

# Introduction to Machine Learning and Big Data

## Exercise 03

Volker G. Göhler

TU Bergakademie Freiberg

April 25, 2024

# Machine Learning

# Checklist

- ☒ Frame the problem and look at the big picture
- ☒ Get the data
- ☒ Explore the data to gain insights
- ☒ **Prepare the data to better expose the underlying data patterns to machine learning algorithms**
- ☐ Explore many different models and shortlist the best ones
- ☐ Fine-tune your models and combine them into a great solution
- ☐ Present your solution
- ☐ Launch, monitor, and maintain your system

# Home Work Assignment

- ▶ take one of the datasets out of the group work exercise (not too big, should be a table (csv))
- ▶ explore the data like with the californian housing prices
- ▶ present your findings now (20 mins total)

# We continue where we bugged out last week

► jupyter

# Data Preparation Addendum and Further Steps

► jupyter lab

# Homework Task

- ▶ Look at the data
  - ▶ prepare the data for a ML Algorithm
  - ▶ Think about the usability of the Data for different ML Tasks
  - ▶ Groups: as you see fit
- 
- ▶ <https://tinyurl.com/imlbde3>
  - ▶ Popular Video Games 1980-2023, Kaggle

# Questions?

?



# Reference I