Combining Predictive Analytics with Business Rules to make better decisions

Swami Balasubramanian, IBM
Richard Cletus, TAX/ITS – State of NY
Setting the Context

Background

• State of New York Department of Taxation is responsible for all state tax matters
• Charged with Mission to efficiently collect tax revenues in support of State services and programs while acting with integrity and fairness in the administration of the tax laws of New York State

Problem

• Historically the department has been reactive in responding to non-compliance which is expensive and time consuming
• Department wanted to move to a model that enables early detection of non-compliance and pro-active action

Approach & Presentation

• In this presentation, we will discuss an approach we took to solve this problem using predictive decision-making
Agenda

• Business Rules
• Predictive Analytics
• Combining Business Rules with Predictive Analytics
• Case Study – State of New York Taxation
• Lessons Learned
Agenda

• Business Rules
• Predictive Analytics
• Combining Business Rules with Predictive Analytics
• Case Study – State of New York Taxation
• Lessons Learned
Operational Decisions

- **Characteristics**
  - Repeatable
  - Consistent
  - Lower value
  - Often relate to a single entity (customer, transaction, etc.)
From Policy to Rule

A business rule is a statement of business logic that

• Business users can author and understand
• Applications can invoke for execution

From a business perspective

• A precise statement that describes, constrains, or controls some aspect of the business

From an IT perspective

• Business rules are a package of executable business policy statements that can be invoked from an application
From Policy to Rule

Rules formalize business policy as ‘if-then’ statements

Business policy

When customers spend more than $1500 in a single transaction, they should be upgraded

Formal rule

If the customer’s category is Gold and the value of the customer’s shopping cart is more than $1500
Then change the customer’s category to Platinum
Empowering the Business User

• A Business Rules Management System (BRMS) should offer the business user
  – Ability to manage the rules within a rule authoring environment
  – Ability to unit and regression test rules
  – Ability to build and deploy rule updates without dependence on the enterprise IT team
  – Ability to govern the rules
Agenda

• Business Rules

• Predictive Analytics

• Combining Business Rules with Predictive Analytics

• Case Study – State of New York Taxation

• Lessons Learned
Data Mining

Current state of many organizations

- Drowning in data
- Starving for information

Exponential data growth culprits

- Social
- IOT
- Enterprise
- Unstructured
Data Mining

• Need for automated data analysis driven by
  – Rapid growth of data: from terabytes to exabytes
  – Rapid increase in percentage of uncertain data
    • unstructured data from IOTs, social media, etc.
Data Mining

Common problem

- More the data; more difficult and time-consuming it is to analyze and draw meaning from the data
- A gold mine of data remains unexplored

Solution

- Employ clear business procedures, and powerful analytical tools
- Explore massive volumes of data
- Extract and present hidden patterns and relationships in past behavior – the heart of business intelligence

This is the essence of Data Mining
Past Behavior to Future Predictions

Past Behavior
- Patterns
- Relationships

Models
- Rules
- Formulas
- Equations

Future Predictions
- Advanced Analytics
- Decision Optimization

- Uses patterns in historical and transactional data to predict future events
  - Is a broad array of techniques that combine data mining and statistics
  - Analyzes current and past facts
  - Makes predictions about future events
Predictive Analytics

Is a business process and set of technologies that combines

- Advanced Analytics
  - Tools and techniques to analyze past and present events; and predict future outcomes
- Decision Optimization
  - Selects the appropriate actions so they are incorporated in the business processes

Holds the key for

- Optimized performance
- Informed decisions
- Actionable insight
- Trusted information

Help answer the questions

- What is happening
- Why is it happening
- What is likely to happen
- How should we plan the future
# Analytical Models

## Classification
- Use the values of one or more **input** fields to predict the value of one or more **output** (target) fields
- Ex – decision trees, neural networks, regression

## Association
- Finds patterns in the data when one or more entities are associated with other entities
- Ex – apriori models, carma models

## Clustering
- Divide the data in buckets based on similar patterns
- Ex – K-means clustering, Kohonen networks
Agenda

• Business Rules
• Predictive Analytics
• Combining Business Rules with Predictive Analytics
• Case Study – State of New York Taxation
• Lessons Learned
Current challenges in decision making

- Evolving decision landscape
- Disparate and incomplete data
- Departmental silos and ‘Digital IT’
- Rearview thinking
Current challenges in decision making

- Data expanding rapidly
  - Volume
  - Velocity
  - Variety
- Spans multiple formats
  - Traditional enterprise
  - Unstructured
  - Machine-generated
- Decision making getting difficult
- Data needs to be managed so it is meaningful in decision-making
  - Sort through the data
  - Understand and connect key pieces of information
  - Drive smarter decision making that equally responds to and anticipates events
Current challenges in decision making

