

STATISTICA

Transformez vos Données en Actions

Pratique des techniques de data mining



Objectifs :

Comprendre le fonctionnement de *STATISTICA* Data Miner et de ses modules, utiliser les principales techniques analytiques. Savoir créer un projet pour une étude et déployer les modèles retenus.

Public concerné :

Testeurs et futurs utilisateurs et de *STATISTICA* Data Miner

Pré-requis :

Bases en analyse des données (Régression Multiple, ACP, Classification).
Connaissance élémentaire de Windows (gestion des fichiers, répertoires).

Durée :

2 jours, soit 12 heures

Organisation :

6 stagiaires maximum
1 personne par ordinateur
projecteur multimédia

Un questionnaire permet au formateur de mieux cibler les attentes des stagiaires pour les exercices d'application. Merci de retourner ce questionnaire avec votre bulletin d'inscription.

Les données

- Les différents types de données
- Faire une requête simple
- Requête avec l'outil IDBD

Les nœuds analytiques

- Insertion d'un nœud
- Modification des paramètres d'un nœud
- Personnalisation d'un nœud

Préparation et transformation des données

- Fusion de fichiers
- Transformations de fichier (ajout de variables, transformations...)
- Préparation (traitement des données manquantes, échantillonnage...)

Exploration des données

- Statistiques descriptives et Drill-Down
- Sélection et filtrage des meilleurs prédicteurs

Analyse : techniques de modélisation numérique

- Arbres de régression, Réseaux de neurones, MARS, Modèles additifs généralisés,
- Apprentissage machine (SVM)...

Analyse : techniques de classification

- Classification supervisée : K-moyennes, Arbres de classification, Réseaux de neurones,
- Apprentissage machine, MARS...
- Classification non supervisée : ACP, CAH, Cartes de Kohonen

Utilisation des projets

- Les Projets pré-paramétrés
- Les Projets personnalisés

Qualité d'un modèle et choix de modèles

- Comparaison de modèles
- Utilisation de plusieurs modèles

Résultats : présentation des sorties

- Différentes sorties disponibles
- Personnalisations des sorties

Déploiement : exploitation des modèles

- Utilisation du modèle sur de nouvelles données par un expert (statisticien)
- Utilisation du modèle par un non-statisticien

Exercices d'application