# Ablauf Projekt

1. Git Repository auf PC und Laptop einrichten
2. Jupyter Notebook Datei erstellt
   1. Bibliotheken hinzugefügt
   2. Datensatz eingelsen
   3. Daten die nicht gebraucht werden rausgeschmissen
   4. Features löschen
   5. Spalten für Lesbarkeit umbenennen
   6. balkendiagramm (-> Vorkommen der einzelnen Werte)
   7. Manche Werte sind im Datensatz falsch -> Transformation

One-Hot-Encoding (-> Binäre Daten) phi koeffizient?

Decision tree (Variablennamen checken ab supervised learning)

Decision tree verbessern

Andere Klassifikationsmethoden

Feature selection

Maybe irgendwie overfitting

Stabilität testen

Visualisierung lernkurven

Neuronales netz vielleicht

# Notizen

FAVC: frequent consumption of high caloric food

FCVC: Frequency of consumption of vegetables

NCP: Number of main meals

CAEC: Consumption of food between meals CH20: Consumption of water daily

CALC: Consumption of alcohol

SCC: Calories consumption monitoring

FAF: Physical activity frequency

TUE: Time using technology devices MTRANS: Transportation used

## Features für Supervised Learning

High caloric food

vegetables

daily meals

eat between meals

smoke

physical activity

technological devices

alcohol

transportation