Mid Term Exam

Date: 07/05/2023  
Time: 80 Minutes

MSCS(2022-2024)

Semester-2

RollNo:.---------------------------------------------------- Name:------------------------------------------------------------

Write the Precise answer. Irrelevant detail is highly discouraged

1. If the weights in a Deep Neural Network are initialized to zeros instead of being randomly initialized, what effect would it have? (2)
2. Write out the formula for the "He initialization" technique, which is a type of random initialization used for initializing weights in neural networks. (2)
3. What does it mean for a model to have both high bias and high variance, and could you provide an example? (2)
4. Dropout regularization is a technique used to regularize deep neural networks. Explain the complete process involved in using dropout regularization? (2)
5. What are the problems of vanishing and exploding gradients in deep neural networks (DNN), and what methods can be used to prevent these issues? (4)
6. What are the advantages and disadvantages of full-batch gradient descent and mini-batch gradient descent, and what would be an appropriate size for the mini-batches? (3)
7. How does the hyper parameter Beta affect the process of Exponentially Weighted Averages? Additionally, it has been observed that EWA initially performs poorly but stabilizes over time. What measures can be taken to address this issue? (4)
8. Write the pseudo code of Adam Optimization algorithm (3)
9. Define learning rate decay, and what role does it play in the convergence of a machine learning model? Additionally, could you list two techniques for learning rate decay that have been proposed by researchers? (2)
10. Write down the list of priorities when tuning the following hyper parameters: Alpha, Beta, Beta1, Beta2, Epsilon, number of layers, number of hidden layer units, learning rate decay, and mini-batch size? (2)
11. Write ReLU and Leaky ReLU activation functions, including their ranges? Additionally, what is the range of their derivative functions? Finally, how does the issue of "dying ReLU" affect the ReLU activation function, and how is this problem addressed by Leaky ReLU? (4)