

CS 411: Artificial Intelligence I  
Spring 2018  
Homework 3  
Due: May 6th, 11:59pm (via Blackboard)

## Programming Portion

This portion of the assignment may be completed individually or in groups of 2.

Complete Project 5 at <http://ai.berkeley.edu/classification.html>, Questions 1–6 (25 points).

Your code for files `perceptron.py`, `mira.py`, `answers.py`, `perceptron.pacman.py`, and `dataClassifier.py` should be submitted to Blackboard for evaluation.

**Submitted code must be your own (or you and your partner's) code and should not be copied from any other source.** We will check for similarity to other submissions and existing resources available on the web for any cheating.

## Written Portion

This portion of the assignment must be completed individually.

For the first four questions, consider the following dataset:

Shape	Color	Label
Circle	Blue	1
Circle	Green	0
Diamond	Blue	0
Diamond	Green	0
Diamond	Green	0
Diamond	Red	1
Square	Blue	1
Square	Red	1
Square	Red	1

1. What are the parameters of the Naïve Bayes model estimated using maximum likelihood estimation? **(3 points)**
2. What is a testing example (Shape, Color) for which the maximum likelihood Naïve Bayes model provides an undefined posterior label probability (i.e., 0/0)? **(2 points)**
3. What are the parameters of the Naïve Bayes model using Laplacian smoothing with a pseudo-count of 1? **(3 points)**
4. What is the posterior label probability for your example from part (2) under the Laplacian-smoothed Naïve Bayes model? **(3 points)**
5. Consider the Perceptron algorithm with parameters  $\alpha_0, \alpha_1, \alpha_2$  for the following binary-labeled dataset with inputs  $x_1$  and  $x_2$  and label  $y$ ,  $(x_1, x_2, y)$ , with positive (+) and negative (-) labels:  
(1,4,+)  
(3,2,+)  
(1,2,-)  
(2,1,-)  
If the algorithm predicts the positive class (+) if  $\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 > 0$ , indicate which examples produce errors and the resulting updates for the weights when (repeatedly) iterating over these examples in order. What are the final parameters when the algorithm converges? **(8 points)**
6. Provide a dataset with 12 datapoints for which the K-means algorithm with  $k = 4$  can produce a non-optimal clustering. First, show the resulting non-optimal clustering and explain why the algorithm has converged. Next, show the optimal clustering. **(6 points)**