ECE M146 — Discussion 6

Introduction to Machine Learning

Instructor: Lara Dolecek

TA: Zehui (Alex) Chen

1. **Naive Bayes Example** You're stuck in a forest with nothing to eat. Suddently, you spot a mushroom but you don't know if its poisonous. Luckly, you've studied some mushrooms as part of a class to fulfill your undergraduate requirements. Your previous knowledge is summarized by the following chart:

Sample #	IsColorful	IsSmelly	IsSmooth	IsSmall	IsPoisonous
1	0	0	0	1	1
2	0	0	0	0	0
3	1	0	1	1	1
4	1	0	0	0	1
5	0	0	1	0	0
6	0	0	1	0	0
7	1	1	0	0	1
8	1	1	1	0	1
9	0	1	1	0	?

   (a) Build a Naive Bayes classifier to classify the unknown mushroom.

(b) Repeat (a) with Laplace smoothing.

2. Discriminative v.s. Generative
   So far, we have learned two approaches for binary classification in class. The generative approach model the prior $P(C_i)$ and class conditional distribution $P(x|C_i)$. The discriminative approach model $P(C_i|x)$ directly. Taking the Naive Bayes classifier for binary feature and label as an example. It model the class prior as a biased coin and model the class conditional distribution for each feature and each class also as a biased coin. From page 122 of the book *A Course in Machine Learning*, we learned that to make a decision, we can use $\text{sign}(w^T x + b)$ for some w and b. Learning w and b is then a discriminative approach!

   (a) What is w and b in terms of the parameters of Naive Bayes classifier?

   (b) Suppose we have D features, how many parameters does the Naive Bayes classifier have? How many parameters does the linear model have?

   (c) Can you compare discriminative approach with generative approach given the above example?

3. More about Discriminative v.s. Generative
   Let $p(x|C_1) \sim 0.4\mathbb{N}(0.2, 0.1) + 0.6\mathbb{N}(0.5, 0.1)$. Let $p(x|C_2) \sim \mathbb{N}(0.7, 0.1)$. In MATLAB, plot the two class conditional distribution and find the decision boundary. Let $P(C_1) = P(C_2) = 0.5$, what is the equation to find the posterior distribution for C_1 and C_2. Find and plot the posterior distribution for C_1 and C_2.
   Find the maximum likelihood decision boundary using both the class conditional distribution and the posterior distribution. Comment on your observation.