CSCI 5521: Machine Learning Fundamentals (Spring 2024)

Quiz 4 (Apr. 18, 2024)

Due on Gradescope by 2:00 pm, Apr. 19

Instructions:
- This quiz has 3 questions, 30 points, on 1 page.
- Please **write your name & ID on this cover page**.

1. (**12 points**) Given the data in the figure below, to build a decision tree for classification, what is the attribute you will use to split. Show your work.

Attribute A	Attribute B	Class Label
1	0	0
0	0	1
1	1	0
1	0	0
0	1	1

2. (**12 points**) Select all correct statement(s) about decision tree and random forest methods:

   (a) Using less decision trees to form a random forest tends to decrease overfitting.

   (b) Entropy, Gini Index and Misclassification Error are all valid options for feature selection in the decision tree method.

   (c) Random forest method is more interpretable than decision tree method.

   (d) In decision trees, the depth of the tree can be adjusted to control model complexity and avoid overfitting.

3. (**6 points**) Describe when we would want to use soft margin SVM and when we would want to use kernel SVM.