasi data dan Scikit-Learn digunakan dalam proses *machine learning*. Proses koding menggunakan Google Colab. Hasil penelitian menghasilkan rumus regresi linear $y = 0.25 X_1 + 0.26 X_2 + 0.27 X_3 + 0.23 X_4 + 0.0001 X_5 - 1e-06 X_6$ dengan nilai koefisien adalah 0.997. Penggunaan Python dengan dukungan *library* memberikan kemudahan dalam implementasi *machine learning* khusus regresi linear.

Kata kunci—regresi linear, python, google colab, data

Abstract

One of the activities that are routinely held in the academic process is filling out questionnaires by students. One of the items in the questionnaire is feedback for lecturers in carrying out the learning process. Students fill out the form provided. And the results of this questionnaire will produce a lecturer's performance value from the student questionnaire side with a range of numbers 0-4. This research is processing prediction analysis using linear regression. The datasheet is processed according to the required data. The data used is the data from the questionnaire from the 2016-2020 data. The datasheet consists of 7 columns and 800 rows. From the datasheet, the data will be separated. From the existing data, 80% will be used

for training data and 20% for test data. The research method uses the waterfall method. The research begins with system analysis by searching and collecting datasheets, system design is carried out by separating unnecessary data and coding and implementation. The implementation in predicting the results of the questionnaire will use Python using various libraries, including Pandas which is used for data analysis, Matplotlib is used for data visualization and Scikit-Learn is used in the machine learning process. Coding process using Google Colab. The results of the study resulted in a linear regression formula $y = 0.24743301702419157 X_1 + 0.257319654680142 X_2 + 0.2662633268986068 X_3 + 0.2332290731664389 X_4 + 9.697801443559612e-05 X_5 - 1.4522339706890541e-06 X_6$ with a coefficient value of 0.997. Using Python with library support makes it easy to implement machine learning specifically for linear regression

Keywords— *linear regression, python, google colab, data*

PENDAHULUAN

Evaluasi kinerja dosen menjadi salah satu parameter yang dievaluasi di setiap semester. Elemen yang digunakan dalam proses penilaian kinerja dosen terdiri dari aktivitas pada bidang pendidikan (mengajar, membuat buku), bidang penelitian (melakukan penelitian atau publikasi ilmiah) dan bidang pengabdian pada masyarakat. Dalam penilaian aspek akademik, salah satu parameter adalah dengan memberikan kuesioner pada mahasiswa. Kuesioner wajib diisi setiap mahasiswa di setiap semester. Pengisian kuesioner dilakukan pada mahasiswa yang mengambil mata kuliah. Hasil kuesioner diolah sesuai dengan rumus yang sudah ditentukan dan proses kuesioner ini dilakukan sebagai upaya untuk menilai kinerja dosen.

Perguruan tinggi sebagai lembaga pendidikan yang mndidik mahasiswa menjadi lulusan bermutu harus selalu memperhatikan sistem serta proses pendidikan. Sistem dan proses yang berjalan dengan baik tentunya akan menghasilkan produk yang berkualitas. [1]. Salah satu faktor utama agar sistem dan proses berjalan dengan baik adalah dosen. Dalam proses melaksanakan fungsi dan tugas pendidikan dengan baik, dosen harus memahami teori-teori pendidikan serta metode yang digunakan dalam proses perkuliahan, Dosen dalam proses perkuliahan harus menjadi sumber belajar, fasilitator dan mediator serta evaluator dalam menentukan output perguruan tinggi. [2]. Di samping unsur pendidikan, kinerja dosen juga dinilai dari peran aktif dalam menghasilkan penelitian dan pelaksanaan pengabdian pada masyarakat. [3]

