RapidMiner Server: Deployment & Management

Course Overview

RapidMiner Server: Deployment & Management is a one-day expert seminar focusing on the RapidMiner Server product from the perspective of a user and an administrator. RapidMiner Server is a collaboration platform for data science teams with comprehensive user management, including sharing via individual access control and version control. It integrates with enterprise data sources such as relational databases, and also includes job scheduling to automate recurring analytics and MIS tasks. Finally, it provides access to data science predictions and results in external systems via modern web services.

After successfully completing this course, participants will have a solid understanding of how RapidMiner Studio and Server work together in an organization to support data science initiatives. The course will cover the installation and configuration of RapidMiner Server in the Amazon Web Services (AWS) environment and connecting it to RapidMiner Studio and external relational databases. It will also show how analysts can use RapidMiner Server to deploy predictive models, collaborate with team members, manage access rights, automate processes, and deploy results via web services. The class exercises and labs are hands-on, so students will internalize the topics covered, which will provide a jumpstart to the real world application of these techniques.

Prerequisites & Target Audience

This class is aimed at Analysts and Data Scientists as well as IT Administrators. It assumes a basic knowledge of computer programming principles and higher mathematics (through calculus). It assumes basic knowledge of RapidMiner Studio and basic data science. Some familiarity with SQL is helpful. ***Note that students must also have access to their own RapidMiner Server, or an AWS account in which they can deploy a new RapidMiner Server instance (which will be covered in this course).

Course Outline

- RapidMiner Server Installation inside AWS (via AMI)
- Collaboration
    - Sharing Data Repositories & Building a Process Library
    - User Access Control and Process Version Control
    - Reading from and Writing to External Relational Databases
    - Parameterizing Processes for Broader Use
- Automation
    - Deploying Tested Models in a Production Environment
    - Scheduling Processes
    - Using Triggers
    - Turning RapidMiner Processes into Web Services