

Tutoriel `scikit-learn` : les classifieurs SVM

Romain Tavenard

Voici l'en-tête à utiliser pour la suite de ce travail :

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats
from sklearn.datasets.samples_generator import make_blobs
```

Pour cette partie du tutoriel, nous allons générer des données synthétiques simples. `sklearn` fournit des générateurs automatiques de données synthétiques, nous allons en utiliser un pour plus de facilité :

```
X, y = make_blobs(n_samples=50, centers=2, random_state=0,
                  cluster_std=0.60)
```

Bien sûr, il n'est pas question de manipuler des données sans les visualiser :

```
plt.scatter(X[:, 0], X[:, 1], c=y)
```

1 TODO list

- fit a linear SVM
- plot support vectors
- compute coordinates of the separating hyperplane
- plot separating hyperplane
- same thing with a dataset for which classification errors occur and growing C values???
- fit a non-linear SVM for a more difficult problem