Spring 2018 ESC Curriculum

< Syllabus >

1) Lecture : Data Mining with Applications in R

2) Course : 3/15 - 5/31, Every Thursday, (P.M 7:00 - P.M 9:00)
(No class : 4/19(Midterm))

3) Lecturer : SeoHyeong Jeong (Economics/Applied Statistics)

4) Objective

(1) to understand statistical learning

(2) to utilize R to proceed data mining and data analysis

5) Preparations

(1) R and Rstudio

6) Reference

(1) Stanford STAT 202(http://web.stanford.edu/class/stats202/content/homework.html)

(2) An Introduction to Statistical Learning with Applications in R (Hastie.T)

(3) The Elements of Statistical Learning (Hastie.T)

7) Plan - Next page

Date	Topic & Contents
3/15(Thur)	(Topic) Exploratory Data Analysis (R) Correlation Covariance Matrix, Boxplot, Histograms, Missing Value Imputation, Bias-Variacne Trade-off
3/22(Thur)	(Topic) Linear Regression (R) Simple and Multiple Linear Regression
3/29(Thur)	(Topic) Model Selection (R) Model Selection, LOOCV, K-fold Validation
3/29(Thur)	(Topic) Logistic Regression (R) Logistic Regression, ROC Curve
4/5(Thur)	(Topic) K-Nearest Neighborhood (R) Bayes Classifier, KNN
4/12(Thur)	(Topic) Regularization - Ridge, LASSO (R) Ridge Regression, The Lasso
4/26(Thur)	(Topic) Decision Tree (R) Decision Tree
5/3(Thur)	(Topic) Bootstrap (R) Bootstrap
5/10(Thur)	(Topic) Ensemble - Random Forests, Bagging, Boosting (R) Random Forest, Bagging, Boosting (XGBoost)
2/3(Thur)	(Topic) Final Project - Prediction on IMDb score