ada\_boost Open Source code

# library link

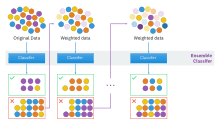
**install :**

[**https://scikit-learn.org/stable/install.html**](https://scikit-learn.org/stable/install.html)

**github :**

**https://github.com/scikit-learn/scikit-learn/blob/7e1e6d09b/sklearn/ensemble/\_weight\_boosting.py#L313**

# basic description



AdaBoost is a statistical classification meta-algorithm formulated by yoav Freund and Robert Schapire. It can be used in conjunction with many other types of learning algorithms to improve performance.

# version

* Python ( >=3.7 )
* NumPy( >= 1.14.6 )
* Scipy( >= 1.1.0 )
* joblib( >= 0.11 )
* threadpoolctl( >= 2.0.0 )
* For example code :
* pandas (pip install pandas)

# dataset

train.scv: The Otto Group is one of the world’s biggest e-commerce companies, with subsidiaries in more than 20 countries, including Crate & Barrel (USA), Otto.de (Germany) and 3 Suisses (France). We are selling millions of products worldwide every day, with several thousand products being added to our product line.

A consistent analysis of the performance of our products is crucial. However, due to our diverse global infrastructure, many identical products get classified differently. Therefore, the quality of our product analysis depends heavily on the ability to accurately cluster similar products. The better the classification, the more insights we can generate about our product range.

For this competition, we have provided a dataset with 93 features for more than 200,000 products. The objective is to build a predictive model which is able to distinguish between our main product categories. The winning models will be open sourced.

Sources : https://www.kaggle.com/c/otto-group-product-classification-challenge

# code description

* Code for separating the same dataset into a learning verification dataset, learning them in the models of decision tree, adaboost, and random forest methods, verifying them through the verification dataset, and comparing the accuracy of each model.

# validation

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

(test\_size = 0.3, random\_state = 30)

* Additionally, the fitness between the actual value and the predicted value is evaluated using Sklearn's metrics.accuracy\_score function.