EECS545 Lecture 19 Quiz Solutions

1. Consider that you are trying to train an ML model, and the training loss does not seem to be dropping (assume we trained this model for a lot of epochs). Which of the following can be the possible reasons? (Choose all that apply)

   (a) Learning rate is too low.

   (b) Weights are not properly initialized.

   (c) The coefficient for the regularization term is too low.

   (d) The dataset is too small.

   Solution: (a), (b).
   (c) and (d) would drop the learning rate faster than expected.

2. Which of the following methods can be used to reduce the high variance problem? (Choose all that apply)

   (a) Using dropout.

   (b) Using the aggregated result from multiple models.

   (c) Using a high coefficient for the regularization term while training.

   (d) Changing the weight initialization.

   Solution: (a), (b), (c).
   Weight initialization (d) does not contribute to the high variance issues, in general.

3. Which of the following is true about ablative analysis? (Choose all that apply)

   (a) It generally tries to explain the difference between the poorer baseline and the current performance.

   (b) It is performed by removing components from your system one at a time.

   (c) It focuses more on explaining the difference between the current and perfect performance (100%) compared to the error analysis.

   (d) It helps to understand the contribution of the component to the overall system.

   Solution: (a), (b), (d).
   (c): Error analysis focuses primarily on explaining the difference between the current performance and the perfect performance (100%).

4. (True/False) It is always possible to tell if an algorithm has converged and finished training by looking at the iteration/loss plot.

Solution: False. Your optimizer sometimes may train with a small learning rate in the beginning (e.g., warm-up when training a Transformer from scratch). Also, your objective sometimes does not well-matched with your problem. It is recommended to check whether you are optimizing the right function.

5. (True/False) There exists one optimization algorithm that would work on any problem. So, we do not need to change the optimizer from the beginning to the end for whatever problem settings.

Solution: False. You can start with a single optimizer, like ADAM, but it doesn't mean this optimizer will work for all problem settings. It is always recommended to check with multiple plausible options.