

Empirical Evaluation of Management Practices I

Predictions and Machine Learning

Prof. Dr. Dirk Sliwka Jesper Armouti-Hansen

02/11/20

- 1 Introduction to Machine Learning
- 2 Regression
- 3 Classification
- 4 Model Selection and Assessment
- 5 Decision Trees and Random Forests

1. Introduction to Machine Learning

Definition

“the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data”. - Oxford Languages

- Sessions (Lectures with integrated exercises):
 - Mondays: 14:00 - 15:30 and 16:00 - 17:30
 - Tuesdays: 10:00 - 11:30
- All sessions take place on Zoom (See ILIAS for instructions).
- Exam:
 - Date: TBA!
 - Format: Remote.
- Course website: github.com/armouthansen/EEMP2020

Organizational Issues (Coding)

For all our coding exercises as well as the exam (most likely), we will be using notebooks in Google Colab (colab.research.google.com/).

If you have a functioning google account, you should be able to:

- Create new notebooks
- Save notebooks to your Google Drive and open them again
- Open notebooks from Github repositories
- Share notebooks with your fellow students

Thus, make sure that you have a Google account!

Working in groups

When doing exercises in the lecture, you will work in groups of three in break-out rooms on Zoom:

- The groups are randomly assigned by us and will stay the same throughout the course
- Within groups you should share your google account email so that you are able to collaborate
- We (the lectures) will go from room to room to check your progress and answer your questions

In order to create the groups and have them ready for the next lecture, please send your Zoom email to me (j.hansen@uni-koeln.de) today!

Questions

- If the question is not urgent, please ask during the lecture
- If urgent, write me an email (j.hansen@uni-koeln.de)

Today's agenda

The goal of today's session is to give you a short overview of working with python for data analysis. This includes:

- An overview of basic data structures and commands in python
- Providing a toolkit for data analysis in python

Note: Today's session is far from exhaustive. More packages, concepts and commands will be introduced later in the course.