Di setiap perguruan tinggi parameter penilaian yang tidak sama tetapi mempunyai maksud dan tujuan yang sama yaitu menilai kinerja dosen dari sisi proses perkuliahan. [4], melakukan evaluasi terhadap penilaian kuesioner yang diisi mahasiswa. Kuesioner diberikan dengan sifat tertutup dan berdasar hasil kuesioner tersebut akan dinilai kepuasan mahasiswa terhadap dosen. [5], melakukan analisis hasil kuesioner yang diisi mahasiswa, khususnya proses

pembelajaran yang berlangsung. Salah satu hasil penelitian adalah adanya korelasi antara nilai kuesioner yang diisi mahasiswa dengan prestasi mahasiswa. Variabel yang diteliti diantaranya adalah kedisiplinan dosen selama proses perkuliahan. Prestasi mahasiswa meningkat dapat dipengaruhi dengan nilai yang tinggi yang dicapai dosen. Hal ini menunjukkan jika proses perkuliahan berjalan dengan baik, maka prestasi mahasiswa dalam perkuliahan juga baik. [6],

Hasil kuesioner yang didapat akan diolah sesuai dengan tujuan yang diinginkan. Hasil kuesioner diantaranya dapat digunakan dalam evaluasi kinerja dosen dan hasil evaluasi tersebut dapat berupa indeks prestasi kinerja dosen. Proses olah data hasil kuesioner dilakukan dengan menggunakan nilai atau rumus yang sudah menjadi peraturan. Di samping olah data tersebut dapat juga dilakukan dengan menggunakan model regresi linear terutama regresi linear ganda. Analisis dengan menggunakan regresi linear berganda merupakan salah satu metode statistika yang dapat digunakan untuk mengetahui keterkaitan hubungan fungsional sebuah variabel tidak bebas dengan dua atau lebih variabel bebas. Analisis regresi linier berganda dapat digunakan untuk mengetahui pengaruh beberapa variabel bebas terhadap variabel tidak bebas dan dapat digunakan untuk melakukan prediksi nilai variabel dependen (tidak bebas) jika semua variabel independen (bebas) yang digunakan nilai variabel tersebut diketahui. Persamaannya regresi linear ganda adalah $y = a_1x_1 + a_2x_2 + \dots + a_nx_n + b$. [7].

Proses perhitungan regresi linear dapat menggunakan berbagai software, diantaranya menggunakan SPSS, Minitab dan Python. Penelitian yang melakukan olah data regresi linear dengan menggunakan SPSS dilakukan [8]. Penelitian [8] mengupas penggunaan regresi linier berganda yang digunakan untuk melakukan estimasi dalam produktivitas tanaman padi di kota Karawang. [9], menggunakan SPSS untuk melakukan analisis regresi linier berganda dalam menyelesaikan skripsi. Penelitian lain yang menggunakan SPSS dilakukan [9], [10],[11],[12].

Aplikasi lain yang digunakan adalah dengan menggunakan pemrograman Python. Dukungan *library* menjadi salah satu penggunaan Python dalam proses analisis regresi linear berganda semakin mudah. Proses implementasi regresi linear dengan python dapat menggunakan berbagai editor, salah satunya adalah dengan menggunakan google colab. [13].

[14] melakukan penelitian dengan Python untuk melakukan analisis menggunakan metode regresi linear ganda. Data yang diolah adalah data yang digunakan untuk melakukan prediksi harga sembako. Proses prediksi dilakukan dengan membagi data menjadi data untuk training dan data untuk test. Hasil penelitian diantaranya adalah aplikasi yang dapat melakukan prediksi harga sembako harian. Aplikasi dikembangkan menggunakan bahasa pemrograman python sebagai *backend*, dan *web framework* django. [15] melakukan penelitian regresi linear untuk menentukan kelas yang dapat digunakan untuk mengetahui rentang jumlah jenis obat tertentu dan melihat jenis obat yang paling banyak digunakan dan yang paling sedikit digunakan pada rentang bulan tertentu. Penelitian lain yang mengupas regresi linear dilakukan [16], [17], [18],[19].

