Supervised Machine Learning

Preparing Data & Deploying Analytic Models for Classification & Prediction

- Decision Trees in Machine Learning
- 1984: Leo Breiman and CART
    - Build Model
    - Decisions Driven by Data and Scores
    - Strategic Considerations
    - Other Uses of Trees
    - Other Advantages of Trees
- How Do Decision Trees Work?
    - Build Model
    - Selecting Records
    - A Decision Tree Algorithm
- CART Classification Trees
    - Gini Coefficient
    - Pruning
    - Balance: Default Cut
    - Forcing Balance
    - Forcing High Purity
    - Surrogates
- CART Regression Trees
    - Regression Tree
    - Overfitting
- Other Tree Algorithms
    - Understanding CHAID
    - Scale
    - Nominal
    - Four Algorithms Compared
    - The Roshomon Effect
- Tree vs. Ruleset
    - Build Model
    - Initial Tree Output
    - Tree Converted Into Rules
    - PMML
- Confidence vs. Propensity
    - Confidence Scores
    - Propensity Scores
- What are Ensembles?
    - Heterogeneous Ensemble

    - Predicting MPG with an Ensemble
    - Predicting MPG with Regression
    - Predicting MPG with a Neural Network
    - Predicting MPG with a Regression Tree
    - A Simple Ensemble
    - Heterogeneous Classification Ensemble
    - Netflix Prize: A Cautionary Tale
    - 8th Law of Data Mining
    - Overfitting
    - Error Decomposition
    - Variance
    - Low Variance
    - Bias
    - Lower Bias
    - Increased Variance
    - Pedro Domingos on Bias & Variance
    - Why Tradeoff?
    - Signal to Noise Ratio
- Famous Ensemble Approaches
    - Heterogeneous Ensembles
    - Bagging
    - What is “Bootstrapping”?
    - From Bagging to Random Forest
    - Stacking
    - Deep Stacking
    - Is a Multi-Layer Perceptron Similar to Stacking?
    - Boosting
    - Boosting is About Changing Weights
    - Wrong Answers Get Increasing Weight
    - Bumping
- Wrap-Up

