
Pandas

Pandas

- Open-source, BSD-licensed Python library
- Provides high-performance, easy-to-use data structures and data analysis tool
- Used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc.

DataFrame

A Data frame is a two-dimensional data structure.

Features of DataFrame

- Potentially columns are of different types
- Size - Mutable
- Labeled axes (rows and columns)
- Can Perform Arithmetic operations on rows and columns

Scikit-learn

Scikit-learn

- Free software machine learning library
- It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means

Iris dataset

- This data set consists of 3 different types of irises' (Setosa, Versicolour, and Virginica) petal and sepal length, stored in a 150x4 numpy.ndarray
- The rows being the samples and the columns being: Sepal Length, Sepal Width, Petal Length and Petal Width.