

# MACHINE LEARNING

Nama : Hanif Ridal Warits

NPM : 41155050210060

Kelas : Informatika A2 – 2021

## Tugas Pertemuan 4

### 1.0. 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

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

### Pengenalan Naive Bayes Classification

#### Studi Kasus 1

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 pesannya adalah **lumpia** dan **bakso**.

### 1.3. Pengenalan Prior Probability

**Prior Probability:**  $P(y)$

- Referensi: [https://en.wikipedia.org/wiki/Prior\\_probability](https://en.wikipedia.org/wiki/Prior_probability)
- $P(Asep) = 0.5$
- $P(Joko) = 0.5$

### 1.4. Pengenalan Likelihood

**Likelihood:**  $P(X|y)$

- Referensi: [https://en.wikipedia.org/wiki/Likelihood\\_function](https://en.wikipedia.org/wiki/Likelihood_function)

- Asep:

$$\begin{aligned} P(lumpia, bakso|Asep) &= (0.1 \times 0.8) \\ &= 0.08 \end{aligned}$$

- Joko:

$$\begin{aligned} P(lumpia, bakso|Joko) &= (0.3 \times 0.2) \\ &= 0.06 \end{aligned}$$

### 1.5. Pengenalan Evidence | Normalizer

**Evidence atau Normalizer:**  $P(X)$

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

## 1.6. Pengenalan Posterior Probability

**Posterior Probability:**  $P(y|X)$

- Referensi: [https://en.wikipedia.org/wiki/Posterior\\_probability](https://en.wikipedia.org/wiki/Posterior_probability)

- Formula:

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

- Asep:

$$\begin{aligned} P(Asep|lumpia, bakso) &= \frac{0.08 \times 0.5}{0.07} \\ &= 0.57 \end{aligned}$$

- Joko:

$$\begin{aligned} P(Joko|lumpia, bakso) &= \frac{0.06 \times 0.5}{0.07} \\ &= 0.43 \end{aligned}$$

## 1.7. Studi kasus dan implementasi Naive Bayes

## ▼ DATASET : Breast Cancer Wisconsin (Diagnostics)

```
[1] print('Hanif Ridal Warits - 41155050210060')
```

```
Hanif Ridal Warits - 41155050210060
```

Load Dataset

```
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)
  - concavity (severity of concave portions of the contour)
  - concave points (number of concave portions of the contour)
  - symmetry
  - fractal dimension ("coastline approximation" - 1)

The mean, standard error, and "worst" or largest (mean of the three
worst/largest values) of these features were computed for each image,
resulting in 30 features. For instance, field 0 is Mean Radius, field
10 is Radius SE, field 20 is Worst Radius.

- class:
  - MIBC-Malignant
  - MIBC-Benign

:Summary Statistics:

=====
:              Min      Max
-----
radius (mean):    6.981  28.11
texture (mean):   9.71   39.28
perimeter (mean): 43.79  188.5
area (mean):      143.5 2501.0
smoothness (mean): 0.053  0.163
compactness (mean): 0.019  0.345
concavity (mean):  0.0    0.427
concave points (mean): 0.0    0.201
symmetry (mean):   0.106  0.304
fractal dimension (mean): 0.05  0.097
radius (standard error): 0.112  2.873
texture (standard error): 0.36  4.885
```

```
[3]
perimeter (standard error): 0.757  21.98
area (standard error):      6.802 542.2
smoothness (standard error): 0.002  0.031
compactness (standard error): 0.002  0.135
concavity (standard error):  0.0    0.396
concave points (standard error): 0.0    0.053
symmetry (standard error):    0.008  0.079
fractal dimension (standard error): 0.001  0.03
radius (worst):              7.93   36.04
texture (worst):             12.02  49.54
perimeter (worst):           50.41 251.2
area (worst):                185.2 4254.0
smoothness (worst):          0.071  0.223
compactness (worst):          0.027  1.058
concavity (worst):           0.0    1.252
concave points (worst):       0.0    0.291
symmetry (worst):            0.155  0.664
fractal dimension (worst):    0.055  0.208
=====
```

:Missing Attribute Values: None

:Class Distribution: 212 - Malignant, 357 - Benign

:Creator: Dr. William H. Wolberg, W. Nick Street, Olvi L. Mangasarian

:Donor: Nick Street

:Date: November, 1995

This is a copy of UCI ML Breast Cancer Wisconsin (Diagnostic) datasets.  
<https://goo.gl/U2Uwz2>

