CS-4850 Machine Learning - Midterm Review Study Guide

Midterm objectives:

- Machine learning concepts
  - Optimizing a cost function
  - Gradient descent optimization
  - Overfitting, underfitting, generalization
  - Residuals and irreducible error
  - KNN
- Linear regression, multivariate analysis
  - Coefficient analysis for linear regression
  - Polynomial regression
  - RSS, R^2, MSE, RMSE
  - Solving by exact solution, normal equation, gradient descent
- Logistic regression
  - Cross-entropy loss
  - Odds, log odds, coefficient analysis for logistic regression
- Decision Trees, Bagging, Boosting
  - Gini Index, cross-entropy/information gain
  - Bagging/Bootstrap aggregation
  - Boosting
  - Ensemble classifiers, Random Forests, Gradient Boosting
- Classification model evaluation and metrics
  - Confusion matrix
  - Accuracy, precision, recall/sensitivity, specificity
  - TPR, FPR, ROC curve