

**1. PCA und Ziffern**

Implementieren Sie Principal Component Analysis (PCA) wie in der Vorlesung bzw. im Tutorium besprochen. Reduzieren Sie die Daten des Digit-Datensatzes<sup>1</sup> auf zwei Dimensionen und plotten Sie f  r alle Paare an Klassen, wie sie im 2D-Raum dargestellt werden. F  r die Berechnung der Eigenvektoren k  nnen sie Bibliotheksfunktionen verwenden, z.B. `numpy.linalg.eigh()` .

F  hren Sie anschlie  end f  r alle Paare von Klassen (auf den 2D-Daten) eine lineare Regression durch und werten Sie auf den Testdaten den Klassifikationsfehler aus. Dazu m  ssen auch die Testdaten entsprechend den Klassen aufgeteilt werden. F  r jeden bin  ren Klassifikator werden nur die beiden relevanten Klassen aus den Test- und den Trainingsdaten verwendet.

**2. PCA und Eigenfaces**

Eine popul  re Anwendung von PCA war traditionell die Gesichtserkennung. Benutzen Sie Ihre Implementierung aus Aufgabe 1 und wenden Sie diese auf den Graustufen-Gesichtsdatensatz<sup>2</sup> an. Entscheiden Sie sich f  r eine Anzahl an Dimensionen (z.B. 20) und visualisieren Sie die entstehenden Hauptkomponenten als Bilder. Dabei werden die einzelnen Werte der Hauptkomponenten als Pixelintensit  ten interpretiert.

**Hinweis:** Bitte bearbeiten Sie die Aufgaben in Zweier-Gruppen und laden Sie alle Ergebnisse (Quelltext + Dokument mit Plots, Tabellen und Erl  uterungen) auf der Vorlesungsseite im Whiteboard hoch<sup>3</sup>. Geben Sie die Namen beider   bungspartner an. Die Bewertung erfolgt bin  r (bestanden/nicht bestanden).

---

<sup>1</sup> <https://web.stanford.edu/~hastie/ElemStatLearn/data.html> → ZIP Code

<sup>2</sup> <http://conradsanderson.id.au/lfwcrop/> → greyscale version

<sup>3</sup> <https://kvv.imp.fu-berlin.de/portal/site/30550a4f-be9f-4be0-8ec2-a35a294ab3a7>