ECSE 4965/6965

Introduction to Deep Learning

**Assignment #8**

Due 4 pm, May 1

Given the discussions on Bayesian Neural Networks (BNN) in the class, let  be the parameters of a NN and its prior distribution be specified as p(****)~N(0,I), where I is an identity matrix. For a NN with input **X** **∈**RN and one dimensional output Y.

Given training data D={Xi, Yi}, i=1,2,..,M

1. For a regression problem, i.e., Y **∈**R1, give the objective function that estimates  by maximizing the posterior probability of  (not maximum likelihood estimation), assuming Y|X follows N(f(x,), 2 (x,)), where f(x,) is the discriminant function of the last hidden layer and 2 (x,)) is the variance of Y. Show how the variance of Y|X is determined, given the estimated For both questions, provide the details for each term.
2. For a binary classification problem, i.e., Y **∈** B1 and B **∈** {0, 1}, give the objective function that estimates  by maximizing the posterior probability of  (not maximum likelihood estimation), assuming Y|X follows Bernoulli((f(x,)), where  () is the sigmoid function. Show how the variance of Y is determined, given the estimated For both questions, provide the details for each term.

Note:

1. Students taking 4000 level class only need attempt problem 1.
2. To balance between homework and programming assignment, this assignment will be counted towards a programming assignment.