Scikit Learn

Scikit-learn strives to have a uniform interface across all methods, and we'll see examples of these below.

Given a scikit-learn *estimator* object named model, the following methods are available...

Scikit Learn

- Available in **all Estimators**
    - model.fit() : fit training data.
    - For supervised learning applications, this accepts two arguments: the data X and the labels y (e.g. model.fit(X, y)).
    - For unsupervised learning applications, this accepts only a single argument, the data X (e.g. model.fit(X)).

Scikit Learn

Available in **supervised estimators**

- model.predict() : given a trained model, predict the label of a new set of data. This method accepts one argument, the new data X_new (e.g. model.predict(X_new)), and returns the learned label for each object in the array.

Scikit Learn

Available in **supervised estimators**

- model.predict_proba() : For classification problems, some estimators also provide this method, which returns the probability that a new observation has each categorical label. In this case, the label with the highest probability is returned by model.predict().

Scikit Learn

Available in **supervised estimators**

- model.score() : for classification or regression problems, most estimators implement a score method. Scores are between 0 and 1, with a larger score indicating a better fit.

Scikit Learn

Available in **unsupervised estimators**

- model.predict() : predict labels in clustering algorithms.

Scikit Learn

Available in **unsupervised estimators**

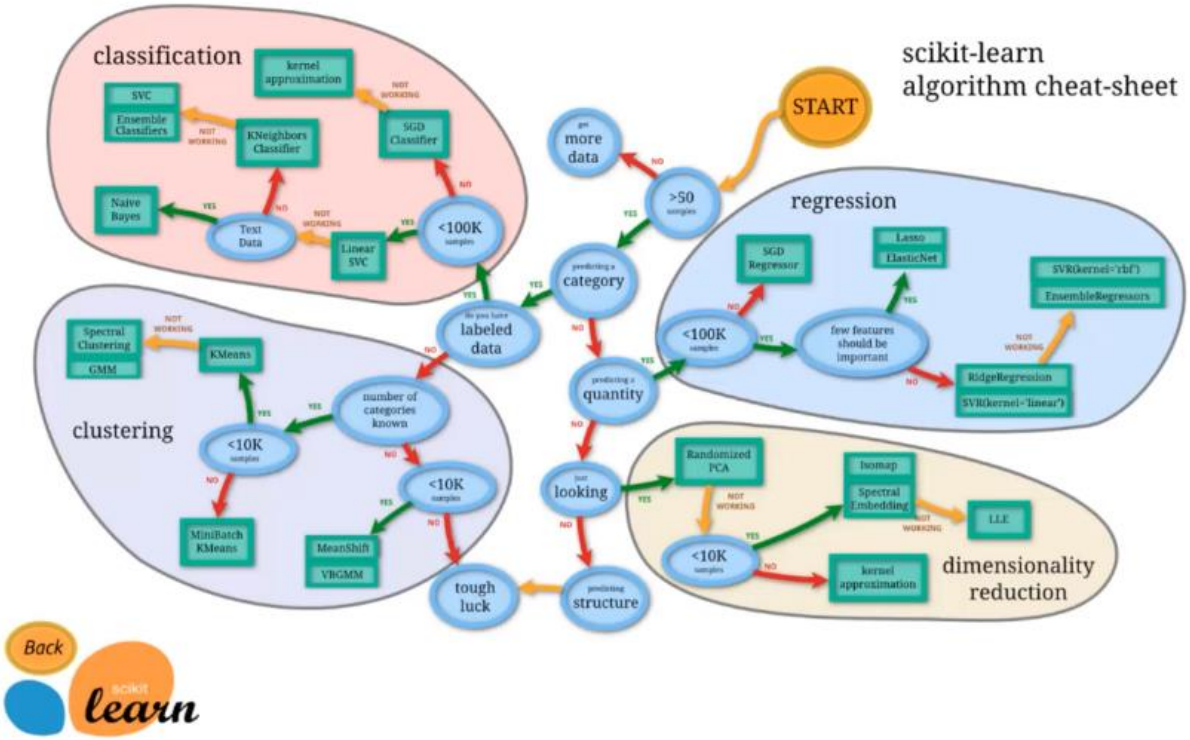
- model.transform() : given an unsupervised model, transform new data into the new basis. This also accepts one argument X_new, and returns the new representation of the data based on the unsupervised model.

Scikit Learn

Available in **unsupervised estimators**

- model.fit_transform() : some estimators implement this method, which more efficiently performs a fit and a transform on the same input data.

Choosing an Algorithm



ML with Python