Berdasar pada latar belakang dan tinjauan pustaka, olah data untuk melakukan prediksi menjadi hal yang penting dalam melakukan analisis data. Hasil kuesioner terdapat kinerja dosen yang dilakukan mahasiswa semestinya tidak hanya

sebagai laporan untuk kinerja dosen tetapi juga dilakukan olah data yang terkait dengan prediksi dan keterkaitan antar data. Hasil kuesioner yang didapat diolah dengan menggunakan analisis regresi linear berganda dengan menggunakan pemrograman Python dan menggunakan editor Google Colab. Batasan masalah diantaranya data yang diolah merupakan data hasil kuesioner mulai tahun 2016-2020.

METODE PENELITIAN

Studi Literatur

Langkah awal dalam penelitian ini adalah melakukan studi literatur dengan melakukan pengumpulan data yang terkait dengan kuesioner mahasiswa. Proses studi literatur dilakukan dengan melakukan mempelajari dan memahami materi yang berkaitan dengan tema regresi linear dan machine learning khususnya regresi linear. Proses lain adalah dengan mencari sumber-sumber studi literatur yang digunakan sebagai studi pustaka.

Metode Waterfall

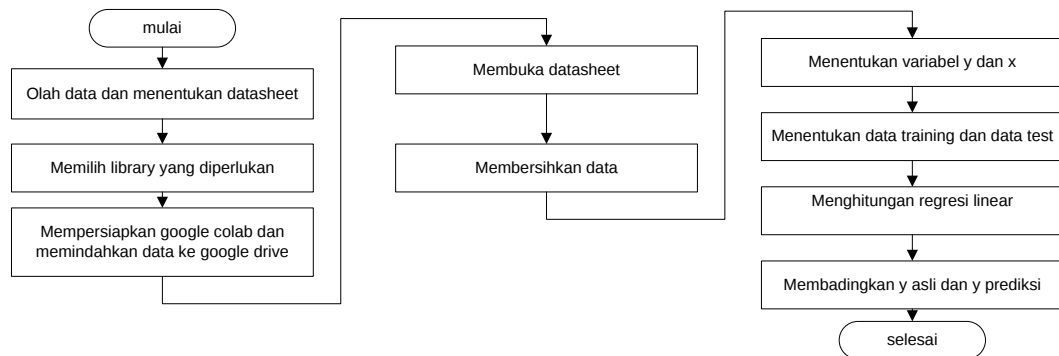
Proses koding analisis regresi linear menggunakan pemrograman Python dan menggunakan Google Collab sebagai editor. Proses pemrograman dilakukan dengan menggunakan *metode waterfall*. Dalam metode ini proses diawali dengan melakukan analisis data, desain data, koding , evaluasi dan implementasi.

Library yang digunakan

Library yang digunakan dalam proses perhitungan regresi linear harus diinstall terlebih dahulu agar dapat digunakan. *Library* yang digunakan diantaranya numpy, pandas, matplotlib, seaborn, sklearn, statsmodels, drive

Langkah-langkah Penelitian

Langkah penelitian diawali dengan pengumpulan data hasil kuesioner dan disimpan dataset dalam bentuk file csv. Sebelum dilakukan pengolahan pada dataset, library yang diperlukan dilakukan proses instalasi. Dari libray yang diperlukan dan sudah dipasang, proses dilakukan proses pengecekan datasheet dari data-data yang mengandung kesalahan seperti data yang kosong. Datasheet yang sudah tidak ada kesalahan dan untuk menvari regresi linear, datasheet akan dibedakan menjadi data x dan data y serta menentukan data yang menjadi data training dan data yang menjadi data test. Hasil dari data training akan menghasilkan rumus regresi linear dan untuk menguji prediksi menggunakan data test. Hasil data test ini akan ditemukan persentasi keakuratan rumus regresi linear. Langkah penelitian lebih detail ada pada gambar 1.



Gambar 1 Langkah Penelitian

HASIL DAN PEMBAHASAN

Persiapan data

Hasil proses pengolahan data yang digunakan disimpan dalam bentuk file CSV. Dari data CSV ini akan dilakukan proses pembersihan data yang tidak digunakan. Gambar 2, file CSV yang disimpan dalam Microsoft Excel. Data yang diolah sebanyak 800 baris dan 7 kolom.

