CSE-6363-007 Assignment 3 (Spring 2024)

Problem 2: [20 points]

Let us review the bias-variance decomposition first. The intuition behind it is straight-forward: if the model is too simple, the learnt function is biased and does not fit the data. If the model is too complex then it is very sensitive to small changes in the data. If we were able to sample a dataset D infinite many times, we will learn different g(x) for each time, and get an expected hypothesis $\bar{g}(x)$. So bias means the difference between the truth and what you expect to learn. It measures how well our model can approximate the truth at best. However, it is impossible to sample the training dataset multiple time, so variance means the difference between what you learn from a particular dataset and what you expect to learn. Now please answer the following questions:

True or False:

(i) If a learning algorithm is suffering from high bias, adding more training examples will improve the test error significantly.

False. High bias usually indicates that the model is too simple and underfitting the data. Adding more training examples will not significantly improve test error for a high-bias model. Instead, increasing the model complexity is likely needed to reduce bias.

(ii) We always prefer models with high variance (over those with high bias) as they will be able to better fit the training set.

False. High variance models overfit the training data and do not generalize well to unseen data. A balance between bias and variance is preferred to achieve good prediction accuracy on both the training set and the test set.

(iii) A model with more parameters is more prone to overfitting and typically has higher variance.

True. More parameters can allow a model to capture more complex patterns, which can lead to fitting the training data too closely and thus having higher variance.

(iv) Introducing regularization to the model always results in equal or better performance on the training set.

False. Regularization typically reduces model complexity, which can decrease training performance as the model may not fit the training data as closely as before. However, it often increases generalization to new data.

(v) Using a very large value of regularization parameter λ cannot hurt the performance of your hypothesis.

False. A very large value of the regularization parameter can lead to underfitting, where the model is too simple to capture the underlying pattern in the data, thus hurting performance.