

Basisschulung Data Science mit Python

Dorian Prill, Olaf Sassnick & Max Schirl
Donnerstag, 3. Dezember 2020

Vorabinformation: All jene, die aktiv am Workshop teilnehmen wollen und Docker nicht installiert haben, folgen bitte diesem [Link](#)!

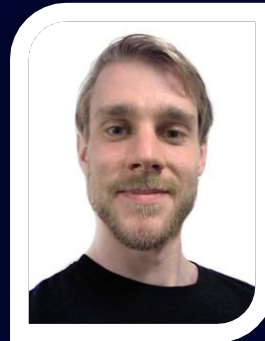


Wer sind wir?

Dorian Prill

Junior Researcher, Studiengang ITS

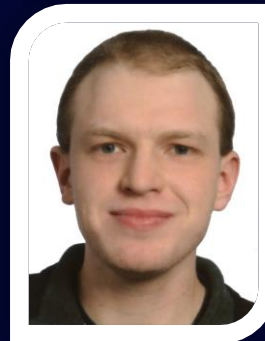
dorian.prill@fh-salzburg.ac.at



Olaf Saßnick

Researcher, Studiengang ITS

olaf.sassnick@fh-salzburg.ac.at



Wer sind wir?

Max Schirl

Junior Researcher, Studiengang ITS

maximilian.schirl@fh-salzburg.ac.at





Wer sind Sie?

- **Vorstellung**
- **Kurzinfo zu Vorkenntnissen**
- **Wenn möglich bitte mit Kamera! 😊**

Was steht auf dem Programm?

1. Das Setup
2. Die Daten
3. Die Herangehensweise
4. Timeline

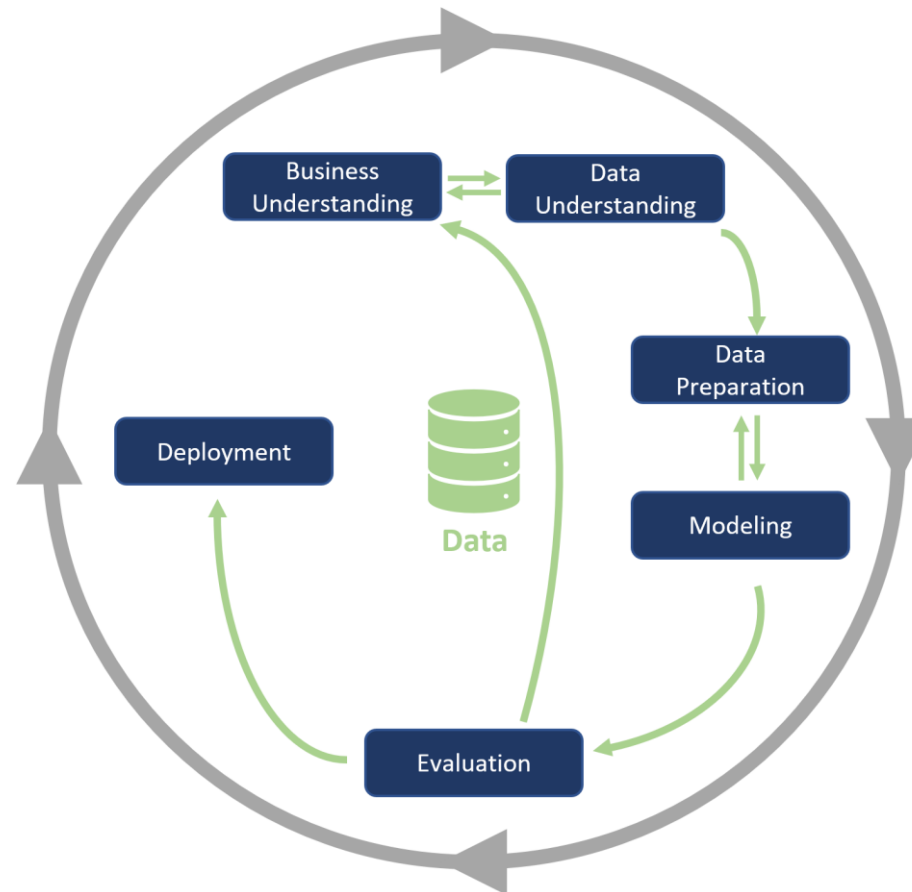
1. Das Setup

- Docker Container
 - mit
 - [python 3.8.3](#)
 - [jupyter-notebook 6.0.3](#)
 - [pandas 1.0.5](#)
 - [seaborn 0.10.1](#)
 - [scikit-learn 0.23.1](#)
 - [tensorflow 2.2.0](#)
 - [Keras 2.4.3](#)
 - bereitgestellt für Sie auf [github.com](#)

2. Die Daten

- „Student Performance Data Set“
 - öffentlich zugänglich ([UCI Machine Learning Repository](#))
 - beschreibt
 - schulische Leistung (Noten per Trimester sowie Abschlussnote)
 - demographische Daten (aus Befragung bzw. Erhebung)
 - soziale und schulbezogene Daten (aus Befragung bzw. Erhebung)
 - von SchülerInnen zweier portugiesischer Schulen ([MS & GP](#))
 - aufgezeichnet im Schuljahr 2005-2006
 - in den Fächern Mathematik & Portugiesisch

3. Die Herangehensweise



CRISP-DM Modell (Shearer 2000, p. 14)

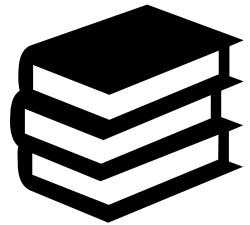
4. Timeline

Timeline	
~2h	Business Understanding Data Understanding Data Preparation
1h	Mittagspause
1-2h	Modeling & Evaluation
Ende	



Fragen?

Gerne jederzeit, der Workshop lebt von unserer Interaktion!



Let's Go!

<https://github.com/mschirl/py-data-science-basics>

Referenzen

Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. *Journal of data warehousing*, 5(4), 13-22.