Machine earning

HW1 - Dimensionality Reduction
Spring 2025

1. Write your name below to acknowedge that you should not include any hand-written solution, code, or pictures (unless explicity requested), in your HW solutions. That everything should be typeset and converted to PDF. (1pt)

2. Convert the follow features to binary, using their means: (2pts)

$$X = \begin{bmatrix} 2 & 1 \\ 3\,5 & 2 \\ 0 & 3 \\ 12 & 4\,1 \end{bmatrix}$$

3. Given the following enumerated features, create a new *one hot encoded* set of features. You may assume that the set of values you see for each feature constitute the entire unique set of enumerations (2pts).

$$X = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

4. Consider the following dataset consisting of 10 observations, each with two features:

$$X = \begin{bmatrix} 0 & 2 \\ 0 & 1 \\ 0 & 2 \\ 0 & 3 \end{bmatrix}$$

(a) What are the principle components of the observed data X for use in PCA? While you **M Y** use a utility function like *eig* or *svd* to verify your answers, the directions for this particular dataset should be rather straightforward. (5pts).

(b) What are the angles of the principle components relative to the original coordinate system? That is, how is the rst principle component a rotation of the x-axis (the rst original feature)? Hint: Use trigonometry using the values of the elements of the components. (3pts).

(c) If we were to project our data down to 1-D using the principle component, what would the new data matrix X be (5pts).