

Tugas Pertemuan 4

Nama : Rizky Muhamad Wicaksana

NPM : 41155050210010

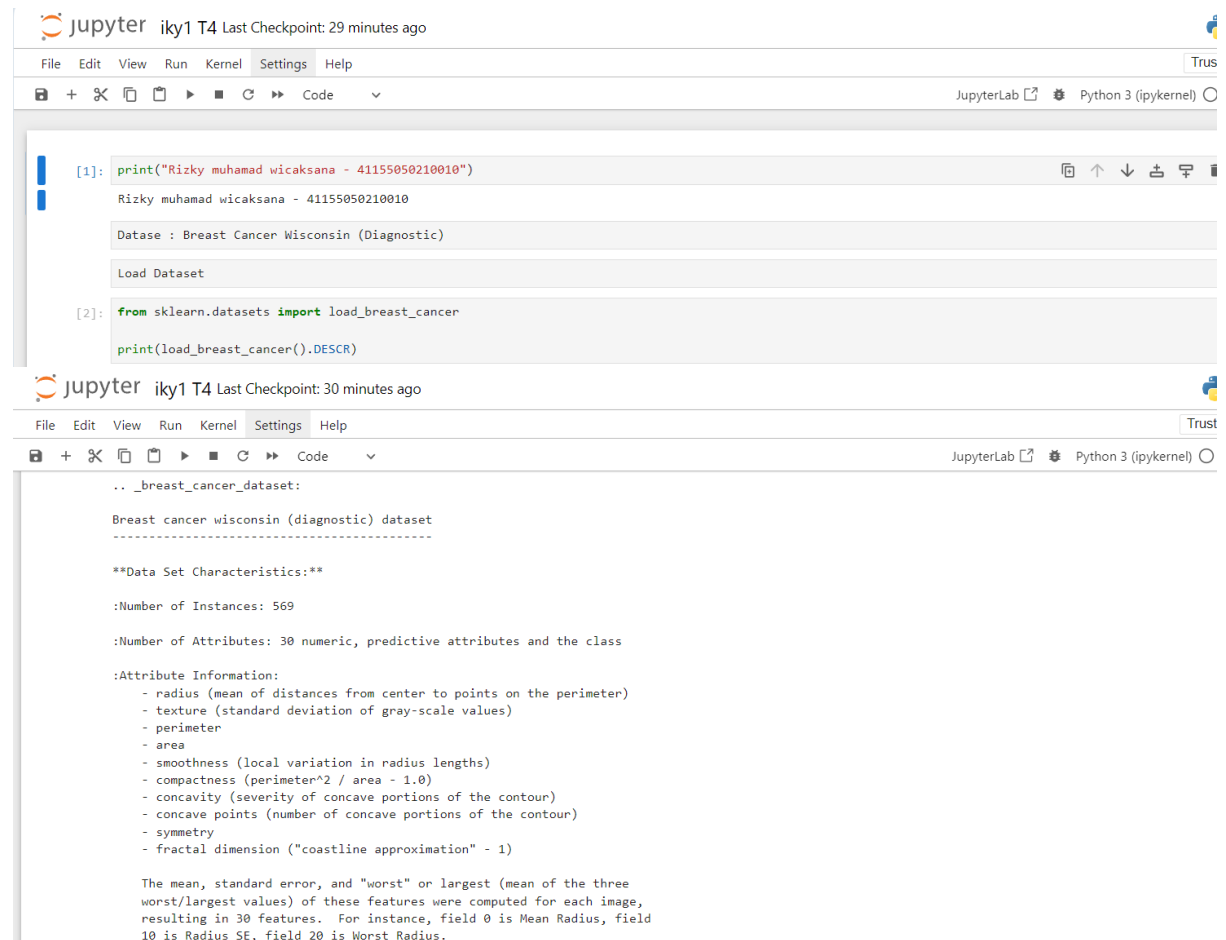
Informatika A1

Machine Learning

- 1.1. Pengenalan Bayes Naive Bayes Classification Probability
- 1.2. Pengenalan Naive Bayes Classification
- 1.3. Pengenalan Prior Probability
- 1.4. Pengenalan Likelihood
- 1.5. Pengenalan Evidence | Normalizer
- 1.6. Pengenalan Posterior Probability
- 1.7. Studi kasus dan implementasi Naive Bayes

