

> Data Mining - PASW Modeler - in Healthcare

This two-day course introduces you to the major steps of the data mining process in healthcare environment. The course goal is for you to be able to begin planning or evaluate your firm's current plan for data mining. Topics include: setting goals; framing research questions; organizing, cleaning and manipulating data; selecting analysis methods; presenting and automating. The course also reviews the basic operations and environment of the PASW Modeler (formerly known as Clementine®) data mining software. Within the context of data mining, you will be introduced to such machine learning techniques as rule induction, neural networks, cluster analysis and association rules, and will apply them to business data.

Following an overview of data mining and an introduction to essential terminology and technology, you will then proceed logically through the following topics:

INTRODUCTION

- ◆ Definition of Data Mining
- ◆ Data Mining and Statistics CRISP-DM
- ◆ Data Mining Examples in Healthcare

CHECKING AND TRANSFORMING DATA

- ◆ Data Cleaning
- ◆ Data Quality
- ◆ Creating New Variables Transformation

BUSINESS UNDERSTANDING (SETTING GOALS)

- ◆ Business Context
- ◆ Consider Constraints
- ◆ Track and Evaluate Results

MODELLING TECHNIQUES- OVERVIEW

- ◆ Neural networks
- ◆ Market Basket or Association Analysis
- ◆ Cluster Analysis
- ◆ Memory-Based Reasoning Decision Trees
- ◆ Regression
- ◆ Which techniques, when?

DATA UNDERSTANDING AND PREPARATION

- ◆ Reading in Data Files
- ◆ Merging and Appending datasets
- ◆ Data Exploration
- ◆ Missing Values
- ◆ Data Selection

EVALUATION AND DEPLOYMENT

- ◆ Choosing among Methods
- ◆ Deployment of Models
- ◆ Deployment of Results

Target Audience

The target audience for this course is the broadly defined Healthcare market. Included are those in government, corporate and educational settings.

Prerequisites

Attendees should be familiar with basic research methods. Training in advanced statistical methods is not necessary, but those with such training will gain a deeper understanding of the models.

Objectives

By the end of the course, you will have learned:

- ◆ How to identify potential data mining applications by using PASW Modeler in healthcare;
- ◆ How to carry out data exploration techniques using visualization and statistics, as well as to model your healthcare data;
- ◆ How to interpret results effectively using both numerical and graphical output enabling appropriate conclusions to be drawn.

Cost

HK\$4,000.00 per person

Duration

2 days (7 hours x 2 days)