

COMPLÉMENT · OUTIL

scikit-learn · la boîte à outils

Une interface cohérente pour apprendre, évaluer et comparer.

M1 · 12 h · cours + TP · scikit-learn

Le même protocole pour plusieurs modèles

DONNÉES

X : lignes × variables
y : cible

ESTIMATEUR

fit(X_train, y_train)
predict(X_test)

ÉVALUATION

métrique + validation
comparaison honnête

Classifier, régresseur, transformeur : API commune, comportements distincts

Les objets essentiels du TP

ESTIMATEUR

LogisticRegression(), DecisionTreeClassifier(), Ridge()

STANDARDISER moyenne et écart-type

StandardScaler().fit_transform(X_train)

MÉTHODES

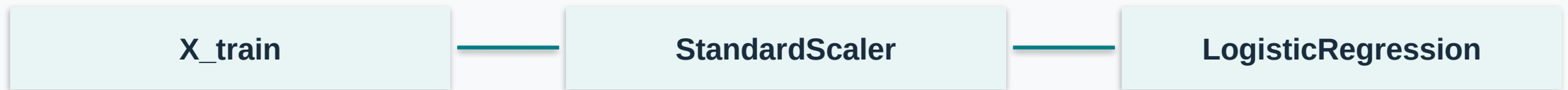
fit apprend ; predict produit des sorties ; score() a une valeur par défaut.

PARAMÈTRES

get_params() les lit ; set_params() les modifie.

Attention : score() vaut généralement accuracy en classification, R^2 en régression.

Pipeline : l'antidote à la fuite de prétraitement



```
modele = make_pipeline(StandardScaler(), LogisticRegression())
modele.fit(X_train, y_train)
```

En validation croisée, chaque pli réapprend son scaler uniquement sur sa partie entraînement.

Comparer et régler sans ouvrir le test

DÉVELOPPEMENT

```
cross_validate(model, X_dev, y_dev, cv=5,  
scoring='accuracy')
```

RECHERCHE

```
GridSearchCV(model, {'logisticregression__C': [0.1, 1,  
10]}, cv=5)
```

TEST FINAL

Choisir le modèle sur le développement ; ouvrir le test une seule fois après ce choix.

Quatre réflexes avant de croire le score

1 · SÉPARER

Réserver le test avant toute sélection ou normalisation apprenante.

2 · ENCAPSULER

Placer les transformations apprenantes dans le Pipeline.

3 · RÉPÉTER

Utiliser des plis adaptés et les mêmes pour tous les modèles.

4 · DÉCRIRE

Rapporter métrique, dispersion, erreurs et coût.

Périmètre et documentation

scikit-learn inclut un MLP simple ; CNN, Transformers et grands modèles génératifs relèvent d'autres frameworks.

https://scikit-learn.org/stable/getting_started.html

https://scikit-learn.org/stable/modules/cross_validation.html

https://scikit-learn.org/stable/common_pitfalls.html

<https://scikit-learn.org/stable/faq.html>