


MANNHEIM DATA SCIENCE CERTIFICATE: BIG DATA & MACHINE LEARNING

This certificate will prepare you to help organizations to incorporate new data sources and emerging technologies in their decision and reporting processes. With the skills you acquire with this certificate you help organize the data pipeline, from data collection or ingestion to final data product.

KEY FACTS

 **STRUCTURE**
Asynchronous learning through video lectures combined with weekly 1-hour live online meetings


 **CERTIFICATE DURATION**
2 courses
4-8 weeks per course

 **TIME COMMITMENT**
Part-Time, full day

 **LANGUAGE**
English

 **APPLICATION REQUIREMENTS**

- An academic degree
- Fluency in English

 **CREDITS**
Certificate of Completion

 **PARTICIPANTS**
Limited to 20

CERTIFICATE DESCRIPTION

The amount of data generated as a by-product in society is growing fast including data from satellites, sensors, transactions, social media, and smartphones, just to name a few. Such data are often referred to as Big Data and used for prediction and classification tasks, both of which can be tackled with machine learning techniques. In this certificate, you will explore, from a social science perspective, how Big Data concepts, processes, and methods can be used within the context of social research. We will illustrate key concepts using specific survey research examples including tailored survey designs and nonresponse adjustments and evaluation. Further, we discuss advanced Machine Learning concepts such as cross validation, class imbalance, boosting, and stacking as well as key approaches for facilitating model tuning and performing feature selection. Code examples will be provided using the statistical programming language R.

KEY BENEFITS

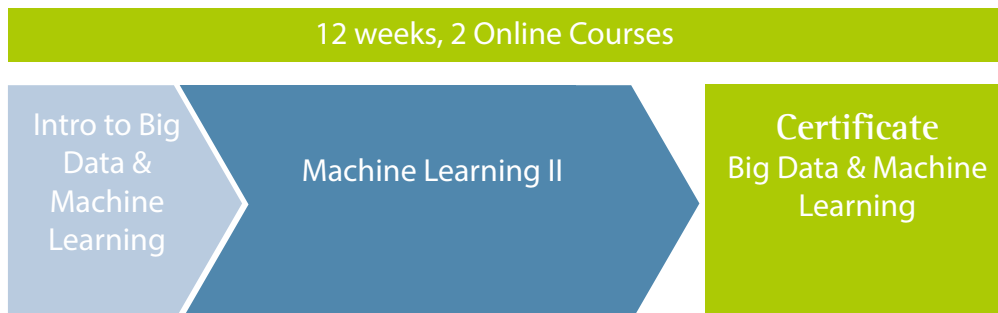
- **Excellence:** Theoretically based and practice-oriented learning from a faculty from world's top ranked universities, statistical agencies and businesses
- **Flexibility:** Join courses online from anywhere in the world, at your own pace.
- **Online learning:** Asynchronous learning experience (pre-recorded lectures, readings) and synchronous learning experience (virtual classroom, weekly live discussions led by the instructor)
- **Participant profile:** Participants benefit from a diverse group of international peers coming from various industries and with different occupational profile




BY COMPLETION OF THE CERTIFICATE YOU WILL...

- understand key Big Data terminology and concepts
- know about common data generating processes
- understand some primary issues with linking Big Data with survey data
- understand issues of coverage and measurement errors within the Big Data context
- be able to contrast inference versus prediction
- understand general concepts from machine learning including signal detection and information extraction
- learn about potential pitfalls for inference from Big Data
- understand key analytic techniques (e.g. classification trees, random forests, conditional forests) to process Big Data using R with example code provided
- have a profound understanding of advanced (ensemble) prediction methods
- have built up a comprehensive ML toolkit to tackle various learning problems
- know how to (critically) evaluate and interpret results from "black-box" models

COURSES



SAMPLE SCHEDULE

	Introduction to Big Data & Machine Learning	Machine Learning II
 Mandatory weekly online meeting	Mondays, 1:00 PM ET/ 7:00 PM CET, June 6 – 27, 2022	Fridays, 1:00 PM EDT/ 7:00 PM CEST, September 29 – November 11, 2022
Final Exam	July 4, 2022	November 25, 2022

CASE STUDIES & EXAMS

There will be a mixture of weekly online quizzes and assignments. Additionally, the courses each conclude with a final exam.

To see all courses in the upcoming term click [here](#).