

Data Mining: summary



SII-298 2 Days (14 Hours)

Description

Data Mining is a technique that allows you to extract knowledge from raw data. This seminar offers a theoretical and practical overview of Data Mining, covering its applications, its advantages for the company, as well as the most commonly used methods and tools. You will learn the theoretical principles and will have the opportunity to participate in practical experiments to better understand this technique and its use in a professional context.

Who is this training for ?

For whom

Infocentre managers, marketing managers, statisticians, IT managers, project managers and decision-making experts. Users and business database managers.

Prerequisites

Basic knowledge of decision analysis. Basic knowledge of statistics.

Training objectives

- Understand the principles and purpose of Data Mining (DM) Identify the main techniques of DM and their use case Implement scoring and geomarketing methods on a simple case Discover the predictive methods and descriptive methods of DM Know the main stages of a Data Mining project

Training program

Le Système d'Information Décisionnel (SID)

- The challenges of SID: needs, areas of application.
- Typical architecture of a SID, state of the art.
- Development of decision-making information.
- Design of a SID: steps, optimization, data organization, dictionaries.

Comprendre le Data Mining (DM)

- Definition and purpose of Data Mining (DM).
- What link between DM and statistics, what dependence between DM and IT? Difference between DM and OLAP ? Business expectations, DM responses.

Les techniques du Data Mining

- The different families of DM.
- Predictive methods and descriptive methods.
- Factor analysis, typological.
- Classification.
- Decision trees, neural networks.
- Classification of DM techniques.

La méthode descriptive du Clustering

- Definition and methodology.
- The criteria for structuring the data to be classified.
- Evaluation and validation of the classes obtained.
- The different subsections -Clustering families.
- Example Presentation of Clustering applications.

Exemples d'application du DM

- Scoring: definition, purpose, methodology.
- Geomarketing: definition, purpose, methodology.
- Example Implementation of the scoring method .
- Practical use case for geomarketing.

Les données de l'entreprise

- Reminder of IS data issues.
- Data quality and data administration.
- Collection and exploration process.
- Creating aggregates and new data.
- Transforming data.

Méthodologie de projet Data Mining

- Definition of the business problem to be solved and the objectives to be achieved.
- Inventory, describe and classify the data.
- Design and populate the Data Mining database.
- Explore, segment analyzed entities.
- Establish and apply analysis models.
- Iterate, deploy to users.
- Maintain the model and associated software.

Panorama des outils

- Main tools on the market: SAS, R, IBM SPSS.
- Focus on the SAS tool and the Powercenter ETL.
- What selection criteria for this type of tools?