



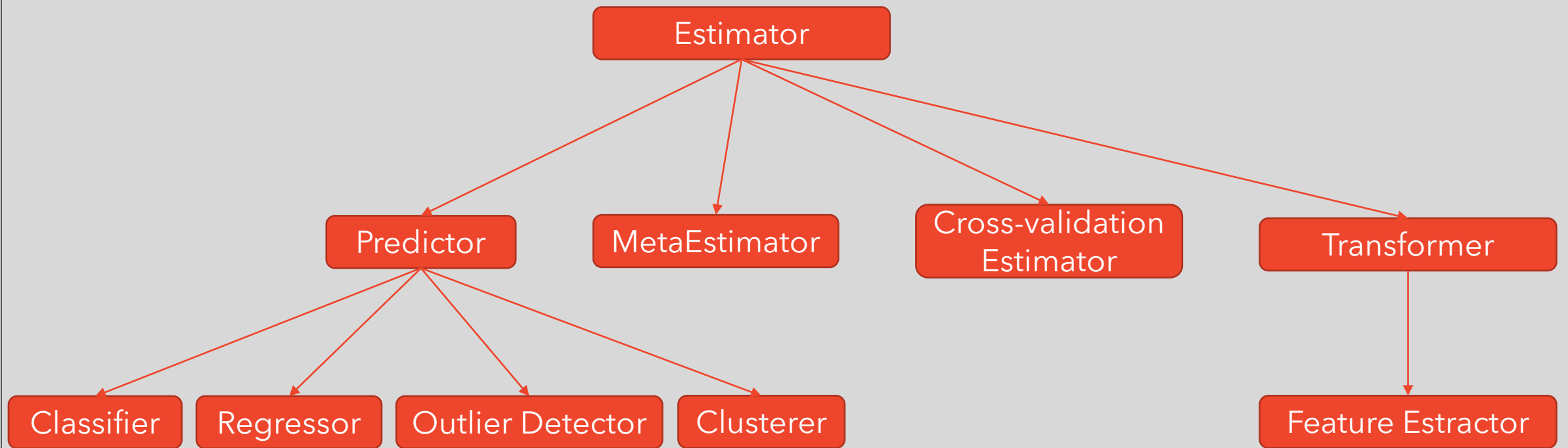
SCIKIT-LEARN

Aspetti Architeturali

SCIKIT-LEARN

- Scikit-learn is an open source **machine learning Python library** that supports **supervised** and **unsupervised learning**. It also provides various tools for **model fitting**, **data preprocessing**, **model selection**, **model evaluation**, and many other utilities.
- Focus: API della libreria
 - Overview degli oggetti principali e loro gerarchia
 - Uso di API come linee guida per lo sviluppo di elementi custom -> utile per lo sviluppo di Pipeline



Gerarchia Oggetti



Estimator

- Object che stima il modello utilizzando un insieme di dati di addestramento e può inferire alcune proprietà su dati «unseen»/nuovi
- Implementano i metodi
 - `fit(data, target)` oppure
 - `fit(data)`

Stima di un modello dipende da:

- Dati passati al metodo `fit()` o `partial_fit()` 
- Parametri del modello = iperparametri. I metodi `get_params()` e `set_params()` permettono l'interazione con l'insieme degli iperparametri, insieme al costruttore dell'oggetto
 - NB. Il costruttore non accetta MAI i dati di training/addestramento 
- Eventuale random state `numpy.random.RandomState`

Predictor

- Estimator che implementa anche il metodo
- `predict()`
- `fit_predict()`

Classifier

- Classificatore supervisionato:
 - Binario - in questa immagine c'è almeno una persona?
 - Multiclasse - se c'è una persona identifica la sua professione
 - Multioutput - se c'è una persona identifica la sua professione e identifica in quale quadrante si trova la persona
 - Multilabel = multioutput con output binari - identifica se c'è almeno una persona e/o almeno un animale nell'immagine.
- Implementano metodi: `fit`, `predict`, `score`
- Possono implementare: `decision_function()`, `predict_proba()`, `predict_log_proba()`
- Memorizzano attributo `classes_`

Regressor

- Predictor supervisionato che restituisce output continui (reali R)

Outlier Detector

- Identifica punti anomali (outlier) rispetto al core
- `fit_predict()` se transduttivo
- `predict()` se induttivo

Clusterer

- Predictor non supervisionato
- `fit_predict()` se transduttivo
- `Predict()` se induttivo
- Memorizza l'attributo `labels_()`

Transformer

- Oggetti che implementano il metodo `transform()` e/o `fit_transform()`
- Applicano una trasformazione ai dati

Feature Extractor

- Trasforma un sample non array-like in un sample array-like. Esempio stringhe, sequenze trasformate in vettori
- Metodi:
 - `get_feature_names()`
 - `get_feature_names_out()`

Moduli SKL

`sklearn.calibration`: Probability Calibration

`sklearn.cluster`: Clustering ☒

`sklearn.compose`: Composite Estimators

`sklearn.covariance`: Covariance Estimators

`sklearn.cross_decomposition`: Cross decomposition

`sklearn.datasets`: Datasets ☒

`sklearn.decomposition`: Matrix Decomposition

`sklearn.discriminant_analysis`: Discriminant Analysis

☒ `sklearn.dummy`: Dummy estimators

`sklearn.ensemble`: Ensemble Methods ☒

`sklearn.exceptions`: Exceptions and warnings

`sklearn.experimental`: Experimental ☒

`sklearn.feature_extraction`: Feature Extraction ☒

`sklearn.feature_selection`: Feature Selection

`sklearn.gaussian_process`: Gaussian Processes

`sklearn.impute`: Impute ☒

`sklearn.inspection`: Inspection

`sklearn.isotonic`: Isotonic regression

`sklearn.kernel_approximation`: Kernel Approximation

`sklearn.kernel_ridge`: Kernel Ridge Regression

`sklearn.linear_model`: Linear Models ☒

`sklearn.manifold`: Manifold Learning

`sklearn.metrics`: Metrics ☒

`sklearn.mixture`: Gaussian Mixture Models

`sklearn.model_selection`: Model Selection ☒

`sklearn.multiclass`: Multiclass classification ☒

`sklearn.multioutput`: Multioutput regression and classification

`sklearn.naive_bayes`: Naive Bayes

`sklearn.neighbors`: Nearest Neighbors

`sklearn.neural_network`: Neural network models

`sklearn.pipeline`: Pipeline ☒

`sklearn.preprocessing`: Preprocessing and Normalization ☒

`sklearn.random_projection`: Random projection

`sklearn.semi_supervised`: Semi-Supervised Learning

`sklearn.svm`: Support Vector Machines ☒

`sklearn.tree`: Decision Trees ☒

`sklearn.utils`: Utilities ☒