

# Analyse de données multivariées – Machine Learning

## Arbres de décision

### Objectifs

- Apprendre les bases des arbres de décision, issus du « Machine Learning »
- Savoir analyser vos propres données multivariées à l'issue du cours.
- Utiliser seul le logiciel d'analyse de données étudié

### Programme

#### Arbres de décision

##### CART

- Principe théorique des arbres de décision
- Optimisation
- Limitations des modèles CART
- Application sur jeu de données et logiciel

##### Random Forests

- Principe des « ensemble methods »
- Optimisation
- Application sur jeu de données et logiciel

##### Boosting

- Discrimination
- Régression
- Application sur jeu de données et logiciel

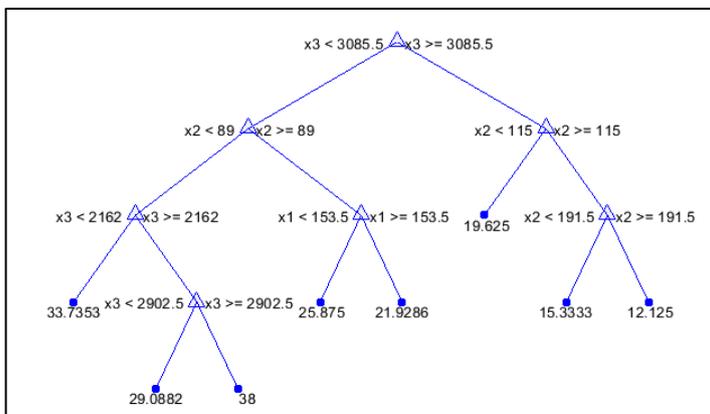
### Les plus

- Les principes des méthodes sont introduits par une approche géométrique
- L'accent est mis sur l'utilisation pratique des méthodes et l'interprétation des résultats
- Exercices d'application sur un jeu de données pour chaque méthode
- Utilisation d'un logiciel de chimiométrie: SPM® (Salford Systems - Minitab), MATLAB® (Mathworks)



### Public concerné

- Pré-requis : Connaissances en analyse de données multivariées (ACP, PLS)
- Profil : Techniciens, ingénieurs ou chercheurs
- Domaines : R&D, développement de produits, contrôle et optimisation de procédés, contrôle qualité...
- Type de données : Instrumentales, données physico-chimiques, paramètres procédés,...
- Secteurs d'activité : Agriculture/Agro-alimentaire, Pétrochimie, Pharmaceutique, Biotechnologies, Chimie, Environnement...



### Contact

Ondalys

4, rue Georges Besse

34 830 Clapiers

Tél : 04.67.67.97.87 / fax : 04.67.67.97.88

✉ : [contact@ondalys.fr](mailto:contact@ondalys.fr)

### Infos pratiques

- Durée du stage : 1 jour
- Nombre de participants : 8 personnes maximum
- Une convention de stage sera établie entre l'entreprise et Ondalys
- N° organisme de formation : 91-34-05332-34 auprès du Préfet de la région Languedoc-Roussillon