

1. PCA und Ziffern

Implementieren Sie Principal Component Analysis (PCA) wie in der Vorlesung bzw. im Tutorium besprochen. Reduzieren Sie die Trainingsdaten des Digit-Datensatzes¹ auf zwei Dimensionen. F  r die Berechnung der Eigenvektoren k  nnen sie Bibliotheksfunktionen verwenden, z.B. `numpy.linalg.eigh()`. Erstellen Sie Plots der projizierten 2D-Punkte f  r alle Paare von Klassen.

F  hren Sie anschlie  end f  r alle Paare von Klassen auf den projizierten Trainingsdaten eine lineare Regression durch. Jeder bin  re Klassifikator wird nur auf den Daten von zwei Klassen trainiert. Werten Sie den Klassifikationsfehler auf den (projizierten) Testdaten dieser beiden Klassen aus.

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-Gesichtsdatensatz² 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 hoch³. Geben Sie die Namen beider   bungspartner an. Die Bewertung erfolgt bin  r (bestanden/nicht bestanden).

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

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

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