Nama : Suci Maria

Kelas : TIF-A1

Npm : 41155050210005

TUGAS PERTEMUAN 4

1. **Lakukan praktik dari https://youtu.be/Sj1ybuDDf9I?si=hCajHe1zasTQ9HGY , buat screenshot dengan nama kalian pada coding, kumpulkan dalam bentuk pdf, dari kegiatan ini:**

   1.1. Pengenalan Bayes Theorem | Teori Bayes | Conditional Probability

   Bayes'theorem menawarkan suatu formula untuk menghitung nilai probability dari suatu event dengan memanfaatkan pengetahuan sebelumnya dari kondisi terkait atau sering kali dikenal dengan istilah conditional probability

   $$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

   $$P(y|X) = \frac{P(X|y) \times P(y)}{P(X)}$$

   $$Posterior = \frac{Likelihood \times Prior}{Evidence}$$

   1.2. Pengenalan Naive Bayes Classification

   Studi Kasus 1

Asep	Joko
+ siomay: 0.1 + bakso: 0.8 +lumpia: 0.1	+ siomay: 0.1 + bakso: 0.8 +lumpia: 0.1

   Misi: Lakukan prediksi siapa pelanggan yang melakukan pemesanan dengan diketahui pesanannya adalah **lumpia** dan **bakso**

   1.3. Pengenalan Prior Probability

   Prior Pribability: p(y)

   - *P(Asep)* = 0.5
   - *P(Joko)* = 0.5

1.4. Pengenalan Likelihood

Likelihood: P(X | y)

- Asep :

$$P(\text{lumpia,bakso} \mid \text{Asep}) = (0.1 \times 0.8)$$
$$= 0.08$$

- Joko :

$$P(\text{umpia,bakso} \mid \text{Joko}) = (0.3 \times 0.2)$$
$$= (0.06)$$

1.5. Pengenalan Evidence | Normalizer

Evidence atau Normalizer: P(X)

$$\text{Evidence} = \sum (\text{Likelihood} \times \text{Prior})$$
$$P(\text{lumpia,bakso}) = (0.08 \times 0.5) + (0.06 \times 0.5)$$
$$= 0.07$$

1.6. Pengenalan Posterior Probability

Posterior Probability: P(y | X)

- Formula:

$$\text{Posterior} = \frac{Likelihood\ x\ Prior}{Evidence}$$

- Asep :

$$P(\text{Asep} \mid \text{lumpia,bakso}) = \frac{0.08\ x\ 0.5}{0.07}$$
$$= 0.57$$

- Joko :

$$P(\text{Joko} \mid \text{lumpia,bakso}) = \frac{0.06\ x\ 0.5}{0.07}$$
$$= 0.43$$

1.7. Studi kasus dan implementasi Naive Bayes

StudiKasus 2

Asep	Joko
+ siomay: 0.1	+ siomay: 0.5
+ bakso: 0.8	+ bakso: 0.2
+lumpia: 0.1	+lumpia: 0.3

Misi : Lakukan prediksi siapa pelanggan yang melakukan pemesanan dengan diketahui pesanannya adalah **siomay** dan **bakso.**

Posterior Probability: P(y | X) (Kasus 2)

- Pesanan : siomay, bakso
- Evidence : P(X)

  P(siomay,bakso) = (0.1 x 0.8 x 0.5) + (0.5 x 0.2 x 0.5)

  = 0.09
- Asep :

$$P(\text{Asep} \mid \text{siomay, bakso}) = \frac{(0.1\ x\ 0.8)\ x\ 0.5}{0.09}$$

$$= 0.444$$

- Joko :

$$P(\text{Joko} \mid \text{siomay, bakso}) = \frac{(0.5\ x\ 0.2) x\ 0.5}{0.09}$$

$$= 0.555$$

Dataset : Breast Cancer Wisconsin (Diagnostic)

- **Load Dataset**

```
[1]: from sklearn.datasets import load_breast_cancer
     print(load_breast_cancer().DESCR)
```

```
.. _breast_cancer_dataset:

Breast cancer wisconsin (diagnostic) dataset
--------------------------------------------

**Data Set Characteristics:**

:Number of Instances: 569

:Number of Attributes: 30 numeric, predictive attributes and the class

:Attribute Information:
    - radius (mean of distances from center to points on the perimeter)
    - texture (standard deviation of gray-scale values)
    - perimeter
    - area
    - smoothness (local variation in radius lengths)
    - compactness (perimeter^2 / area - 1.0)
```

```
[2]: print("Nama: Suci Maria")
```

```
Nama: Suci Maria
```

[4]:
```
load_breast_cancer?
```

```
Signature: load_breast_cancer(*, return_X_y=False, as_frame=False)
Docstring:
Load and return the breast cancer wisconsin dataset (classification).

The breast cancer dataset is a classic and very easy binary classification
dataset.

=================   ==============
Classes                          2
Samples per class    212(M),357(B)
Samples total                  569
Dimensionality                  30
Features            real, positive
=================   ==============

The copy of UCI ML Breast Cancer Wisconsin (Diagnostic) dataset is
downloaded from:
https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic

Read more in the :ref:`User Guide <breast_cancer_dataset>`.
```

[5]:
```
print("Nama: Suci Maria")
```

```
Nama: Suci Maria
```

[6]:
```
X, y = load_breast_cancer(return_X_y=True)
X.shape
```

[6]:
```
(569, 30)
```

[7]:
```
print("Nama: Suci Maria")
```

```
Nama: Suci Maria
```

➢ Training & Testing set

[8]:
```
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=0)
print(f'X_train shape {X_train.shape}')
print(f'X_test shape {X_test.shape}')
```

```
X_train shape (455, 30)
X_test shape (114, 30)
```

[9]:
```
print("Nama: Suci Maria")
```

```
Nama: Suci Maria
```

➢ Naive Bayes dengan Scikit Learn

[10]:
```
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import accuracy_score

model = GaussianNB()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
accuracy_score(y_test, y_pred)
```

[10]:
```
0.9298245614035088
```

[11]:
```
print("Nama: Suci Maria")
```

```
Nama: Suci Maria
```

[12]:
```
model.score(X_test, y_test)
```

[12]:
```
0.9298245614035088
```

[13]:
```
print("Nama: Suci Maria")
```

```
Nama: Suci Maria
```