- Traditional systems rely on broad segmentation of historical data that generalize, not personalize
  - Often expressed in structured formats enabling fundamental analysis
- Unstructured data contains rich insights often overlooked in fundamental analytics
  - Such data provide critical insights into motivations behind various behaviors
- Such attitudinal data key to differentiation
  - Understanding the underlying reasons for a specific behavior unlocks new insights allowing for prediction of future behaviors
- However, if the data is incomplete (that is, limited to historical instances only), the ability to make high-quality decisions throughout the enterprise becomes severely limited
Current challenges in decision making

- Organizations develop natural silos along Lines of business
- Lack of data sharing
  - Missed opportunities to realize synergistic benefits
- Multiple expressions of the same data
  - Disparate systems articulate critical data in different ways
  - Perpetuates silos
  - Inhibit cross-department intelligence
- Data across sources and systems needs to be integrated
  - Common vocabulary
  - Common representation
- Traditional centralized IT diverging into two very different camps
  - Enterprise IT
    - Manage business critical systems
    - Maintain service level agreements
    - Provide data integrity and secure governance
  - Digital (Lines of business) IT
    - Rapidly create new business applications
    - Integrate new sources of data – social, IOT, etc.
    - Current explosive growth in data often attributable to Digital IT
- Data management needs to factor in the role played by Digital IT
Current challenges in decision making

- Traditional decision making grounded on historical data
  - Expected future outcomes based on what has happened in the past
- This approach is becoming increasingly unreliable
  - Conditions and variables of a scenario increasingly volatile
  - People, economies, environment, technologies that impact every decision point increasingly dynamic and connected
- Historical data remains critical in decision-making when
  - Combined with current data
  - Continuously refreshed with new information
- Data, both historic and current, needs to be combined in decision-making to
  - React – Respond to past events
  - Anticipate – Predict future events
## Evolution of decision making

<table>
<thead>
<tr>
<th>Decisions from ‘Intuition’</th>
<th>Automated decision making</th>
<th>Predictive decision making</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Based on anecdotal experience, or hunches</td>
<td>• Based on business knowledge, policies and practices embodied in business rules</td>
<td>• Based on analysis of historic patterns (data) and current conditions (data)</td>
</tr>
<tr>
<td>• Based on instinct</td>
<td>• Objective, consistent and reliable</td>
<td>• Leverage all available data</td>
</tr>
<tr>
<td>• Often subjective and inconsistent</td>
<td>• Static rules lack adaptive capabilities and quickly obsolesce</td>
<td>• Continuously adapt to new information; flexible in ever changing volatile environments</td>
</tr>
</tbody>
</table>
In Summary

Make insightful decisions

- ‘what your data knows’ + ‘what your organization knows’ = insightful decisions

Anticipate and Respond

- Analytical models based on historical data predict future behavior
- Decisions made with the advantage of ‘hindsight’

Apply rules to direct Big Data processing

- Use decision services to select, classify and filter information for Big Data processing

Act with situational awareness

- Combine analytic insight with rules to decide ‘when to act’
Notional View of Predictive Decision-Making

Capture → Predict → Act

Feedback

Association
Clustering
Classification

Predictive Models

Scoring

Actionable Insights

Decision Services

Enterprise Data
- Data at rest
- Data in motion

Unstructured Data
- Social data, IOT data

Current data

Enterprise Data

Data in motion

Enterprise Data

Data at rest

Unstructured Data

Social data, IOT data

Current data

Predictive Models

Association
Clustering
Classification

Actionable Insights

Decision Services

Capture → Predict → Act

Feedback
Traditional vs. Predictive decision making

Traditional (automated) decision-making

- If the cart total is greater than $1000 then Offer a discount of 10%

Predictive decision-making

- If the cart total is greater than $1000 and the propensity of the customer to complete checkout is greater than 0.75 and the likelihood of recurring purchases is greater than 0.7 and the likelihood of up-sale is greater than 0.5 then
  - Offer a discount of 15%
  - Offer a coupon for recurring purchases
  - Offer a coupon for future purchases
Agenda

• Business Rules
• Predictive Analytics
• Combining Business Rules with Predictive Analytics
• Case Study – State of New York Taxation
• Lessons Learned
Our Mission:
To efficiently collect tax revenues in support of State services and programs while acting with integrity and fairness in the administration of the tax laws of New York State.

