

In The Name of God, The Compassionate, The Merciful

Name:

Student ID#:

Statistical Pattern Recognition (CE-725)
Department of Computer Engineering
Quiz #1 (Overview & Introductory Materials) - Spring 2011

1. (5 points) What is the difference between following terms: "Data Reduction", "Dimension Reduction", "Feature Selection", "Feature Extraction" and "Sampling"?

2. (5 points) Suppose we have the following data set. All variables, including the labels, are binary.

X1	X2	X3	Y
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	1
1	0	1	0
1	1	0	1
1	1	1	0

Suppose we can't store all 3 features in the memory and hence we have to remove one of the features. Which of the 3 features, if any, do you suggest for removal? Why? Why not others?

3. Bonus (5 points) Explain the role of feature dimensionality in the problem of model overfitting?