

## Devoir : Atelier de Machine Learning

Vous allez faire le même exercice de l'atelier à savoir étudier l'utilisation d'une méthode de classification supervisée pour prédire une variable Y ou Y vaut 1 si spam et 0 sinon.

Vous devez :

- Utiliser la base de données *spam* de la librairie **kernlab** (les données utilisées dans l'atelier).
- appliquer une autre méthode de classification supervisée pour prédire la variable Y (différente de la méthode utilisée dans l'atelier)
- Donner l'erreur de prédiction estimée par validation croisée.
- Comparez le résultat obtenu avec celui de la méthode utilisée dans l'atelier.

Rappel :

La base de données spam se trouve dans la librairie kernlab

**>library(kernlab)**

Le paquet **caret** (abréviation de Classification And REgression Trainig) est un ensemble de fonctions qui visent à rationaliser le processus de création des modèles prédictifs.

**> library(caret)**

Ce paquet contient des outils pour :

- Le fractionnement de données, à l'aide de la fonction `createDataPartition()`
- Fixer quelques paramètres du processus d'apprentissage à l'aide de la fonction `trainControl()` .
- L'apprentissage d'un modèle à l'aide la fonction `train()`. Cette fonction optimise aussi les paramètres du modèle.
- La construction d'un prédicteur à l'aide de la fonction `predict()`
- fonctionnalité sélection de variables
- estimation d'importance d'une variable
- ainsi que d'autres fonctionnalités.