The four pillars supporting our mission are:
- Customer Service
- Fiduciary Responsibility
- Loss Prevention
- Enforcement

Our Goal:
Current and Evolving Challenges

Vision

• Make it easier to do business with NYS
• Less complex, more efficient
• Improve usability for DTF staff
• Use data that is gathered during the process to improve analysis and decision making for managing accounts

Success Criteria

• Full in-stream processing for fast response to taxpayer and reduced manual work
• Accelerate intervention with tax payer – both internal and external
• Know the vendor
• Automated evaluation of data to determine follow-up actions
• Increase in voluntary compliance
Vision around Data Analytics

**Business Definition:**
Creating data driven business solutions by building a customizable user enabled process.

**Benefits:**
- Big Opportunities, Identification and creation of new services
- Quicker capability to respond to trends
- Timely decision-making based on data
- Better enablement of key strategic initiatives
- Enhance relationships with customers and business partners for increased voluntary compliance
- Better sense of our risk and ability to react to changes in the economic environment
- Reusability and agility leading to Cost effectiveness
Problem Summary

**Reactive**
- Wait for entities to meet obligations before checking for compliance
- Process expensive and time consuming
- Lengthy Audit and Appeals processes

**Proactive**
- Early detection of non-compliance (Risk)
- Pro-active action
Solution Approach – Overview

- **Capture**
  - Multiple predictive models based on identified ‘Risk’ scenarios
  - Historic tax event data used to build models
    - Models refreshed on a schedule
  - At runtime, input ‘Tax Event’ data drives model selection

- **Predict**
  - Scoring model corresponding to each predictive model generates ‘Risk’ predictors for input ‘Tax Event’

- **Act**
  - Rules determine action plan based input ‘Tax Event’ and ‘Risk’ predictors
  - Rules implement specific legal and policy restrictions to the action plan

---

**Historic Tax Data**

**Current Data**
(Tax Event – Registration, Filing, Action Feedback, etc)

**Predictive Models**
Multiple models based on ‘Risk’ scenarios

**Scoring Models**
Multiple Scoring models

**Decision Services**
Determine Actions, Apply legal restrictions, Apply policy restrictions

**Corrective Actions / Action Plan**
Solution Approach – Design Choices

Option #1
- Client invokes the predictive analytics component that internally calls decision services after scoring
- Tight coupling, no cohesion, no separation of duties – not a good choice

Option #2
- Client invokes decision services that internally calls the predictive analytics component for scoring
- Tight coupling, no cohesion, no separation of duties – not a good choice

Option #3
- Client invokes an orchestrator
- Orchestrator calls the predictive analytics component for scoring, next invokes decision services with scoring data
- Analytics & Decision Services loosely coupled, clear separate of duties

SELECTED PATTERN
Solution Approach – Need for data augmentation

Multiple predictive models developed corresponding to different ‘Risk’ scenarios
- Corresponding scoring model

Two Challenges
- At runtime when processing a Tax Event, policies exist that determine the ‘Risk’ scenario based on event data
- Additional ‘system of record’ data needed for scoring based on the identified scenario (data augmentation)

Approach Taken
- Use business rules for policy-based scenario determination
- Use business rules to drive the data augmentation

Rationale
- This is an intra-process technical decision (not a traditional decision point) BUT
- We wanted to empower business users to manage the solution end-to-end, and this allowed them to be fully empowered
  - Manage Predictive and Scoring models
  - Manage Business Rules
Implementation – Products

- IBM Operational Decision Manager
- IBM SPSS
- IBM WebSphere Process Server (BPM)

More Information –
### Implementation – Flow

<table>
<thead>
<tr>
<th>IBM WPS</th>
<th>IBM ODM</th>
<th>System of Record</th>
<th>IBM SPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read Tax Event data from event queue</td>
<td>Invoke ODM</td>
<td>Invoke SOR service for augmentation</td>
<td>Invoke SPSS with augmented Tax Event data</td>
</tr>
<tr>
<td>Determine ‘Risk’ scenario based on event data</td>
<td>Select data augmentation</td>
<td>Perform data augmentation</td>
<td>Invoke ODM for action plan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Select Actions / Action Plan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Apply legal and policy restrictions on plan</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Perform scoring</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Return action plan to end-user application</td>
</tr>
</tbody>
</table>
Where we are today ...

• Development and testing of solution completed
• Solution deployed to production in Sept
• Predictive modeling and rule authoring is an on-going process
  – The two elements go together
  – An initial set of scenarios, and corresponding business rules developed
  – Business user enablement underway
Lessons Learned

Developing predictive models is a complex undertaking

• Determining correct predictors, scoring weights, etc. requires a good sample set of initial data and time set aside for tuning
• Start with simple predictive models and progress to more complex models as expertise is built

Combining predictive scoring with business rules allows the creation of effective rules

• Personalize vs. Generalize
• Tailor-able to individual classes of tax payers

Involving the business from the start critical to success
THANK YOU