FINAL EXAM TOPICS

Note: the list below is not exhaustive, and there may be some questions outside this list at the discretion of the instructor. You are responsible for all topics covered in class

The final exam will primarily focus on the following topics:

Pre-midterm topics:

- Know the decision boundary and basic concepts of each of the classifiers e.g. Decision Trees, Perceptron, Neural Net, linear SVM, etc.
- MLE technique for parameter estimation
- MAP technique using naïve Bayes
- Expected value and variance of common probability distributions, such as Bernoulli, Binomial, Uniform, and Gaussian.

Post-midterm topics:

1. Logistic Regression
   - Know the equation and rule for classification
   - Decision Boundary
   - How parameters are calculated

2. Model Evaluation
   - Understand the definition and calculation of:
     Accuracy, Precision, Recall, F-measure
   - k-fold cross-validation
   - ROC curve and how to infer classifier performance from a given chart

3. Instance Based Learning (k-Nearest Neighbor)
   - Concept of lazy evaluation
   - Decision Boundary
   - Given distance metric and training data, classify a test point for several values of k

4. Ensemble Methods
   - Training and True Errors
   - How to calculate expected value, bias, and variance of a random variable
   - Intuitive definition of bias and variance
   - Bias-Variance tradeoff and its application to various scenarios – e.g. simple vs complex models, constant vs highly changing output, etc.
   - Expected value and variance of sum, average, and other variants of N variables
   - Bootstrap, Bagging, and Random Forest classifiers
   - Boosting – concept, calculations, weight update of points, what type of questions can it solve, number of iterations needed, final decision boundary of a group of weak classifiers

5. Unsupervised Learning
    - Concept
    - K-means including calculations, K-means along with n-fold cross validation.
    - Strength and weakness of k-means
    - Hierarchical clustering using single, complete, or average link
    - Basic idea of EM algorithm (no calculations)

6. PCA
    - Just basic understanding and aim (no long calculations)

7. Bayes Network
    - Number of parameters needed for complete dependence, complete independence, and with a given Bayes Net
    - Given a Bayes net and probability tables, compute the joint probability
    - Given value of one node, how does it influence probability values of other nodes
    - d-separation for various scenarios

8. Hidden Markov Model
    - Concept of hidden and observable variables
    - Joint probability of hidden and observables
    - Forward algorithm for evaluating probability of observables
    - Viterbi algorithm for predicting the most likely hidden state at every step