Assignment on Explorative Data Analysis

Machine Learning

Prof. Dr. Marc Pouly & Dr. Tim vor der Brück

Lucerne University of Applied Sciences and Arts

Dear students,

this introductory course to machine learning has not started yet. However, an industry partner just provided us with a real data set and is now quite eager to hear whether machine learning applied to this data can open up new opportunities or improve their current business. We therefore ask you to conduct an explorative data analysis and to elaborate on how machine learning could bring our partner to a next level. We award points for non-trivial insights that emerge from your data analysis as well as for reasonable advises to our industry partner. Try your best even when you feel like being thrown in at the deep end. Hand in your solution as a Jupyter notebook by e-mail to Marc Pouly no later than February 27, 2018. In case you need assistance or maybe just a little help on how to start off, feel free to get in touch with Tim vor der Brück.

Exercise 1: Install Party (0 Points)

1. Install Jupyter Notebook from here

2. Organize yourselves in teams of at most 4 members

Hint: For team exercises, Jupyter Notebooks can easily be shared over Git.

Exercise 2: Schifffahrtsgesellschaft Vierwaldstättersee (2 Points)

We are given a data set containing the number of passengers per day along with a lot more information. Conduct an explorative analysis, uncover interesting and non-trivial insights and speculate on how your insights would boost our partner's business.

Use Python and Jupyter Notebooks to hand in code along with the conclusions you draw from the analysis and the findings that we shall transfer to our industry partner. We award 1 point for an explorative data analysis that uncovers non-trivial and interesting insights and 1 point for useful and detailed advises on how our partner could turn these insights into business.