

Big Data Analytics in Smart Manufacturing

Course objective

In the Industry 4.0 era, machines, devices, sensors, and people are connected via IoT and results in enormous exchange of data. Such revolution necessitates the systematic analytics on data to transform them into information that enables “informed” decisions.

Therefore, organisations must be able to adapt to big data technologies to meet the expectations of smart manufacturing. However, big data analytics is a relatively new phenomenon and its potential applications on manufacturing activities are wide-reaching and diverse.

In this 2-days course, we will walk you through multiple big data analytics use cases in the manufacturing industry, via hands-on exercises incorporating state-of-the-art analytics techniques.

We understand that just having vast quantities of data at hand doesn't mean one can extract the needed insights. Therefore, the key deliverables of this course – the analytical methodologies to turn big data into useful information, is, in turn, the key to sustainable innovation in a smart manufacturing environment.

What will you learn?

Address production challenges:

- Yield ramp-up is a challenge or
- Waste reduction or
- Throughput optimization

Address analytical challenges:

- Equipment and process complexity
- Process dynamics
- Data quality

After the training, students will have the ability to:

- Perform all common data preparations
- Build sophisticated predictive models
- Evaluate model quality with respect to different criteria
- Deploy analytical predictive models

Pre-requisite

Basic knowledge of computer programs and mathematics.

Training methodology

Hands-on exercise, lecture, group discussion, and case study.

Course Outline:

1. Overview

- Importance of Data Analytics
- Business scenario – Semiconductor Manufacturing environment
- CRISP-DM – Articulate a business problem to data science problem

2. Basic usage

- Introduction to User interface
- Creating and handling RapidMiner repositories
- Starting a new RapidMiner project
- Operators and processes
- Loading data
- Storing data, processes, and results

3. Data Cleansing and Preparation

- Normalization and standardization
- Basic transformations of value types
- Handling missing values
- Sampling
- Filtering examples and attributes
- Handling attribute roles
- Joining Data from different Sources

4. EDA: Exploratory Data Analysis

- Data Types
- Data Hierarchy
- Quick Summary Statistics
- Visualizing Data
- Charting

5. Building better processes

- Organizing
- Renaming
- Relative Path
- Flow Control
- Subprocesses
- Breakpoints

6. Predictive models

- K – Nearest Neighbour
- Naive Bayes
- Linear Regression
- Decision Trees
- Optimization
- Importance of attributes – Feature Selections

7. Model evaluation

- Applying models
- Splitting data
- Evaluation methods
- Performance criteria

8. Model Application

- Scoring on new set of data

9. Sharing and collaboration

- RapidMiner Server
- Automation