STAT 6306 – Some Final Exam Topics

In no particular order, here are some important concepts (membership on this list is neither necessary nor sufficient for inclusion on the exam). The exam will be closed book/notes/computer.

- (all the topics listed on the first exam topics list)
- Be able to sketch the optimization problems behind ridge and lasso regression. Understand how to include on the sketch the least squares solution, and the tuning parameter/.
- Trees can be viewed either as a type of partition (binary splits perpendicular to the axes) or a dendrogram. Know both of these representations and how they related to each other.
- How can you create classifications out of B iterations of bagging applied to trees?
- What type of procedures are amenable to bagging? boosting?
- We haven't talked about K nearest neighbors. It is a classic, straight-forward procedure that is commonly used in practice. Read through that section of the book to get a general idea of the concept.
- What do the slack variables do for SVMs? What happens if the budget is zero? Can you sketch the decision boundary for SVMs in a simple problem were $p \leq 2$?
- Know the ingredients to Neural Networks, (including what is an activation function, hidden unit, ...). The hidden units are a type of a feature transformation.
- What is the relationship between logistic regression and LDA? How are they exactly the same? In spite of this, why don't they produce the same solution?
- What are the benefits behind using GAMs? What are the limitations?