Mathematics, Statistics, and Computer Science   **@ UIC**

Statistics and Data Science Seminar

JMP and the Predictive Modeling Workflow

Kevin Potcner (JMP Statistical Discovery Software)

Abstract: As the size and sources of data becomes more available in today's business environments, data analysts are beginning to add more sophisticated predictive statistical modeling techniques to their analysis toolkit.

A typical real-world predictive modeling workflow includes data cleaning and exploration, model fitting, model validation, model comparison, final model selection and deployment of the final predictive model.

In this presentation, a statistical scientist from JMP will illustrate the predictive modeling workflow by analyzing a real dataset. After data preparation and initial exploration, we will create a number of predictive models such as Multiple Linear Regression, Regression tree, Neural Net, and K-Nearest Neighbors.

We will evaluate each model and select the best model using the Prediction Profiler and JMP's Model Comparison tool.

Code will be automatically created in a variety of programming languages (e.g., SAS, SQL, Python, et al.) in order to implement that model in a production environment.

Wednesday, September 26 at 4:00 PM in 636 SEO