CS589 HW1

Yifu liu

February 2018

1 Qusetion 1

Let K = 1, which means in this case, for a single point, that always find itself to be the result for the point. So that the model will be really close to the training data and the bias will be in the lowest point when K = 1. When k increases, the bias will increase.

2 Question 2

2.1

$P(A|B) = \frac{P(A \cap B)}{P(B)}$, if P(B) $\neq 0$
$P(B|A) = \frac{P(A \cap B)}{P(A)}$, if P(A) $\neq 0$
$P(A) = P(A|B) * P(B) = P(B|A) * P(A)$
$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$, if P(B) $\neq 0$

2.2

posterior = $\frac{prior * likelihood}{evidence}$

Because the denominator does not depend on C and the values of the features are given, so that the denominator is effectively constant. The numerator is equivalent to the joint probability model.

2.3

Omitting the constant terms will not affect the result of the function, since during the calculation, we only need to get the relationship of each term, but we don't have to get the exact probabilities and that will be ordered the same relative to each other.

2.4

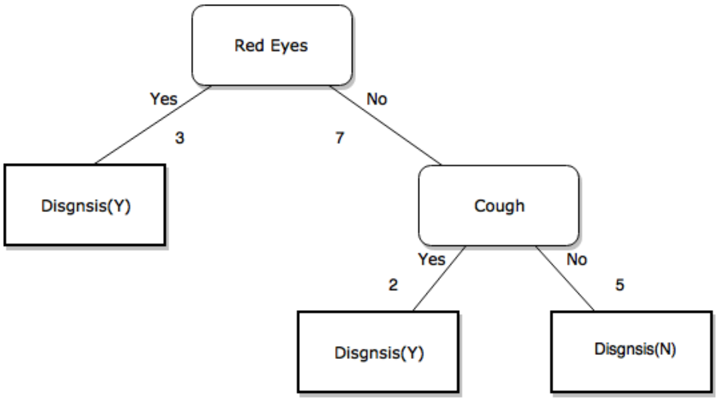
From the PPT slides, NBCs assume that all features are probabilistically independent given the value of the class variable. In practice, some features are correlated, so they probably have similar distributions, so the NBCs tend to push probability values to the extremes.

3 Question 3

3.1

Decision tree is discriminative model. Since for every layer of the tree, the conditional model always is given an unobserved variable y and an observed variables x. So this can be done with conditional probability distribution $P(y|x)$.

3.2



4 Question 4

4.1

Bias(DEC), Variance(INC). If add a new decision node at the bottom of the tree, the decision boundary will be more similar or close to the features of training data, that will lead overfiiting and lower the bias.

4.2

Bias(INC), Variance(DEC). This question is against with question a). Since we cut(prune) the leaf, so the decision boundary will not capture the features of the training set, so the Bias will be increased and the Variance will go the other way.

4.3

Bias(DEC), Variance(DEC). Increase the data set will improve the accuracy of the model, so Bias decreases. The first tree is more shallower than the new tree, so the variance will be decreased.

4.4

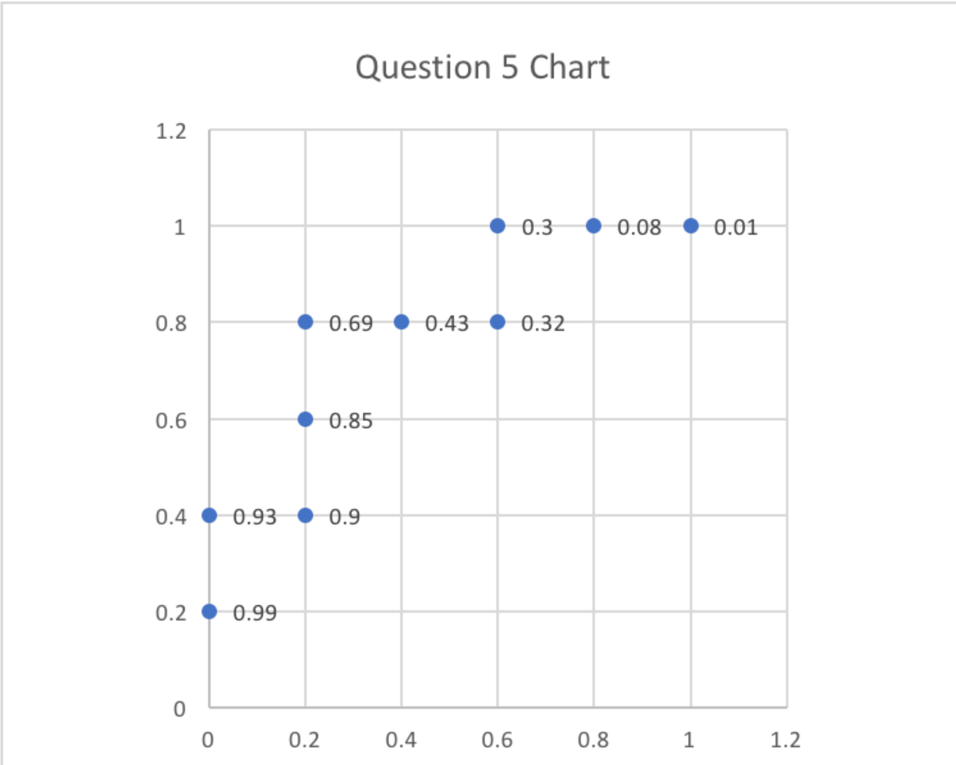
Bias(INC), Variance(NO). Changing the value of leaives, that will lower the similarity between the decision boundary and actual data, so the Bias will be increased. However, we cannot say anything about Variance since the Bias and Variance don't have absolute relationships, and changing the data, we cannot say anything about whether the classification line will close or away from the training data.

4.5

Bias(NO), Variance(NO). Renormalizing will not affect the bias and variance.

5

5.1



Note:**X-axis is False Positive Rate(X); Y-axis is True Positive Rate(Y);**