Statistical Machine Learning

Description

This lecture provides an overview of the most widely used methods of statistical machine learning. We will focus on classification and regression: Linear regression, classification via logistic regression, linear discriminant analysis, SVM and kernel methods, cross validation, model selection, regularization (ridge regression and LASSO), regression and classification trees, boosting, bagging and random forest, dimensionality reduction, principal component analysis and clustering. During the labs, we will apply these methods to real data extracted from UCI website.

Learning Objectives and Outcomes

Students who pass the course should be able to:

- Use and develop linear and nonlinear models for classification and regression
- Explain how the quality of a model can evaluated by use of cross validation
- Explain the challenges with high dimensional data and have basic understanding of dimensionality reduction.
- Use dimensionality reduction, principal component analysis and clustering to visualize data and find groupings in data

Course Schedule and Contents

Session 1: Introduction - Linear Regression - Gradient descent - Régularization Lasso and Ridge
Session 2: Linear Classification - Logistic Regression - MultiClass Classification
Session 3: Support Vector Machine : Linear version
Session 4: Support Vector Machine : Kernel Method
Session 5: Decision Trees, Bagging, Boosting, Random Forest
Session 6: Dimensionality Reduction
Session 7: Unsupervised Learning
Session 8: Final Exam

Grading <Feel free to modify/adapt…)

Ass#1 Course case study:	20%
Quiz:	50%
Ass#2 Course Project:	30%

Policies

- I expect you to turn-in your reports on time to receive proper credit/grade.
- Any work submitted must be your own.
- I expect everyone to contribute equally to group assignments

- Attendance in every class is expected and class participation and discussion is strongly encouraged.
- Late work will not be accepted unless prior arrangements have been made directly with me.
- Cases will be decided on an individual basis.

Good Luck!