Foundations of Machine Learning and Unstructured Data in Business Contexts

Dr. Jochen Hartmann
(jochen.hartmann@uni-hamburg.de)

Course Abstract

The rise of machine learning is disrupting how firms do business. In this introductory course you will gain foundational knowledge of a large spectrum of machine learning methods. Moreover, you will gain insights on diverse real-world applications as well as ethical considerations of machine learning.

In the second part of the course, we will focus on extracting insights from unstructured data such as texts, images, and videos using state-of-the-art deep learning methods to inform profound business decisions. The course is designed to give you an end-to-end perspective on applied machine learning, learn from leading companies, and to discuss the common pitfalls and opportunities that arise when mining (unstructured) data in business contexts.

Course Objectives

The main course objectives are:
1. Understand the foundations of machine learning
2. Get exposure to numerous practical examples from leading companies
3. Learn about responsible machine learning, biases, and challenges
4. Build up a toolbox of diverse supervised and unsupervised machine learning methods
5. Apply different approaches to solve real-world problems

Course Composition and Teaching Methods

The course blends theoretical foundations, multiple applied case studies, hands-on coding, real-world examples, and break-out sessions. Overall, you will experience an interactive learning environment.

Evaluation and Grading

- Group presentation (40%)
- Individual case study (60%)

Group presentation:

- In the kick-off, you are assigned to groups of 3-5 students. The assignment questions will be announced in class on Sun, November 14, 2021.

Individual case study:

- Following November 14, 2021, you will individually work on a quantitative case study. The case study materials and assignment questions will be announced in class.
- To solve the individual case study, you will be required to run analyses in Python, visualize, and interpret your results. Please analyze the data using a Jupyter Notebook.
- **Please use the standardized Jupyter Notebook submission template distributed in class.**
- Your analyses and answers will be graded with respect to the following criteria:
  - **Scope & Performance:** *Application and evaluation of diverse, adequate machine learning methods (incl. hyperparameter tuning)*
  - **Insight:** *Logical interpretation of results with compelling managerial recommendations*
  - **Comprehensibility:** *Meaningfully annotated Python code and supporting visualizations*

Literature

Mandatory Readings
- Improving Worker Safety in the Era of Machine Learning (A) (HBS Case 9-618-019).
- Improving Worker Safety in the Era of Machine Learning (B) (HBS Case 9-618-064).
- Chapter 1 of Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems.* 2nd Edition. O'Reilly Media.

The materials will be provided in the kick-off session on Wed, June 30, 2021.

Acknowledgements

The seminar is supported by the Exist program, led by Prof. Dr. Michel Clement.

Important Dates

Kick-off: Wed, June 30, 2021, from 18:00-20:00

Foundations of Machine Learning 1: Sat, November 13, from 09:00-18:00

Foundations of Machine Learning 2: Sun, November 14, from 09:00-18:00

Group Presentations: Fri, December 3, from 16:00-19:00

Discussion of Individual Case Study: Fri, January 21, from 16:00-19:00

All meetings take place via Zoom.

Dr. Jochen Hartmann
Postdoctoral Researcher at the University of Hamburg

Jochen Hartmann is a postdoctoral researcher at the University of Hamburg and a frequent external lecturer at Mannheim Business School. He holds a B.Sc. and M.Sc. degree in Management from WHU Vallendar. Before his doctoral studies, he worked as a management consultant at McKinsey & Company, specialized on marketing and machine learning. Jochen is a regular visiting scholar at Columbia Business School.

Jochen's substantive research interests include digital marketing, human-machine interactions, and customer insights. He is fascinated by the opportunities arising from the automatic analysis of unstructured data such as images, texts, and videos. He has published a paper on video mining in *Journal of Business Research.* His work on automated text classification has appeared in *International Journal of Research in Marketing.*