DATA SCIENTIST WITH PYTHON

Professional machine learning with big data

The goal of the Data Scientist-Training is the application-oriented implementation of data models with Python for the prediction of different business scenarios. Unsupervised and supervised machine learning algorithms are implemented and iteratively optimized. Visualization methods and Data Storytelling are also taught in order to be able to take on the role of a Data Scientist after the training.

MODULES



MACHINE LEARNING BASICS

- · Refresh how to perform data analyses with pandas and matplotlib.
- · Learn how to create your first machine learning models.
- Performance metrics and assumptions of models of supervised and unsupervised learning with *sklearn*.

2

SUPERVISED LEARNING

- Deepened application of algorithms of supervised learning, such as decision trees and random forests, support vector machines.
- Understanding the Python module *TensorFlow* in the context of neural networks.
- Understanding challenges in the context of machine learning projects and developing models.

3

ADVANCED TOPICS IN DATA SCIENCE

- · Understanding the foundations of data storytelling.
- Model interpretability: model specific and model agnostic methods.
- Implement best practices in the informative design of visualizations with *bokeh*.
- Get to know big data procedures with *pyspark*.
- Conduct independent data science projects with an industrial data set.



ORGANIZATIONAL

Format: Online-Training **Duration:** 96 h in 4 months

(4-6 h/week)



PARTICIPANTS

The training is aimed at employees who want to specialize in the field of data science or expand their existing knowledge with data and predictions. The Data Science Training is important for departments that want to automate more data—driven decisions and predictions.



BENEFITS

- Enable data-based (automated) decisions.
- Get to know relevant data science projects with the help of knowledge from the subject domain.
- Make data-based predictions in the domain.