BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

Department of Computer Science and Information Systems

I SEMESTER 2017-2018

BITS F464 – Machine Learning

12th October, 2017 Mid-semester Exam (CLOSED BOOK) Weightage: 30%

---

1. Compare and contrast Principal Component Analysis (PCA) & Singular Value Decomposition (SVD). Give your answer in tabular form.

[3]

2. (a) Fisher's Linear Discriminant (FLD) gives the best discrimination of data. Explain. Give a 2-dimensional data set having two classes, which shows that separation between the means of classes is not a good criterion for preserving discrimination. How does FLD handle this?
   (b) Principal Component Analysis (PCA) gives the best representation of data. Explain. Give a 2-dimensional data set having two classes, which shows that PCA does not preserve discrimination. PCA preprocessing suits which classification algorithm?

[3+3]

3. In the perceptron model, suppose there are two neurons. In one particular case, all the three inputs are zero. The threshold for firing a neuron is 0.25. It is required that one neuron fires and the other does not. Give a solution to the problem. Why it is not recommended that the threshold be made adjustable? Draw the perceptron network.

[3]

4. Consider two perceptron classifiers. Each perceptron classifier has only a single neuron. The first perceptron classifier is built using the points C1: (0,1), (1,0), and C2:(1,1), while the second perceptron classifier is built using the points C1:(0,1), (1,0), and C2:(0,0). Use a fixed bias input of -1 and a learning rate of η=0.20. The initial classifier in both the cases is the vertical line x_1=0.5. Process the points in the given order. The neuron fires for class C1 in both the cases. Can the two perceptron classifiers be combined to solve the XOR problem?

[5]

5. Do Support Vector Machines (SVM) suffer from overfitting? Pictorially illustrate and justify your answer. What are the ways of overcoming overfitting?

[3]

6. Consider an ensemble of 5 base classifiers, C_i, with error rates ε_i = 1/(2+i). Determine the error rate of the ensemble.

[5]

7. Consider a two-dimensional input space $X \subseteq \mathbb{R}^2$ together with the feature map:

$$\emptyset: \boldsymbol{x} = (x_1, x_2) \longmapsto \emptyset(\boldsymbol{x}) = \left(x_1^2, x_2^2, \sqrt{2}\, x_1 x_2\right) \in F = \mathbb{R}^3$$

   (a) Find the Kernel corresponding to $\emptyset$ and F
   (b) Also find the Kernel function corresponding to:

$$\emptyset: \boldsymbol{x} = (x_1, x_2) \longmapsto \emptyset(\boldsymbol{x}) = (x_1^2, x_2^2, x_1 x_2, x_2 x_1) \in F = \mathbb{R}^4$$

   (c) What conclusions you can draw from the results of (a) & (b)

[2+2+1]
