SML 2024, Monsoon, Endsem, Duration 2 hours

Q1. a. Clustering. Consider the data $D = \{x_1, x_2, y\} = \{(0,1,1),(0,2,1),(4,0,-1),(5,0,-1)\}$. Assume the initial seed points as $(0,0)$ and $(3,0)$. Perform one iteration of k-means clustering. [2] [CO1]

b. What are updated centroids after (a)? [1] [CO1]

c. Find the purity of clustering obtained in (a). [1.5] [CO1]

Q2. MLP. Consider the following network in Figure 1. The weights of the connections are given on the edges. Where the weights are missing, the weights are 0. There are no activation functions on hidden layer. Use the input (x_1, x_2, y), y is the label. Assume there are two classes {0,1}. Find [CO1]

a. the multiclass cross-entropy loss for the input $(1,1,1)$. Note that h_1 and h_2 are before activation. [1.5]

b. mse loss for the input $(1,1,0)$. [1.5]

c. mean and unbiased variance for batch normalization when the inputs are $(1,1,0),(-1,-1,1)$ [2]

Q3. Gradient boosting. Find the labels, when gradient boosting is preformed, in terms of residues when the loss is given by [2] [CO1]

$$L(y, F(x)) = e^{.5(y-F(x))^2}$$

Q4. Consider a hypothetical pdf $e^{-\theta x}$. Note that it is not a real pdf but for this question, we will assume it behave like a pdf. Assume n iid. samples. [CO1]

a. Give an expression for likelihood function. [1]

b. Find an expression for MLE of θ. [2]

Q5. Consider a fully connected network with two input nodes, one hidden node, and two output nodes. Hidden node has a sigmoid activation. The weights connecting input to hidden node are $W = [w_1, w_2]^\top$. Show that the input $x = W/\sqrt{W^\top W}$ will make the hidden node maximally active? [2] [CO3]

Q6. a. Consider the regression model $\hat{f}(x) = wx + b$. Using the dataset $D_1 = \{(x,y)\} = \{(1,0),(2,1)\}$, find w, b. [1.5] [CO2]

b. Find the bias of this regressor when two datasets D_1 from (a), and $D_2 = \{(x,y)\} = \{(-1,-1),(-2,0)\}$ are available. Expression should be in terms of x. Assume the true model from which data is generated is $y = x + \eta$, η is white noise. Thus, true $f(x) = x$. [2] [CO2]

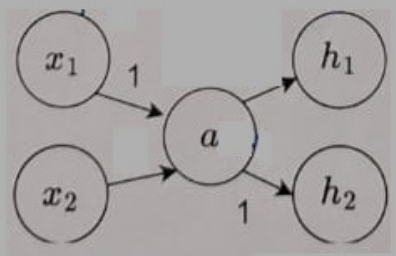

Figure 1: