

Notice du démonstrateur

Ce démonstrateur comporte 5 pages. Pour chacune d'entre elles, vous trouverez les explications ci-dessous.

Comment utiliser ce démonstrateur

Page d'accueil. Cliquez sur le bouton de téléchargement pour arriver sur cette notice.

Qu'est ce qu'un SVM?

Si vous n'avez pas de notion en terme de machine learning, nous vous expliquons ici brièvement la méthode de machine à vecteur de support (SVM). Cette courte introduction est inspiré du cours de Christophe Hurlin, que vous pouvez retrouver en intégralité en cliquant le lien ci-dessous:

<https://sites.google.com/view/christophe-hurlin/teaching-resources/support-vector-machine>.

Présentation de nos données

Vous trouverez ici toutes les informations sur les données que nous avons utilisées pour construire ce démonstrateur. Vous pouvez retrouver la base de données sur le site de kaggle, en cliquant sur le lien ci dessous:

<https://www.kaggle.com/mlg-ulb/creditcardfraud>.

Interaction avec le SVM

Ici, on montre les matrices de confusions sur l'échantillon d'apprentissage et sur l'échantillon test (les observations bien classées sont en bleu, les autres en oranges), la sensibilité, la spécificité et le taux d'erreur qui leurs sont associées, ainsi que la courbe ROC obtenu par la méthode du SVM.

Cette page est interactive. Ainsi, vous pouvez choisir:

- **La fonction noyau** qui permettra de rendre l'échantillon linéairement séparable (linéaire, polynomial, radial ou sigmoïde).
- **Le degré du polynôme** que l'on peut fixer à 3, 4 ou 5 que si nous avons choisi le kernel polynomial.
- **Le coût de pénalisation**, que l'on peut fixer à 1,3,5 ou 10. Pour rappel, plus ce dernier est important, plus on fait d'erreur, plus il est grand, plus on risque de faire de l'overfitting. Si on choisit la fonction noyau polynomial, on doit également fixé le nombre de degrés du polynome (3,4 ou 5)

Comparaison du SVM avec les autres méthodes de machines learning

Cet onglet se compose de 3 parties distinctes:

Recherche du meilleur SVM

Page d'introduction de l'onglet. On explique brièvement la comparaison des méthodes.

Régression logistique

On compare les résultats du SVM à ceux de la régression logistique.

Arbre de régression

On compare les résultats du SVM à ceux de l'arbre de classification.