BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

Department of Computer Science and Information Systems

I SEMESTER 2017-2018

BITS F464 – Machine Learning

09th December, 2017 Mid-semester Exam (Make Up) Weightage: 30%

---

1. What are the problems associated with Curse of Dimensionality (CoD)? In context of Machine Learning, how we work our way around CoD? Name the ML algorithms, covered so far, which do not suffer from CoD.

[4]

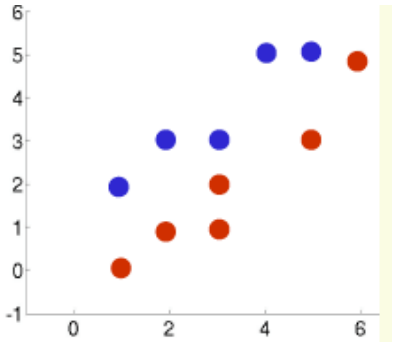
2. If you walked into a a hardware store and asked – have you got fork handles?, then you would be surprised to be presented with four candles!! Fork handles and four candles are acoustically almost identical.
The salesperson knows that he sells many more many more candles than fork handles. As a result, he/she probably did not hear the words fork handles, but instead heard four candles.
   (a) Model the above problem using Bayes' Theorem. Some number to help you are given below:
   Likelihood of four candles = 0.6
   Likelihood of fork handles = 0.7
   (b) Does likelihood give you any answers?
   Priors:
   p(four candles) =0.9
   p(fork handles) =0.1
   (c) What is the MAP estimate of the phrase that was spoken? Given: p(data)=0.61

[6]

3. Mathematically define Kernel functions. Find a Kernel corresponding to the mapping: $\varphi: (x_1, x_2) \rightarrow (x_1^2, x_2^2, \sqrt{2}\, x_1, \sqrt{2}\, x_2, \sqrt{2}\, x_1 x_2, 1)$
Which methods discussed in the course so far can be Kernelized? List them and graphically depict the difference between the normal (without Kernels) and their corresponding Kernelized versions.

[4]

4. Fisher's Linear Discriminant (FLD) vs. PCA
Data:
Class 1 has 5 samples c1=[(1,2),(2,3),(3,3),(4,5),(5,5)]
Class 2 has 6 samples c2=[(1,0),(2,1),(3,1),(3,2),(5,3),(6,5)]



For the data given above, which is better in context of classification, PCA or FLD? Pictorially justify your answer. Calculate J(**w**) for both.

[6]

5. How can one determine whether a given data is linearly separable or not? Why it is important to figure this out? Give specific examples.

[4]

6. (a) What is the main difference between Bayesian and Frequentist approaches to probabilitiy?
   (b) What criticisms the Bayesian Approach to probability face as compared to the frequentist approach? In Machine Learning context, how does the Bayesian Approach overcomes them?
   (c) Compare and contrast Maximum a posteriori (MAP) and Maximum Likelihood Estimation (MLE).

[2+2+2]