XGBoost

# library link

**install :**

https://xgboost.readthedocs.io/en/stable/install.html

**github :**

https://github.com/dmlc/xgboost/tree/master/python-package

# basic description



XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable. It implements machine learning algorithms under the Gradient Boosting framework.XGBoost provides a parallel tree boosting that solve many data science problems in a fast and accurate way.

# version

* xgboost == 1.5.2 (pip install xgboost)

# dataset

* Using Sklearn.datasets.load\_breast\_cancer.
* Sources : https://scikit-learn.org/stable/datasets/toy\_dataset.html#breast-cancer-dataset

# code description

* The dataset is converted into DMatrix object format so that the dataset imported from sklearn can be learned in the XGBoost model, and then learned in the XGBoost model. Afterwards, the test dataset is predicted through the learned model, and the results of the first 10 portions are output.
* In addition, accuracy, precision, record, F1,AUC between predicted and actual values are calculated and output separately, and then the importance of each Features inside the model is shown through a graph.

# validation

* Inside the code, the dataset is divided into learning datasets and verification datasets to verify this.

(test\_size = 0.2, random\_state = 30)

* Additionally, the fitness between the actual value and the predicted value is evaluated using Sklearn's acuity\_score,precision\_score,recall\_score,f1\_score,roc\_auc\_score function.