CS 269/362 Machine Learning: Homework 1
Learning Foundations
Due: 1/19/2016
50 Points Total   Version 1.0

1 (3 points) Explain why you agree or disagree with the following statement.
It is always best to select a hypothesis class that contains the optimal hypothesis.

Answer: I disagree with this statement. One reason is that it may require knowledge of many probabilities to estimate based on the background knowledge. Another reason is that it may take significant computational cost.

2 (3 points) True or False: *An infinite hypothesis class always contains the optimal hypothesis.* If true, why? If false, give a counter example.

Answer: False. The optimal hypothesis and the infinite hypothesis class may be in different dimensions. For instance, the infinite hypothesis class may consist of two parameters. So it's just a linear function. However, the optimal hypothesis may be a cubic function with three parameters.

3 (3 points) Consider the following development scenario. A researcher collects 1000 labeled examples for learning, dividing them into a training set (500 examples) and a test set (500 examples). To develop a classifier, the researcher experiments by adding new features. To guide feature construction, the research tests each set of features by training on the train data and testing on the test data, measuring the resulting change in accuracy and only keeping features that help. When the researcher is finished, he collects 1000 new labeled examples and evaluates the classifier on these new examples.

Do you expect accuracy on these 1000 new examples to be the same, better or worse than the original 500 test examples? Why?

Answer: It may be either the same, better or worse than that of the original test examples depending on what the one thousand new examples look like. For instance, if the new examples doesn't behave like the original ones, the accuracy will be worse because the features may not work. if the new examples are almost same with the 500 test examples, in this case, it will be the same. If, however, the new examples occasionally are compatible with the constructed features perfectly, in this case, it will be better.

4 (3 points) At the start of the semester, you arrive home to find three packages containing your three new textbooks. Each package was sent via Fedex or UPS with equal probability. Of the three packages, one of the packages is a brown box delivery by UPS. What is the probability that you have one UPS package and two Fedex packages? Why?

Answer: 0.25. Since you have three packages and one of them is from UPS, we should calculate the probability of the rest two packages are from Fedex. So it should be 0.5*0.5=0.25.

5 (3 points) True/False (and why): Suppose you know your hypothesis class contains the optimal hypothesis, and you observe that changing any one small part (e.g., a single weight) of your current hypothesis makes it worse than before. You can safely conclude that the current hypothesis must be optimal.

Answer: False. It may be the local minimum rather than the global minimum. For instance, the current hypothesis may be 0.2+0.4x. In this case, if you change it to 0.3+0.4x or 0.2+0.5x, it may be worse. However, the best hypothesis is 0.5+0.8x, which obviously shows that the current hypothesis isn't the optimal despite it's better than some of other similar hypothesis.