	A	B	C	D	E	F	G
1	Aspek1	Aspek2	Aspek3	Aspek4	IPK_respo	jumlah	Rate-rata
2	2.88	2.93	2.87	2.93	3.76	99	2.9
3	2.87	3.02	2.97	3	3.7	43	2.96
4	2.72	2.91	2.94	2.95	3.37	112	2.88
5	2.98	2.97	2.97	2.97	2.9	40	2.98
6	3.21	3.12	3.12	3.12	2.8	26	3.14
7	2.95	2.99	2.99	2.99	5.26	184	2.98
8	2.97	2.95	2.96	2.93	3.06	77	2.95
9	3.16	3.06	3.05	3.05	5.91	87	3.08
10	3.15	3.01	3	3	3.29	81	3.04
11	2.94	2.95	2.95	3.01	3.07	109	2.96
12	3.06	3.09	3.06	3.04	5.1	84	3.06
13	2.99	3.01	3	2.99	3.38	115	3
14	2.88	2.99	3	2.99	3.6	77	2.97
15	3.06	3.03	3.03	3.06	5.99	80	3.04
16	2.99	2.99	2.99	2.99	2.86	63	2.99
17	3.27	3.15	3.15	3.15	4.28	208	3.19
18	3.1	3.03	3.06	3.05	5.05	216	3.06
19	3.11	3.04	3.04	3.01	3.24	55	3.05
20	3.14	3.04	3.06	3.06	9.25	126	3.08
21	2.95	2.94	2.89	2.95	2.98	224	2.93
22	2.96	2.95	2.98	3	3.76	150	2.97
23	3.15	3.07	3.06	3.07	6.07	182	3.09

Gambar 2 Data CSV

Menyiapkan library dan data

Proses awal dalam melakukan perhitungan analisis regresi berganda dengan menggunakan bahasa pemrograman python adalah menentukan *library* yang digunakan dalam proses analisis dan memindahkan data ke *drive*. Langkah-langkahnya adalah

1. Library yang digunakan adalah,


```

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
from sklearn.preprocessing import LabelEncoder
      
```

```

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error, mean_squared_log_error, r2_score
import warnings
warnings.filterwarnings("ignore")

```

2. Memindahkan data dari *drive* komputer ke drive google.

Proses ini diperlukan karena penggunaan google colab harus terkoneksi dengan internet dan data yang ada harus disimpan di drive google. File data2.csv yang disimpan di komputer lokal akan diunggah di drive dengan nama file yang sama.

```

from google.colab import files
uploaded = files.upload()
with open('data2.csv', 'r') as data:
    df3 = pd.read_csv(data, encoding = ('ansi'))
    print(df3)

```

Hasil proses running skrip di atas ditampilkan pada gambar 3

```

import pandas as pd
from google.colab import files
uploaded = files.upload()
with open('data2.csv', 'r') as data:
    df3 = pd.read_csv(data, encoding = ('ansi'))
    print(df3)

```

Pilih File data2.csv

- data2.csv(application/vnd.ms-excel) - 29690 bytes, last modified: 17/1/2022 - 100% done

Saving data2.csv to data2 (2).csv

	Aspek1	Aspek2	Aspek3	Aspek4	IPK_responden	jumlah_responden	Rata-rata
0	2.88	2.93	2.87	2.93	3.76	99.0	2.90
1	2.87	3.02	2.97	3.00	3.70	43.0	2.96
2	2.72	2.91	2.94	2.95	3.37	112.0	2.88
3	2.98	2.97	2.97	2.97	2.90	40.0	2.98
4	3.21	3.12	3.12	3.12	2.80	26.0	3.14
...	...	...	...	...	...	...	...
795	3.17	3.00	3.00	3.00	2.97	4.0	3.04
796	2.89	3.00	3.00	3.00	2.41	6.0	2.97
797	3.00	3.00	3.00	3.00	3.94	1.0	3.00
798	3.00	3.00	3.00	3.00	3.10	4.0	3.00
799	3.00	3.00	3.00	3.00	2.67	8.0	3.00