Dataset : Breast cancer Wisconsin (Diagnostic)

Load Dataset



```
[1]: print("Rizky muhamad wicaksana - 41155050210010")

Rizky muhamad wicaksana - 41155050210010

Dataset : Breast Cancer Wisconsin (Diagnostic)

Load Dataset

[2]: 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:
  - WDBC-Malignant
  - WDBC-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
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):      0.156  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

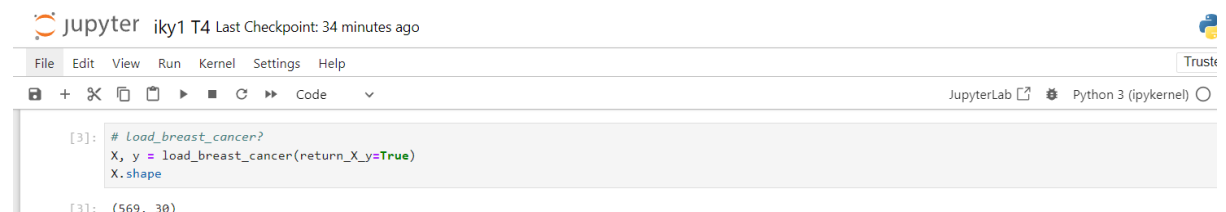
:Donor: Nick Street

:Date: November, 1995

```

Deskripsi data set

Bila ingin mengakses dokumen tesion untuk function “load_breast_cancer?” maka tinggal panggil tidak perlu dikomen karna jupyter notebook memiliki keleluasaan atau kebebasan untuk mngakses dokumen tesion dari suatu function tertentu



The screenshot shows a Jupyter Notebook window with the title "jupyter iky1 T4 Last Checkpoint: 34 minutes ago". The interface includes a menu bar (File, Edit, View, Run, Kernel, Settings, Help) and a toolbar with icons for saving, running, and other actions. The code cell [3] contains the following Python code:

```

[3]: # load_breast_cancer?
X, y = load_breast_cancer(return_X_y=True)
X.shape

```

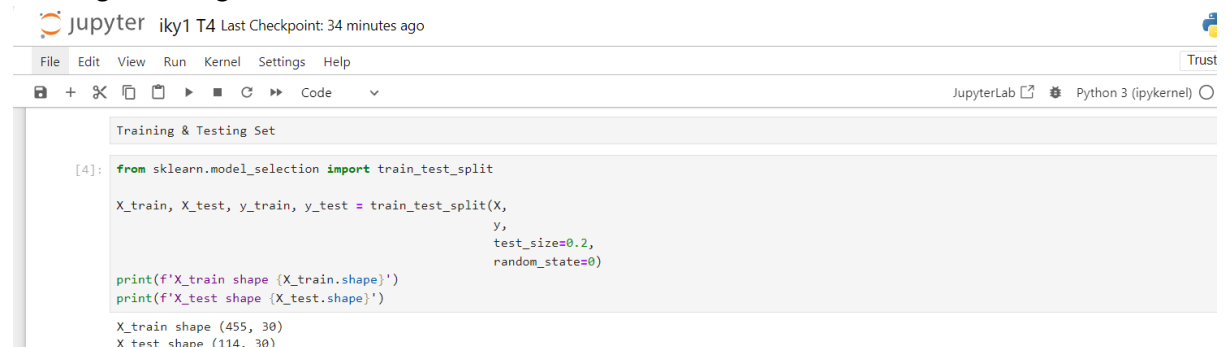
The output of the code cell is displayed below the code:

```

[3]: (569, 30)

```

Training & Testing Set



The screenshot shows a Jupyter Notebook window with the title "jupyter iky1 T4 Last Checkpoint: 34 minutes ago". The interface includes a menu bar (File, Edit, View, Run, Kernel, Settings, Help) and a toolbar with icons for saving, running, and other actions. The code cell [4] contains the following Python code:

```

[4]: 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)

```

Naïve Bayes dengan Scikitn Learn

Jupyter iky1 T4 Last Checkpoint: 35 minutes ago 

File Edit View Run Kernel Settings Help Trust

 JupyterLab 

Naive Bayes dengan Scikitn Learn

```
[5]: 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)
```

[5]: 0.9298245614035088

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

[6]: 0.9298245614035088