CS 59000SA **Fall 2018**

Assignment 3

Instructor. Ninghui Li

Due: 11:59pm Oct 12th, 2018

Please typeset your homework solutions, ideally using LaTeX. No hand written assignment will be graded.

1. **Decision Tree** Consider the following snapshot of a survey about what makes students interested in security. The last column stands for the (prediction) label and the rest columns are attributes.

CERIAS Student	Took CS555	Took CS526	**Interested in Security**
Y	Y	Y	Y
Y	Y	N	Y
N	N	N	N
N	Y	Y	Y
Y	N	N	N
N	Y	N	N
Y	Y	Y	Y
Y	N	Y	Y
N	N	Y	N
N	Y	Y	Y

   Draw two **full decision trees** based on information gain and Gini impurity, you must show intermediate steps to receive full credits.

2. **SVM** Bob is a CS Ph.D. student at Qvsevf University. Bob recalled from his undergraduate experience and realized CS and Math students have distinguishing features: lines of code written and number of derivatives calculated. Bob reads the SVM slides from our CS590SA and decides to start off on an epic journey of classification using SVM. Bob collected a few data points from friends and made a plot(Figure 3.1: SVM Problem) for it. Each blue dot stands for a Math student and each red dot stands for a CS student.

   - Compute the decision boundary. Express the decision boundary using a mathematical formula. Justify your answer.
   - What is the width of the margin? Justify your answer
   - Bob received a funny data point (-10,5) from a CS student. Help Bob to build a kernel so the new SVM can classify all the data points.

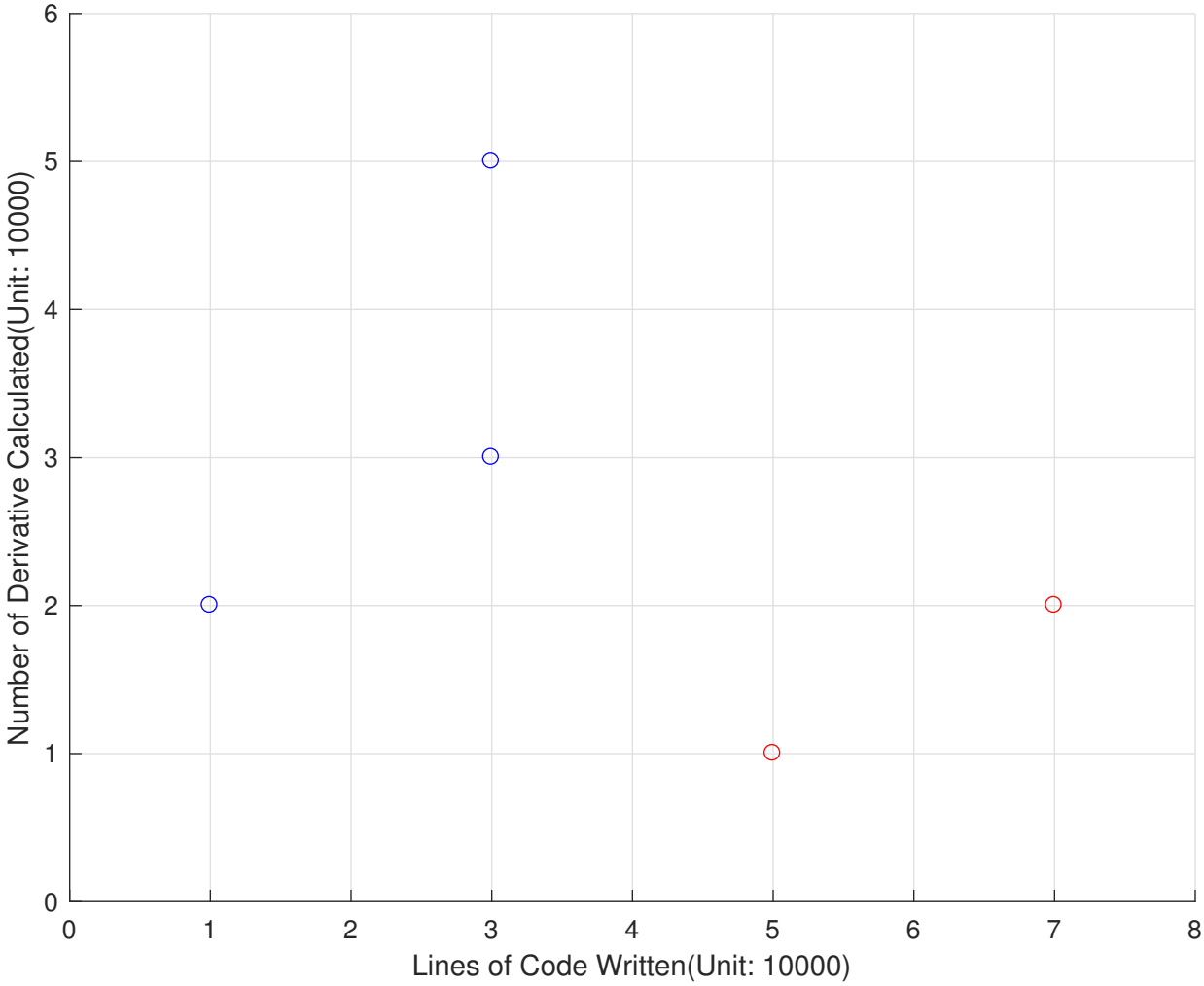

Figure 3.1: SVM Problem

3. Perceptron Algorithms The Perceptron learning algorithm can be parameterized by a learning rate α. If the current weight vector $\mathbf{w}$ classifies an instance $(\mathbf{x}, 1)$ as -1, we do $\mathbf{w}+ = \alpha\mathbf{x}$. If $(\mathbf{x}, -1)$ is classified incorrectly as 1, we do $\mathbf{w}- = \alpha\mathbf{x}$.

3.1) Short answers.

- Assume Perceptron algorithm fails to converge on a training data D. Does adding more training data help Perceptron to converge? Justify your answer.
- What problem will result from using a learning rate that's too large, and how can one detect that problem? Justify your answer.
- What problem will result from using a learning rate that's too small, and how can one detect that problem? Justify your answer.

3.2) Consider the following question about training phrase of Perceptron algorithm with training data (2 dimensional vectors) and their label.

- (0,1), 0
- (1,1), 1
- (1,0), 1

Assume the bias is 0, threshould value is 0.5, learning factor is 0.3 and the initial weight vector $\mathbf{w} = (0, 0)$. Assume samples are trained in the above listed order.

- Write down the value of $\mathbf{w}$ after training with each sample, show how you calculate $\mathbf{w}$. (first iteration only)
  (0,1),0:
  (1,1),1:
  (1,0),1:

- After first iteration, can learned vector $\mathbf{w}$ accurately label all 3 samples?