

**Midterm 1 Practice Problems***(find and work with a partner)*

1. *Algorithm Overview.* For each of the following methods, write down the type of features (continuous, discrete, or both), the type of label (continuous, binary, or multi-class). Then also write down how we defined model complexity (if applicable).

(a) Naive Bayes:

(b) K-nearest neighbors and KD-trees:

(c) Decision Trees:

(d) Random Forests:

(e) AdaBoost:

(f) Logistic Regression:

(g) (optional) Linear Regression:

2. *Evaluation Metrics.* Consider the confusion matrix below:

		Predicted class	
		1	2
True class	1	18	2
	2	10	30

(a) What is the classification error? What is the classification accuracy?

(b) Normalize the confusion matrix (each row should sum to 1).

(c) Compute the FPR, TPR (recall), and precision.

(d) Can we make a ROC curve from this confusion matrix? Why or why not?

(e) What is AUC and how do we (approximately) compute it?