Theoretical task 6.

Recommendations: all solutions should be short, mathematically strict (unless qualitative explanation is needed), precise with respect to the stated question and clearly written. Solutions may be submitted in any readable format, including images.

1. There is a dataset X containing 10^6 100-dimensional objects. It has been compressed with PCA (full transformation matrix $P \in \mathbb{R}^{100 \times 100}$ is stored in memory) to compressed dataset $\hat{X}$ which contains 10^6 20-dimensional objects.

    (a) Can one transform new unseen dataset Y to the same compressed feature space? How?

    (b) Can one reconstruct the original uncompressed dataset given only P and $\hat{X}$? How?

2. Consider a dataset which was compressed with PCA from D to d features ($D >> d$). What would happen if we train and run transformation with another PCA (with $c \le d$ components) on compressed dataset once again? Why?

3. Provide an example of the two-dimensional ($d = 2$) dataset in the binary classification problem for which the preliminary application of PCA compression to dimensionality $d = 1$ would hurt the classification accuracy dramatically. Explain, why PCA can hurt the classification accuracy?