Course Number

STA-364-01

Course Description

This course focuses on the analysis of large data sets in diverse application areas using statistical programming languages. Students will develop an understanding of the role of machine learning methods within the context of the scientific method. They will analyze real data sets using downloadable statistical programming packages, including on a course project of their own choosing. This analysis will include exploratory data analysis, visualization, and use of more sophisticated classification and predictive algorithms including nearest neighbor, nave Bayes, classification and regression trees (CART), neural networks, and others. During the course we will pay special attention to validating models using the train and test regimen, as well as through cross validation and bootstrapping. In the process of studying the machine learning methods themselves, students will develop an ability to manipulate big data to accomplish the previous objectives. This includes downloading, merging, appending and reshaping data, and creating new variables. Successful completion of this course would be advantageous for those considering graduate study or employment in the areas of statistics, data science, machine learning, computer science, econometrics, or related disciplines. Course Types: JDQR

Academic Term

22/SP

Instructor

Hoerl, Roger

Location & Meeting Time

Wold Center-128+ M/W/F 11:45AM-12:50PM LEC

Credits

1.00

Capacity

18

Total Students

12

Common Curriculum

JDQR

Academic Department

Mathematics

Field Of Study

Statistics (STA)