Catalog Description:
An introduction to some basic concepts and methods in statistical learning with emphasis on the mathematics behind these concepts and methods.

Prerequisite:
A C- or above in Stat 4202; and a C- or above in Math 2568; or by permission of department.

Text:
An Introduction to Statistical Learning: with Applications in R, by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, published by Springer, ISBN: 9781461471370

Topic List:

1. Introduction: supervised/unsupervised learning, loss function, regression and classification
2. Model Accuracy: training and test error, bias-variance trade-off, KNN
3. Linear Regression: simple and multiple linear regression, hypothesis testing, variable selection, model fit
4. Introduction to GLM: general linear regression and link function
5. Classification: logistic regression, linear and quadratic discriminant analysis
6. Resampling: cross--validation, booststrap
7. Model Selection: variable selection, ridge regression and Lasso, dimension reduction, principal component analysis, partial least squares
8. Nonlinear Regression: polynomial regression, regression splines, smooth splines, generalized additive model
9. Tree Based Methods: decision trees, pruning, classification and regression trees, error rate, bagging, random forests, boosting
10. Neural Networks
11. Support Vector Machines: maximal margin classifier, support vector classifer, kernel functions, support vector machines
12. Introduction to Unsupervised Learning: PCA in unsupervised learning, K-means and hierarchical clusterings