

# Tugas 1: Laporan Praktikum Tugas Mandiri

Aan Adriyana - 0110224014

<sup>1</sup> Teknik Informatika, STT Terpadu Nurul Fikri, Depok

\*E-mail: [0110224014@student.nurulfikri.ac.id](mailto:0110224014@student.nurulfikri.ac.id)

**Abstract.** Regresi linier adalah alat statistik dasar untuk memodelkan hubungan antara variabel dependen dan satu atau lebih variabel independen. Regresi linier banyak digunakan dalam ilmu data dan pembelajaran mesin untuk memprediksi hasil dan memahami hubungan antar variabel. Dalam Python, implementasi regresi linier dapat dilakukan dengan mudah menggunakan pustaka pihak ketiga seperti scikit-learn dan statsmodels.

## 1. Prediksi dari kasus Dataset

### 1.1 Import Library dan Load Dataset

```
import pandas as pd

path = "/content/drive/MyDrive/praktikum_ml/praktikum03"

# Baca dataset
df = pd.read_csv(path + "/data/day.csv")

# Lihat 5 data pertama
print(df.head())
```

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	\
0	1	2011-01-01	1	0	1	0	6	0	
1	2	2011-01-02	1	0	1	0	0	0	
2	3	2011-01-03	1	0	1	0	1	1	
3	4	2011-01-04	1	0	1	0	2	1	
4	5	2011-01-05	1	0	1	0	3	1	

  

	weathersit	temp	atemp	hum	windspeed	casual	registered	\
0	2	0.344167	0.363625	0.805833	0.160446	331	654	
1	2	0.363478	0.353739	0.696087	0.248539	131	670	
2	1	0.196364	0.189405	0.437273	0.248309	120	1229	
3	1	0.200000	0.212122	0.590435	0.160296	108	1454	
4	1	0.226957	0.229270	0.436957	0.186900	82	1518	

  

	cnt
0	985
1	801
2	1349
3	1562
4	1600

### 1.2 Pisahkan X dan Y

```
# Pilih fitur (independen)
X = df[['season', 'yr', 'mnth', 'holiday', 'weekday', 'workingday',
        'weathersit', 'temp', 'atemp', 'hum', 'windspeed']]

# Target (dependen)
y = df['cnt']
```

### 1.3 Split Data Untuk Training dan Testing

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)
```

### 1.4 Buat model Prediksi (Contoh Linear Regression)

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
import numpy as np

# Buat model
model = LinearRegression()
model.fit(X_train, y_train)

# Prediksi
y_pred = model.predict(X_test)

# Evaluasi
print("R2 Score:", r2_score(y_test, y_pred))
print("RMSE:", np.sqrt(mean_squared_error(y_test, y_pred)))
```

```
R2 Score: 0.8276670090367212
RMSE: 831.2851545662686
```

#### Link Colab:

[https://colab.research.google.com/drive/1ml\\_JoZgsUCDN8BuX8CvB0GeClgNqo3u?usp=sharing](https://colab.research.google.com/drive/1ml_JoZgsUCDN8BuX8CvB0GeClgNqo3u?usp=sharing)

**Referensi:**

<https://realpython.com/linear-regression-in-python/>