Praktikum 4

Naive Bayes Classifier

In [ ]:
```
import numpy as np
import pandas as pd
```

Data

In [ ]:
```
df = pd.read_excel("DataPlayTennis.xlsx")
```

In [ ]:
```
df.head()
```

In [ ]:
```
df['Outlook'].unique()
```

In [ ]:
```
df['Humidity'].mean()
```

In [ ]:
```
df['Humidity'].median()
```

In [ ]:
```
X = df.drop('Class',axis=1)
y = df['Class']
```

In [ ]:
```
outlook_enc=pd.get_dummies(X['Outlook'],prefix='Outlook')
```

In [ ]:
```
windy_enc=pd.get_dummies(X['Windy'],prefix='Windy')
```

In [ ]:
```
encoded_outlook_windy=pd.concat([encoded_outlook,windy_enc],axis=1)
```

In [ ]:
```
encoded_outlook_windy
```

In [ ]:
```
encoded_outlook_windy.drop(['Windy','Outlook_overcast','Windy_False'],axis=1,inplace=True)
encoded_outlook_windy
```

In [ ]:
```
X=encoded_outlook_windy
```

In [ ]:
```
# memisahkan dataset ke dalam set training dan set testing
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.4, random_state=1)
```

In [ ]:
```
# melatih model dengan set training
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(X_train, y_train)
```

In [ ]:
```
# membuat prediksi hasil set testing
predictions = gnb.predict(X_test)
```

In [ ]:
```
# membandingkan hasil prediksi dan label sesungguhnya
from sklearn.metrics import classification_report,confusion_matrix
```

In [ ]:
```
print(classification_report(y_test,predictions))
```

In [ ]:
```
print(confusion_matrix(y_test,predictions))
```