CS 6362 Machine Learning, Fall 2017: Homework 1

Jiani Li

Question 1: Disagree.

(a) A hypothesis class that contains the optimal hypothesis will be every large. And to search the optimal hypothesis will be very hard.

(b) To get the optimal hypothesis, we need large enough data. Limited data will case over-fitting, high variance.

Question 2: False.

Counter example: Suppose the optimal hypothesis is $h(x) = x^3$, but the hypothesis class is linear functions (Linear hypothesis class contain an infinite number of functions).

Question 3: Worse.

Because overfitting problem. The 1000 new example will have lower accuracy.

Question 4: $\frac{3}{7}$.

The probability of one package being delivered by Fedex $P(F) = \frac{1}{2}$, being delivered by UPS $P(U) = \frac{1}{2}$. The probability of any outcome in the domain {UUU, UUF, UFU, UFF, FUU, FUF, FFU, FFF} is $\frac{1}{8}$.

$$P(3F) = \frac{1}{8} \tag{1}$$

Given one of the packages is delivered by UPS, the probability of one UPS and two Fedex packages is

$$P(1U, 2F|\text{at least } 1U) = \frac{P(1U, 2F)}{1 - P(3F)} = \frac{\frac{3}{8}}{\frac{7}{8}} = \frac{3}{7} \tag{2}$$

Question 5: False.

The hypothesis may be optimal for the current data set. But the data set may be limited or biased, so that the hypothesis may not be optimal for another data set and thus may not be the optimal hypothesis.