[800 rows x 7 columns]

Gambar 3 Hasil proses running data yang digunakan

Membersihkan data

Pada tahapan ini dilakukan pengecekan terhadap isi datasheet yang digunakan, Proses pengecekan diantaranya pengecekan data kosong, tipe data, data yang sama, outliers yaitu data yang nilainya terlalu jauh dengan data lainnya dan proses membersihkan data lainnya. Gambar 4, contoh proses dan hasil proses untuk melakukan data kosong

```
#Mencari dan menangani missing values.
df3.isnull().sum()

Aspek1          2
Aspek2          2
Aspek3          2
Aspek4          2
IPK_responden   2
jumlah_responden 2
Rata-rata       2
dtype: int64
```

Gambar 4, Hasil Pencarian data kosong

Hasil pada gambar 4 terdapat 2 data yang kosong dan data kosong ini akan dihapus. Gambar 5 proses penghapusan data dan proses pengecekan kembali apakah ada data kosong.

```
[ ] df3=df3.dropna()

df3.isnull().sum()

Aspek1          0
Aspek2          0
Aspek3          0
Aspek4          0
IPK_responden   0
jumlah_responden 0
Rata-rata       0
dtype: int64
```

Gambar 5 Perintah dan hasil penghapusan data

Menentukan variabel y dan variabel x

Proses awal adalah membedakan data yang menjadi variabel dependen dan variabel independen. Kolom yang menjadi variabel y adalah rata-rata dan variabel x adalah aspek1, aspek2, aspek3, aspek4, rata-rata IPK responden dan jumlah respon

```
#Menentukan variabel x dan y.
x = df3.drop(columns='Rata-rata')
y = df3['Rata-rata']
```

Menyiapkan Data training dan data test

Data sekitar 800 data yang diolah, data tersebut dibagi 80% menjadi data training dan 20% digunakan untuk data test. Data training digunakan untuk menghitung prediksi regresi berganda dan data test digunakan untuk perhitungkan data prediksi dan dapat dibandingkan antara data asli dengan data hasil prediksi

```
# Melakukan proses split data training dan testing dengan komposisi 80:20
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=4)
```

Korelasi Antar Data

Analisis korelasi digunakan untuk melihat keterkaitan tentang derajat hubungan variabel sehingga dapat mengetahui hubungan variabel yang ada. Gambar 6 Perintah dan hasil analisis korelasi data

```
df3.corr().style.set_precision(3)
```

	Aspek1	Aspek2	Aspek3	Aspek4	IPK_responden	jumlah_responden	Rata-rata
Aspek1	1.000	0.481	0.467	0.456	0.035	-0.016	0.856
Aspek2	0.481	1.000	0.955	0.935	0.035	-0.023	0.852
Aspek3	0.467	0.955	1.000	0.947	0.037	-0.009	0.847
Aspek4	0.456	0.935	0.947	1.000	0.035	-0.005	0.836
IPK_responden	0.035	0.035	0.037	0.035	1.000	0.452	0.041
jumlah_responden	-0.016	-0.023	-0.009	-0.005	0.452	1.000	-0.016
Rata-rata	0.856	0.852	0.847	0.836	0.041	-0.016	1.000

Gambar 6 Perintah dan hasil analisis korelasi data

Menghitung hasil regresi linear

Langkah selanjutnya adalah menghitung regresi linear. Library yang digunakan adalah library sklearn. Gambar 7, proses penghitungan nilai koefisien.

```
[31] #mencari cari tau nilai intercept (b).
print(lin_reg.intercept_)

-0.012897918703604017

[33] #mencari nilai koefisien dan menampilkan hasilnya
feature_cols = ['Aspek1', 'Aspek2', 'Aspek3', 'Aspek4', 'IPK_responden', 'jumlah_responden']
X = df3[feature_cols]
y = df3['Rata-rata']
list(zip(feature_cols, lin_reg.coef_))

[('Aspek1', 0.24743301702419157),
 ('Aspek2', 0.257319654680142),
 ('Aspek3', 0.2662633268986068),
 ('Aspek4', 0.2332290731664389),
 ('IPK_responden', 9.697801443559612e-05),
 ('jumlah_responden', -1.4522339706890541e-06)]
```

Gambar 7, proses penghitungan nilai koefisien.

