

DHBW Karlsruhe, Angewandte Informatik  
Vorlesung: Wissensbasierte Systeme



Dr. Sina Keller ([sina.keller@kit.edu](mailto:sina.keller@kit.edu))  
Felix M. Riese ([felix.riese@kit.edu](mailto:felix.riese@kit.edu))

**Abgabe:** bis 22.04.2018, 23:15 Uhr MESZ

11. April 2018

# Übungsblatt 02

## Klassifikation mit SOMs

Alle Programmierungen sollen in einem Jupyter Notebook **und** einer Python-Datei (**SOM.py**) erfolgen und entsprechend nachvollziehbar dokumentiert werden. Die Dokumentation sowie die Beantwortung der Fragen kann direkt im Notebook im Markdown-Format verfasst werden. Dieses Notebook reichen Sie mit der Python-Datei bis spätestens zum Abgabedatum (**22.04.2018, 23:15 Uhr MESZ**) über das Moodle-Portal zur Bewertung ein. Das Übungsblatt ist Pflicht und bildet einen Teil Ihrer Note für das Teilmodul „Wissensbasierte Systeme“.

**Datensatz:** LUCAS-Datenbank (Download, Dokumentation)

### 1 Implementierung

In diesem Aufgabenblatt geht es um Self-Organizing Maps (SOM). Implementieren Sie eine eigene SOM-Klasse zur Klassifikation von Daten. Es ist nicht erlaubt, vorhandenen Code von SOMs zu verwenden. Optimal wäre eine Ressourcen-schonende Implementierung. Dafür bieten sich als Arrays **numpy.array** an, da Numpy in C implementiert ist.

- Implementieren Sie die Klasse `class SOM` in einer Python-Datei **selforganizingmap.py**.
- Implementieren Sie einen Konstruktor darin: `__init__(self, ...)` für die Klasse.
- Lesen Sie sich in die Funktionsweise einer self-organizing map (SOM) ein.
- Implementieren Sie eine einfache SOM, der man die folgenden Parameter übergibt:
  - **df** : pandas.DataFrame das den Datensatz enthält
  - **label** : string der den Namen enthält der label-Variable(n) zur Klassifikation
  - **nRows** : integer als Anzahl Zeilen des SOMs

- **nColumns** : integer als Anzahl Spalten des SOMs
- e) Implementieren Sie die einzelnen Teile einer SOM in einzelnen Funktionen der Klasse.
- f) Implementieren Sie eine Funktion, mit der die SOM trainiert werden kann (Parameter: Anzahl Iterationen)
- g) Implementieren Sie geeignete Funktionen zur bestmöglichen Visualisierung der Daten. Das ist eine der Stärken einer SOM, hier kann viel herausgeholt werden.
- h) Klassifizieren Sie den Datensatz aus der ersten Übung.

## 2 Klassifikation

- a) Erarbeiten Sie sich ein einfaches Klassifikationsverfahren, mit der die SOM die Daten klassifizieren könnte. Tipp: Mehrheitsentscheidung.
- b) Implementieren Sie eine Funktion, mit der die SOM klassifizieren kann.
- c) BONUS: Erweitern Sie die SOM-Klassifikation für eine bestmögliche Klassifikation.
- d) **Trainieren** Sie beide Klassifikatoren auf dem Trainingsdatensatz und **predizieren** Sie auf den Testdatensatz.

## 3 Evaluierung und Visualisierung der Ergebnisse

- a) **Evaluieren** Sie die Vorhersagen (Prediktionen) aus Aufgabe 2 anhand der Qualitätsmaße, die Sie im ersten Übungsblatt erarbeitet und implementiert haben. **Interpretieren** Sie die Ergebnisse.
- b) **Visualisieren** Sie die Klassifikationsergebnisse in geeigneter Form.
- c) BONUS: **Werten** Sie den Einfluss der Vorprozessierungen aus Übungsblatt 1 aus.