

# Tugas 1: Introduction Machine Learning – Read CSV file using pandas library

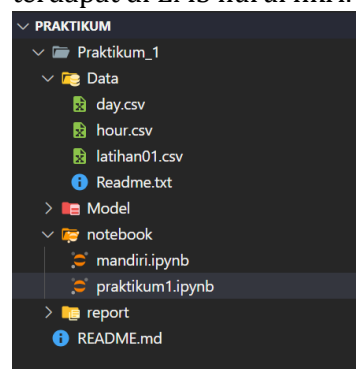
Syayid Muhammad Akbar - 0110222086 <sup>1\*</sup>

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

\*E-mail: [syay22086@student.nurulfikri.ac.id](mailto:syay22086@student.nurulfikri.ac.id)

## 1. Read CSV in python using pandas (Praktikum)

Langkah pertama untuk melakukan analisis data atau pemrosesan data menggunakan machine learning adalah membaca data itu sendiri, untuk tugas 1 pada praktikum machine learning ini menggunakan data yang terdapat di LMS nurul fikri.



Gambar 1. Contoh direktori

Setelah mendownload data yang dibutuhkan, dan load data kedalam code editor VScode atau collab, seperti contoh dibawah:

```
1 import pandas as pd
2 import numpy as np
3 import os
4
5 df = pd.read_csv('c:/Users/syay2/OneDrive/ドキュメント/NF/SEM 7/praktikum/Praktikum_1/data/day.csv')
6 df.head()
```

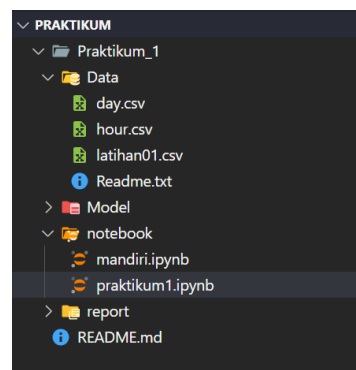
Lalu jalankan code tersebut, code tersebut akan menghasilkan:

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	reg
0	1	2011-01-01	1	0	1	0	6	0	2	0.344167	0.363625	0.805833	0.160446	331	
1	2	2011-01-02	1	0	1	0	0	0	2	0.363478	0.353739	0.696087	0.248539	131	
2	3	2011-01-03	1	0	1	0	1	1	1	0.196364	0.189405	0.437273	0.248309	120	
3	4	2011-01-04	1	0	1	0	2	1	1	0.200000	0.212122	0.590435	0.160296	108	
4	5	2011-01-05	1	0	1	0	3	1	1	0.226957	0.229270	0.436957	0.186900	82	

kode tersebut menghasilkan 731 baris (row) dan 16 kolom, data tersebut mengenai bike rental yang sudah diotomasi baik dari pendaftaran awal membership hingga pengembalian sepeda, data ini digunakan untuk analisis correlation antara bike rental process dengan environment seasonal.

## 2. Read CSV in python using pandas (Mandiri)

Langkah pertama untuk melakukan analisis data atau pemrosesan data menggunakan machine learning adalah membaca data itu sendiri, untuk tugas 1 pada praktikum machine learning ini menggunakan data yang terdapat di LMS nurul fikri.



setelah mendownload data yang dibutuhkan, lalu masukan data yang ada kedalam code editor, dengan menuliskan kode seperti :

```
1 import pandas as pd
2 import numpy as np
3 df = pd.read_csv('c:/Users/syay2/OneDrive/ドキュメント/NF/SEM 7/praktikum/Praktikum_1/data/hour.csv')
4 df.head()
```

Gambar 5. Load data mandiri dataset

Dari kodingan diatas menghasilkan, seperti berikut :

	instant	dteday	season	yr	mnth	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	reg
0	1	2011-01-01	1	0	1	0	6	0	2	0.344167	0.363625	0.805833	0.160446	331	
1	2	2011-01-02	1	0	1	0	0	0	2	0.363478	0.353739	0.696087	0.248539	131	
2	3	2011-01-03	1	0	1	0	1	1	1	0.196364	0.189405	0.437273	0.248309	120	
3	4	2011-01-04	1	0	1	0	2	1	1	0.200000	0.212122	0.590435	0.160296	108	
4	5	2011-01-05	1	0	1	0	3	1	1	0.226957	0.229270	0.436957	0.186900	82	

Gambar 6. Hasil load data

Data tersebut menghasilkan 17379 baris data dan 17 kolom.

Lalu dilakukan slicing terhadap kolom-kolom yang diperlukan untuk memudahkan proses analisis, lalu dapat disimpan ke dalam file csv atau bentuk data yang lain.



```
1 df_save = df[['instant', 'dteday', 'season', 'temp', 'hum', 'windspeed']]
2 df_save.head()
3
4 df_save.to_csv('c:/Users/syay2/OneDrive/ドキュメント/NF/SEM 7/praktikum/Praktikum_1/data/latihan01.csv', index=False)
```

*Gambar 7. Kode untuk menslicing data dan save ke dalam bentuk csv*

	instant	dteday	season	temp	hum	windspeed
0	1	2011-01-01	1	0.24	0.81	0.0
1	2	2011-01-01	1	0.22	0.80	0.0
2	3	2011-01-01	1	0.22	0.80	0.0
3	4	2011-01-01	1	0.24	0.75	0.0
4	5	2011-01-01	1	0.24	0.75	0.0

*Gambar 8. Hasil dari load slicing data*

### 3. Link Repo github

Repository Syayid Muhammad Akbar : [Repo Machine-Learning-NF](#)