Berdasar hasil pada perhitungan di atas, nilai regresi linear adalah :
 $y = 0.25 X_1 + 0.26 X_2 + 0.27 X_3 + 0.23 X_4 + 0.0001 X_5 - 1e-06 X_6$

Evaluasi untuk model linear regresi

Evaluasi untuk model linear regresi dapat dilakukan dengan beberapa cara evaluasi seperti menggunakan *mean absolute error (MAE)*, *mean squared error (MSE)* atau *root mean squared error (RMSE)* . Gambar 8 merupakan hasil perhitungan MSE dan RMSE

```
[ ] ypredict=lin_reg.predict(x_test)
print ("Coefficient of determination :",r2_score(y_test,ypredict))
print ("MSE: ",mean_squared_error(y_test,ypredict))
print("RMSE: ",np.sqrt(mean_squared_error(y_test,ypredict)))

Coefficient of determination : 0.9979745042302917
MSE: 8.352756861727575e-06
RMSE: 0.002890113641663174
```

Gambar 8 Hasil perhitungan MSE dan RMSE

Data Asli dan data prediksi

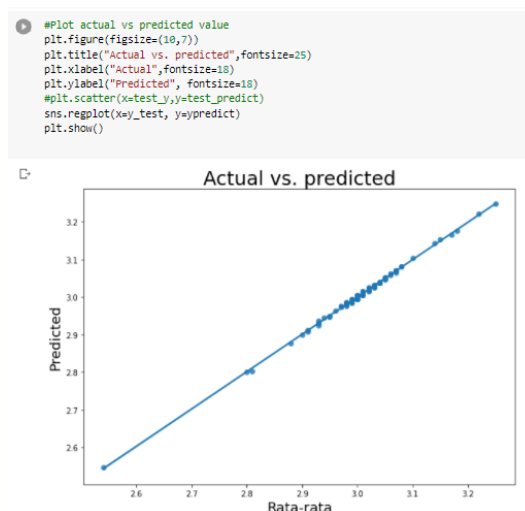
Pada tahap ini akan dilakukan perbandingan antara nilai y awal dengan nilai y prediksi. Nilai y prediksi dihitung pada rumus regresi linear berganda yang telah dihitung pada langkah di atas. Gambar 9 Proses dan hasil perhitungan proses membandingkan data asli dengan data prediksi.

```
df_best_predict = pd.DataFrame({'Actual': y_test, 'Predicted': ypredict})
df_best_predict.head(10)
```

	Actual	Predicted
267	3.15	3.152753
637	3.04	3.040133
464	2.96	2.962937
721	3.00	3.000125
796	2.97	2.972845
669	3.02	3.017842
791	3.00	3.000167
504	2.99	2.990139
539	3.01	3.010014
147	3.01	3.015032

Gambar 9 Hasil perhitungan membandingkan data asli dengan data prediksi

Hasil proses pada gambar 9 dapat dibuat dalam bentuk grafik. Gambar 10. Hasil y asli dengan y prediksi dalam bentuk grafik. Hasil grafik memperlihatkan antara data asli dan data prediksi menghasilkan selisih angka antara data aktual dan data prediksi yang sedikit



Gambar 10. Grafik y asli dengan y prediksi

Perhitungan dengan *library statsmodels*

Library lain yang dapat digunakan untuk melakukan analisis regresi linear adalah dengan menggunakan *library statsmodels*. Hasil perhitungan dengan *library statsmodels* ada pada gambar 11

```
[ ] import statsmodels.api as sm
x = x_train.to_numpy()
y = y_train.to_numpy()
x = sm.add_constant(x)
model = sm.OLS(y, x).fit()
print(model.summary())
```

