

ITU Computer and Informatics Faculty
BLG 454E Learning From Data, Spring 2018
Homework #3
Due May 3, 2018 10pm

Kadir Emre Oto
(150140032)

1: PCA

- (a) (3 pts) What are the main motivations for reducing a datasets dimensionality?
- (b) (3 pts) How can you evaluate the performance of a dimensionality reduction algorithm on your dataset?
- (c) (2 pts) What do you say about the performance of PCA in Figure 1 in terms of classification?
- (d) (2 pts) What is/are drawback(s) of PCA?
- PCA may not success if the input data lies on a complex manifold.
 - PCA makes an assumption that the normality of the input space distribution is comparative.
 - For better results, the input data should be real and continuous.
- (e) (40 pts) Implement a PCA projection on given the data.txt. The last attribute of the data.txt is the class label, range from 0..9.

2: SVD

- (50 pts) You are going to look at compressing the given RGB image, data.jpg, through computing the singular value decomposition (SVD). Each channel (red, green, blue) has 1537 2500 pixels which is a 1537 2500 matrix A .
- (a) (35 pts) Find the SVD of A (one for each channel).
- (b) (15 pts) Display the original image and image obtained from a rank (term) of 1, 5, 20, 50 SVD approximation of A as shown in Figure 3.