

# Projekt: Binäre Klassifikation

Nicolai Ülger, Niklas Mäschke, Suri Volz

Computerpraktikum, Juli 2020

# Aufgaben

- Implementierung eines Verfahrens zur binären Klassifikation
- Visualisierung für zweidimensionale Datensätze
- Einfügen eines Kommandozeilen-Interface

# K-nearest-Neighbours

Bestimmung der  $k$  nächsten Nachbarn eines Punktes über kd Trees

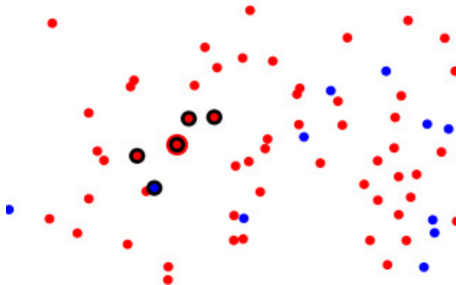


Abbildung 1: knN, wobei  $k=4$

$f_{D,k}$ 

Klassifikation eines Punktes über  $f_{D,k} = \text{sign}(\sum_{j=1}^k y_{ij})$ , wobei  $\text{sign}(0) := 1$ .

# Fehlerklassifikationsrate $R$

Danach wird die Fehlklassifikationsrate  $R_{Di}(f_{D \setminus i, k})$  für jede Teilmenge bestimmt und der Mittelwert gebildet. Der kleinste Mittelwert gibt  $k^*$ .

# Ergebnis

Dieses  $f_{D,k*}$  wird auf die Trainingsdatei angewandt, wobei zusätzlich die Klassifikationsfehelerrate ausgegeben wird.

# Live Demonstration

## Live Demonstration

# Helper Methods

- Die csv-Datei wird ausgelesen und in  $l$  Teilmengen gesplittet
- Die Ergebnisse werden geplottet



# GitHub

Code zum downloaden auf Github:

<https://github.com/uelgerni/KI-Projekt>