

INTRODUCTION

What Is Predictive Analytics?

- Introduction & Core Concepts

Predictive Analytics: Process Models

- Predictive Analytics is a Process
- Modeling Practice Framework™ (“MPF”) Motivations Over CRISP-DM
- MPF is More Linear than CRISP-DM

Overview: Modeling Practice Framework™

- Assess Phase
- Plan Phase
- Prepare Phase
- Model Phase
- Validate Phase
- Deploy Phase
- Monitor Phase

Assess Phase

Enterprise Assessment

- Defined
- Rush to Modeling
- No Project Implementation
- Data Mining ROI Survey
- Right-Sizing Initial Projects
- A Realistic CRISP-DM Timeline

Problem Definition

- Access Phase
- Assemble Team
- Initial Meetings: Candidate Projects
- Translate Into a Predictive Analytics Problem
- Decisions Driven by Data and Scores
- Example
- Prioritize Viable Projects
- Case Study: Windshield Repair
- Initial Cost Benefit Analysis
- What is Your Intervention Strategy?
- Assess Deployment

- Access Data

Documentation

- Initial Report of Practice Readiness

Plan Phase

Team: Resources

- Managing Data Scientist
- Where Do You Find Them?
- Data Science “Unicorn” vs. Teams

Tools: Resources

- Sampling of Open Source and Commercial Data Mining Software Products
- Gartner
- Gartner BI
- Rise of R Usage
- Top Data Analytics / Data Science Tools

The Data

- Viability
- Select Data

Documentation

- Documentation Tasks
- Data Reports
- Create Implementation Strategy

Prepare Phase

Prepare Phase Overview

- Project Roles
- Create Analytic Sandbox
- Integrate
- Aggregations: Numerous Options
- Restructure: Numerous Options

Feature Extraction

- Data Preparation Law
- Two Winning Approaches
- Handcrafted Features: Car Color

Prepare Data Subtasks: Modeling Practice Framework™

- Construct
- Ratios and Deltas
- Date Math
- Extract Subtask

- Strings

Modeling

What are AI and Machine Learning?

- What is AI?
- Traditional Machine Learning
- Supervised Machine learning
- Two Kinds of Supervised learning
- 2015 Algorithm Usage
- Some Common “ML Platforms”

Three Contrasting ML Techniques

- What is Unsupervised Learning?
- Computers “Teaching Themselves”?
- Google Brain
- All Roads Lead to Binary Classification

What is Unsupervised Learning?

- Labeled and Unlabeled
- Examples

Model Phase

- MPF™ Model Phase Objectives
- Tasks Related to Model Selection
- Converging Toward a Model
- One Task – Many Algorithms
- Algorithms
- Model Phase: Generate Test Design

Strategic Considerations

- Value Law
- Bias / Variance Trade Off
- Black Box or Not?
- Balancing
- Roshoman Effect
- Pros and Cons of Statistical Models
- 4th Law of Data Mining

Overfitting

- Training and Testing
- Select Modeling Technique

Black Box Models and Deep Learning

- Strategic Considerations
- Deep Learning’s Grandfather

- Multi-Layer Perceptron
- Deep Learning

Balancing

- Selecting Records

Ensembles and the "Roshomon Effect"

- The Problem
- The Roshomon Effect
- Two Models

Statistical Models

- Statistics vs. Predictive Analytics

Model Phase: Build Models

- Modeling Techniques
- Supervised Learning
- Unsupervised Learning
- Algorithms

Model Phase: Decision Trees

- Decision Trees: Recursive Partitioning
- Decision Trees: Rules-Based

Common Unsupervised Learning Techniques

- Two Main Types of Unsupervised
 - Clustering
 - Association

Hierarchical Clustering

K-means Clustering

Silhouette

K-Nearest Neighbors (KNN)

Model Phase: Evaluate Candidate Model Results

- Assess Models
- Model Performance

Document Models

Validate, Deploy, and Monitor

Validate Phase

- Top Attributes
- Evaluate Candidate Models
- Survey Reveals Flawed Strategy

- 8th Law of Data Mining
- Dress Rehearsal
- Validate Phase Report
- Project Review

Deploy Phase

- Deployment Strategies
- Five Deployment “Locations”
- PMML

Monitor Phase

- Develop Modeling & Maintenance Plan
- Three Levels of Maintenance
 - Refit
 - Remodel
 - Rebuild

RESOURCES

- Analytic Glossary
- Recommended Books
- LinkedIn Groups
- Data Repositories
- Predictive Analytics Across Social Media
- Webinars, Courses, Conferences