



DATA MINING and ANALYTICS

Data Mining and Analytics

Data mining is simply the process of turning raw data into useful information. It involves analyzing data from different perspectives and summarizing them into useful information. The general goal of data mining is to extract the information and transform it into an understandable structure for further use.

Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. It involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

For **tall** data, the course covers a wide range of techniques for supervised learning from linear regression through various classes of more flexible models to fully nonparametric regression models, both for the regression problem and for classification. For **wide** data discussion will center on various forms of regularized model fitting, including many of those used routinely in the analysis of genomic data.

The course includes many examples and case studies, and participants should leave the course well-armed to tackle real problems with realistic tools.

Who should Attend?

Engineers, data analysts, scientists, managers and others who would like to understand the basic principles underlying statistical learning and data mining.