CS580 Take-Home Final

Yang Zhang 11529139

1. (a) Adapting exact imitation learning algorithm as the base. Define two lists called sampleLostList and sampleCount. SampleLostList is used to track of the current loss (result returned from calling black-box function). The list sampleCount is used to count how many times the loss difference of a specific sample within a threshold. (If the difference is within the threshold, the sample is saturated. I.e. there is not a lot can be learned from this sample). Then, modified the base algorithm by removing any sample that has sampleCount 5. If a the difference of sample within the threshold in continues 5 times, we are confident to say that there is nothing we can learnt from this sample. Therefore, removing such kind of samples can reduced the amount of calls of black-box function, which is sample efficient.

   (b) If the black-box function is stochastic, it is possible that returned loss is inaccurate. In other words, the variance of each sample is huge. To deal this problem, Bagging algorithm can be applied. Bagging can minimize the variance, and return an average weighted model.

2. (a) Initialize weight vectors based on different type (first-order, second-order, etc.). For each sample, generate the search space by using flip-bit method. For each variant of the sample, call the black-box function to test the change bits. (If Yes returned, the bit flip has positive effect, add positive value to the weight according to the bits changed. Vice versa)

   (b) If the black-box function is stochastic, it is possible that returned wrong answer. Vote method can be applied. For each variant of a specific sample, call the black-box function N times. Make the decision by selecting the answer with most counts.

3. Combine the cascades and search-based methods in the following way:
   Repeat the following steps:
   i.   Using cascades to train a base classifier that can achieve partial of the goal (for example 60% of the expected performance)
   ii.   Then tuned the base classifier using Search-based methods for N iterations
   iii. If the tuned classifier still doesn't meet the goal, refine the tuned classifier by using cascades again