OLS Regression Results

Dep. Variable:	y	R-squared:	0.997
Model:	OLS	Adj. R-squared:	0.997
Method:	Least Squares	F-statistic:	3.794e+04
Date:	Tue, 18 Jan 2022	Prob (F-statistic):	0.00
Time:	06:43:40	Log-Likelihood:	2828.6
No. Observations:	638	AIC:	-5643.
Df Residuals:	631	BIC:	-5612.
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	-0.0129	0.008	-1.687	0.092	-0.028	0.002
x1	0.2474	0.001	184.485	0.000	0.245	0.250
x2	0.2573	0.009	29.622	0.000	0.240	0.274
x3	0.2663	0.010	26.449	0.000	0.246	0.286
x4	0.2332	0.008	28.395	0.000	0.217	0.249
x5	9.698e-05	0.000	0.760	0.447	-0.000	0.000
x6	-1.452e-06	2.01e-06	-0.723	0.470	-5.4e-06	2.49e-06

Omnibus: 3.278 Durbin-Watson: 1.964
 Prob(Omnibus): 0.194 Jarque-Bera (JB): 2.910
 Skew: 0.088 Prob(JB): 0.233
 Kurtosis: 2.720 Cond. No. 9.97e+03

Warnings:
 [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
 [2] The condition number is large, 9.97e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Gambar 11. Hasil perhitungan dengan *library Statsmodel*

KESIMPULAN

Penggunaan Google colab sangat membantu dalam proses perhitungan regresi linear. Google colab digunakan secara online, sehingga dapat dilakukan kplaborasi dengan tim yang lain. Dalam penelitian ini data yang digunakan disimpan dalam drive dan proses harus dilakukan secara online. Penelitian menghasilkan rumus regresi linear dengan nilai dengan nilai koefisien adalah 0.997. Hasil analisis korelasi antara variabel jumlah korespon dengan variabel yang lain menunjukan nilai negatif. Hal menunjukkan variabel jumlah respondn tidak ada kaitannya dengan variabel yang lain.

SARAN

Penggunaan google colab cukup mudah dalam penggunaan tetapi kelemahannya diantaranya aplikasi yang dijalankan tidak berbasis aplikasi yang pengguna dapat langsung menjalankannya. Pengembangan berikutnya adalah

menggunakan library grafis sehingga tampilan dan hasil lebih menarik. *Library* yang dapat digunakan untuk tampilan berbasis grafis diantaranya Pyqt, Tkinter ,

