Course STAT 656 overview including terminology and notation. General coding practice, R, split-apply-combine

- Data preprocessing, cross validation, & bootstrap. Parallelism. The Caret package in R. Measuring performance in regression & classification.

- Multiple linear regression, Logistic regression, Linear discriminant Analysis

- Penalized/regularized models
- Additional topics on penalized/regularized models such as relaxed lasso and sparse data

- Nonlinear methods including Neural Networks, Support vector machines, Splines, K-Nearest neighbors
- Decision trees

- Bagging and Random Forest

- Boosting
- A very brief introduction to text processing
- Classification with severe class imbalance