

**TUGAS KONSEP APLIKASI DATA MAINING**  
**NAÏVE BAYES**



**Aminurachma Aisyah Nilatika**  
**17.52.001**

**KEMENTRIAN RISET DAN TEKNOLOGI PENDIDIKAN TINGGI**  
**SEKOLAH TINGGI MANAJEMEN INFORMATIKA KOMPUTER**  
**PRADNYA PARAMITA**  
**MALANG**  
**2020**

1. Yang pertama adalah melakukan import library, yang pertama adalah pandas nanti di sebut sebagai pd, matplotlib.pyplot nanti di sebut sebagai plt.

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

2. Selanjutnya adalah membaca file excel menggunakan library pandas. Dimana nama file excel yang di gunakan adalah ('D:/Aminurachma Aisyah Nilatika/STIMATA/Semester 6/Data Mining/UAS/datamining/Uas/dataset\_soal No. 1.xls',header) lalu, header sudah berhasil maka akan muncul, karena merupakan file csv saya menggunakan delimiter=';') sebagai pemisah.

```
In [2]: data = pd.read_csv('D:/Aminurachma Aisyah Nilatika/STIMATA/Semester 6/Data Mining/UAS/datamining/Uas/dataset_soal No.1.csv', deli
```

```
In [3]: data.head()
```

Dengan output :

```
Out[3]:
```

	Age	Income	Student	Credit_rating	Class (buy_computer)
0	<= 30	High	No	Fair	No
1	<= 30	High	No	Excellent	No
2	31..40	High	No	Fair	Yes
3	> 40	Medium	No	Fair	Yes
4	> 40	Low	Yes	Fair	Yes

3. Hitung jumlah dataset yang kita miliki dengan menggunakan syntax berikut ini :

```
In [6]: data.shape
```

Dengan output :

```
(51, 5)
```

4. Selajutnya, hitung jumlah yang membeli dan tidak membeli

```
In [5]: data['Class (buy_computer)'].value_counts()
```

```
Out[5]: Yes    29
No    22
Name: Class (buy_computer), dtype: int64
```

5. Langkah selanjutnya adalah mencari knowledge dari dataset

```

In [34]: print("PNoStudentNo= ",PNoStudentNo)
print("PYesStudentNo= ",PYesStudentNo)

print("PNoStudentYes= ",PNoStudentYes)
print("PYesStudentYes= ",PYesStudentYes)

print("PNoStudent= ",PNoStudent)
print("PYesStudent= ",PYesStudent)

print("PExcellentNo= ",PExcellentNo)
print("PFairNo= ",PFairNo)

print("PExcellentYes= ",PExcellentYes)
print("PFairYes= ",PFairYes)

print("PExcellent= ",PExcellent)
print("PFair= ",PFair)

print("P31_40No= ",P31_40No)
print("PKur30No= ",PKur30No)
print("PLeb40No= ",PLeb40No)

print("P31_40Yes= ",P31_40Yes)
print("PKur30Yes= ",PKur30Yes)
print("PLeb40Yes= ",PLeb40Yes)

print("P31_40= ",P31_40)
print("PKur30= ",PKur30)
print("PLeb40= ",PLeb40)

print("PExcellentHigh= ",PExcellentHigh)
print("PExcellentLow= ",PExcellentLow)
print("PExcellentMedium= ",PExcellentMedium)

print("PFairHigh= ",PFairHigh)
print("PFairLow= ",PFairLow)
print("PFairMedium= ",PFairMedium)

print("PNoStudentExcellent= ",PNoStudentExcellent)
print("PNoStudentFair= ",PNoStudentFair)

print("PYesStudentExcellent= ",PYesStudentExcellent)
print("PYesStudentFair= ",PYesStudentFair)

```

Dengan output :

```

PNoStudentNo= 0.45454545454545453
PYesStudentNo= 0.5454545454545454
PNoStudentYes= 0.4827586206896552
PYesStudentYes= 0.5172413793103449
PNoStudent= 0.47058823529411764
PYesStudent= 0.5294117647058824
PExcellentNo= 0.36363636363636365
PFairNo= 0.6363636363636364
PExcellentYes= 0.41379310344827586
PFairYes= 0.5862068965517241
PExcellent= 0.39215686274509803
PFair= 0.6078431372549019
P31_40No= 0.2857142857142857
PKur30No= 0.5
PLeb40No= 0.47058823529411764
P31_40Yes= 0.7142857142857143
PKur30Yes= 0.5
PLeb40Yes= 0.5294117647058824
P31_40= 0.27450980392156865
PKur30= 0.39215686274509803
PLeb40= 0.3333333333333333
PExcellentHigh= 0.25
PExcellentLow= 0.4
PExcellentMedium= 0.35
PFairHigh= 0.1935483870967742
PFairLow= 0.41935483870967744
PFairMedium= 0.3870967741935484
PNoStudentExcellent= 0.3333333333333333
PNoStudentFair= 0.6666666666666666
PYesStudentExcellent= 0.4444444444444444
PYesStudentFair= 0.5555555555555556

```

```
In [38]: data.describe()
```

```
Out[38]:
```

	Age	Income	Student	Credit_rating	Class (buy_computer)
count	51	51	51	51	51
unique	3	3	2	2	2
top	<= 30	Low	Yes	Fair	Yes
freq	20	21	27	31	29

- 3 -