CS551: Introduction to Deep Learning

Mid Semester, Spring 2018

IIT Patna

Attempt all questions. Do not write anything on the question paper.

Time: 2 Hrs Full marks: 30

1. Prove or disprove: $tr(AB) = tr(A) \times tr(B)$ where A and B are $n \times n$ matrices and tr denotes the trace. (3)

2. A diagnostic test has a probability 0.95 of giving a positive result when applied to a person suffering from a certain disease, and a probability 0.10 of giving a (false) positive when applied to a non-sufferer. It is estimated that 0.5% of the population are sufferers. Suppose that the test is now administered to a person about whom we have no relevant information relating to the disease (apart from the fact that he/she comes from this population). Calculate the following probabilities: (a) that the test result will be positive; (b) that, given a positive result, the person is a sufferer; (c) that, given a negative result, the person is a non-sufferer; (d) that the person will be misclassified. (1+1+2+2)

3. Consider a neural network where the last output layer uses softmax as activation function for some classification problem. Following is an observation on three training examples where o_i denotes the probability of each class from softmax and t_i denotes the labeled target for the examples. Find out mean cross entropy error for these three examples. (5)

o_1	o_2	o_3	t_1	t_2	t_3
0.1	0.3	0.6	0	0	1
0.2	0.6	0.2	0	1	0
0.3	0.4	0.3	1	0	0

4. Consider maximal margin classifier for the following toy data set which has two features X_1 and X_2. Y is the target label . (a) Sketch the optimal separating hyperplane, and provide the equation for this hyperplane. [Equation for hyperplane may be derived logically or mathematically.] (b) On your sketch, indicate the margin for the maximal margin hyperplane. (c) Indicate the support vectors for the maximal margin classifier. (3+2+1)

Sl. No.	1	2	3	4	5	6	7
X_1	3	2	4	1	2	4	4
X_2	4	2	4	4	1	3	1
Y	R	R	R	R	B	B	B

5. Given a set of m points $\{\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(m)}\}$ in $\mathbb{R}^n$ and we want represent these points in k dimension where $k < n$. Propose a suitable methodology for it. (10)