DAFTAR PUSTAKA

- [1] M. S. Sholehuddin, A. Akhwanudin, and U. Khasanah, *PENGELOLAAN KINERJA DOSEN DAN BUDAYA AKADEMIK*. Penerbit NEM, 2018.
- [2] N. N. Dewi, *KIAT-KIAT MERANGSANG KINERJA DOSEN PTS. MEDIA SAHABAT CENDEKIA*, 2019.
- [3] U. Haryaka, *Faktor-faktor yang Mempengaruhi Kinerja Dosen*. CV. Amerta Media, 2021.
- [4] T. Sulastri, "Analisis Kepuasan Mahasiswa Terhadap Kinerja Dosen," *Optimal: Jurnal Fakultas Ekonomi Universitas Islam "45" Bekasi*, vol. 10, no. 2, pp. 167–184, 2017.
- [5] B. D. Purba, "Analisa Korelasi Kinerja Dosen Terhadap Prestasi Mahasiswa Dengan Menggunakan Logika Fuzzy," *Jurnal Teknik Informatika Unika St. Thomas (JTIUST)*, vol. 02, no. 479, pp. 64–71, 2017.
- [6] K. Sa'adah and A. Febrianto, "ANALISA KINERJA DOSEN BERDASARKAN TINGKAT KEPUASAN MAHASISWA DI POLITEKNIK MANUFAKTUR BANDUNG," *JIMEA | Jurnal Ilmiah MEA (Manajemen, Ekonomi, dan Akuntansi)*, vol. 5, no. 1, pp. 1339–1351, 2021, doi: 10.31955/mea.vol3.iss2.p.
- [7] D. S. SUSANTI, Y. SUKMAWATY, and N. SALAM, *ANALISIS REGRESI DAN KORELASI*. IRDH, 2019.
- [8] T. N. Padilah and R. I. Adam, "Analisis regresi linier berganda dalam estimasi produktivitas tanaman padi di kabupaten karawang 1,2)," *Fibonacci Jurnal Pendidikan Matematika dan Matematika*, vol. 5, no. 2, pp. 117–128, 2019.
- [9] A. Wibisono, M. Rofik, and E. Purwanto, "Penerapan Analisis Regresi Linier Berganda dalam Penyelesaian Skripsi Mahasiswa," *Jurnal ABDINUS: Jurnal Pengabdian Nusantara*, vol. 3, no. 1, p. 30, 2019, doi: 10.29407/ja.v3i1.13512.
- [10] I. Sonata, Nuraida, and E. Sofian, "Pengaruh Kepribadian Dan Continuance Commitment Terhadap Kinerja Pegawai Di Kantor Dinas Kelautan Dan Perikanan Provinsi Sumatera Utara," *Tijarah: Jurnal Ekonomi dan Bisnis*, vol. 1, no. 21, pp. 12–22, 2021.
- [11] E. Karamazova, T. Jusufi Zenku, and Z. Trifunov, "Analysing and Comparing the Final Grade in Mathematics by Linear Regression Using Excel and SPSS," *International Journal of Mathematics Trends and Technology*, vol. 52, no. 5, pp. 334–344, 2017, doi: 10.14445/22315373/ijmtt-v52p549.
- [12] S. N. Waghmare and C. N. Sakhale, "Formulation of Experimental Data Based model using SPSS (Linear Regression) for Stirrup Making Operation by Human Powered Flywheel Motor," *International Research Journal of Engineering and Technology (IRJET)*, vol. 02, no. 04, pp. 461–468, 2015.
- [13] J. Enterprise, *Mengolah Data dengan Python dan Pandas*. Jakarta: Elex

- Media Komputindo, 2021.
- [14] K. Puteri and A. Silvanie, "MACHINE LEARNING UNTUK MODEL PREDIKSI HARGA SEMBAKO," *Jurnal Nasional Informatika*, vol. 1, no. 2, pp. 82–94, 2020.
 - [15] A. R. Shidiq, I. Ermawati, M. M. Santoni, J. Barat, O. Keluar, and R. Linier, "Penerapan Algoritma Regresi Linier Untuk Prediksi Pada," pp. 102–111, 2020.
 - [16] E. Triyanto, H. Sismoro, and A. D. Laksito, "IMPLEMENTASI ALGORITMA REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI PRODUKSI PADI DI KABUPATEN BANTUL," *RABIT : Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 4, no. 2, pp. 73–86, 2019.
 - [17] A. S. T. Nishadi, "Predicting Heart Diseases In Logistic Regression Of Machine Learning Algorithms By Python Jupyterlab," *International Journal of Advanced Research and Publications*, vol. 3, no. 8, pp. 69–74, 2019.
 - [18] S. C. Agrawal, "Deep learning based non-linear regression for Stock Prediction," *IOP Conference Series: Materials Science and Engineering*, vol. 1116, no. 1, p. 012189, 2021, doi: 10.1088/1757-899x/1116/1/012189.
 - [19] Sudaryatno *et al.*, "Multiple linear regression analysis of remote sensing data for determining vulnerability factors of landslide in PURWOREJO," *IOP Conference Series: Earth and Environmental Science*, vol. 500, no. 1, 2020, doi: 10.1088/1755-1315/500/1/012046.