

Blatt 10

In dieser Übung geht es um die Nicht-Negative-Matrixfaktorisierung.

Daten: US Postal Service Handwritten Digits

<http://www.inf.fu-berlin.de/lehre/WS12/ME/ueb/ueb10/usps.ascii.zip>

Aufgabe 1: NMF

Lies den ASCII-Datensatz ein und stelle zunächst ein paar Ziffern dar. Achte darauf, dass die Ziffern richtig herum dargestellt werden, eventuell muss man transponieren. Implementiere die Nicht-Negative-Matrixfaktorisierung. Teste dann Deine Implementierung, indem Du für einige Beispiele versuchst, die Originale zu rekonstruieren. Variiere auch die Anzahl der Prototypen, die Du zur Rekonstruktion verwendest.

(10 Punkte)

Aufgabe 2: Klassifikation und Vorverarbeitung

Verwende knn, um die Daten zu klassifizieren. Vergleiche die Erkennungsraten für:

- Vorverarbeitung mit NMF
- Vorverarbeitung mit PCA
- keine Vorverarbeitung

(10 Punkte)

Aufgabe 3: Matrix W

Stelle die Spalten aus Matrix W als Bilder dar. Verwende eine gute Größe für die Anzahl der Spalten. Diese kann sich z.B. aus Aufgabe 2 ergeben (Klassifikation funktioniert besonders gut) oder aus Aufgabe 1 (Rekonstruktion der Originale liefert ein visuell zufriedenstellendes Ergebnis trotz merklicher Dimensionsreduktion).

(5 Punkte)