PREDICTIVE ANALYSIS

1: Preparation

What is supervised learning

Defining Business Goals for Predictive Analytics

Effective Data Preparation

Data partitioning and holdout samples

Choosing variables (features)

Handling missing data

Visualization and exploration

2: Classification and Prediction

Assessing classification models

Confusion matrix

Misclassification costs

Lift

Assessing prediction models

Common metrics

K-Nearest-Neighbors (KNN)

Measuring distance

Choosing k

Generating classifications and predictions

3: Bayesian Classifiers; CART

Full Bayes classifier

Naive Bayes classifier

Classification and Regression Trees (CART)

Growing the tree

Avoiding overfit - pruning

Using trees for classifications and predictions

4: Ensembles

Combine multiple algorithms

Improve results