Actuarial Modernization: Expedited Approaches for Reserving, Pricing, and Financial Reporting

Actuarial analysis in the digital age
NAIC Predictive Analytics Book Club

December 4, 2018
Agenda

Topic

Reserve Modernization – Refresh

Industry Overview – Current Pilots and Uses

Applications of Reserve Modernization

Case Studies

Results Interpretation & Visualization
The components are interconnected and build on each other

<table>
<thead>
<tr>
<th>Robotic process automation</th>
<th>Cognitive Analysis</th>
<th>Results Presentation</th>
<th>Enterprise Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automation of repetitive tasks</td>
<td>Machine learning techniques applied to claims valuation</td>
<td>New techniques to tailor and present results</td>
<td>Operating model to translate new insights into action</td>
</tr>
<tr>
<td>Use of “bots” – a kind of super macro that operates across systems</td>
<td>Results allocated at a granular claim level</td>
<td>Enhanced ad hoc analytics</td>
<td>Mobilization across core departments – pricing, underwriting, claims, finance</td>
</tr>
<tr>
<td>Review existing process flows, identify automation points</td>
<td>Leverages new statistical software</td>
<td>Applies new visualization tools to the granular data</td>
<td>Identifies processes, structure, roles, and governance to communicate, interpret, and respond to insights/trends</td>
</tr>
<tr>
<td>Develop and test ‘bot’ macros</td>
<td>Uses structured and unstructured data, including individual claim characteristics</td>
<td>Combination of standard, tailored, and ad hoc reports</td>
<td></td>
</tr>
<tr>
<td>Shorter cycle times and faster close process</td>
<td>Faster identification of trends</td>
<td>Better, user-friendly reports with more granular insights</td>
<td>Common view of issues</td>
</tr>
<tr>
<td>Less resources needed – deploy to other priorities or eliminate to save costs</td>
<td>Results at granular claim level allows for deeper root cause analysis</td>
<td>Stronger engagement by business-side consumer of the information</td>
<td>Coordinated cross-unit action</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Effective, timely response to issues</td>
</tr>
</tbody>
</table>

© 2017 KPMG LLP, a Delaware limited liability partnership and the U.S. member firm of the KPMG network of independent member firms affiliated with KPMG International Cooperative (“KPMG International”), a Swiss entity. All rights reserved.
Agenda

**Topic**

- Reserve Modernization – Refresh

**Industry Overview – Current Pilots and Uses**

**Applications of Reserve Modernization**

**Case Studies**

- Results Interpretation & Visualization
State of the Industry

Commentary

- Market Leading - One global P&C insurer is currently rolling out cognitive reserving across multiple segments
  - By perhaps 1-2 years ahead of other leading adopters

- Several more sophisticated companies are piloting claims level reserving approaches
  - Often technical exercises on troubled segments

- Some have taken the approach that these tools should be used to enhance traditional methods

- Many companies are using Cognitive approaches outside the reserving function
Agenda

Topic

Reserve Modernization – Refresh

Industry Overview – Current Pilots and Uses

Applications of Reserve Modernization

Case Studies

Results Interpretation & Visualization
A Range of Applications are being Considered

- Traditional Methods
- Pricing
- Financial Reporting
- Distribution Channel Evaluation
- Target Markets
- Claims Handling Enhancement
- IFRS 17
Squaring the triangle

10 years

AY 2009 - 2018

Predict the unpaid losses to calculate ultimate losses and compare to actual ultimates
Reserve Modernization - Applications

Pricing

Application of Cognitive Analytics allows for more granular results that can yield significant results in pricing as well as reserving

- Granular evaluation of loss experience can be applied across many company functions
- The granularity of calculations will help close the feedback loop between core actuarial functions
- Numerous applications in areas that are now siloed such as:
  - Relativities
  - Risk Classification
  - Establishment of tiers
  - Expense loading and allocation
The results of Cognitive tools have numerous applications to financial reporting and planning as well.

- Cognitive methods provide detailed future cash flows
- Drillable profitability data by profit center and book of business
- More accurate capital allocation
- Partial automation of financial reporting packets
Reserve Modernization - Applications

Underwriter & Broker Evaluation

Examine the performance of distribution channels in an intuitive and systematic manner

- Evaluate performance on a loss cost basis of longer periods of time
- Consider variation of results
- Layer in other key evaluation metrics such as retention and production
Target Market Evaluation

Evaluate profitability along numerous metrics to identify areas to grow and segments to avoid

- Target segments
- Target regions
- Target business units
- Target distribution channels
Cognitive methods have the potential to help insurers with many aspects of claims handling

- Setting more accurate case reserves
- Reserve strength & claim closure rates
- Claims triage
- Automation potential with RPA
- Reinsurance recoveries
Claim level IBNR and projected cash flows provide a solution to related issues under IFRS 17:
• Cash flows at a granular level may be required for discounting and risk margin calculations
• Onerous contract testing can be required in various segments
• For segments under the General Measurement Model, establishment of the CSM can require evaluation of various segmentations
• The flexibility of experience at a granular level can provide critical input into IFRS calculations

Key components / building blocks

- Fulfilment cash flows (FCF)
  Risk-adjusted present value of future cash flows – e.g., premiums, claims

- CSM
  Represents unearned profit – results in no gain on initial recognition

Diagram:

1. Future cash flows
   - Inflows
   - Outflows

2. Discounting

3. Risk adjustment

4. CSM
Agenda

Topic

 Reserve Modernization – Refresh

 Industry Overview – Current Pilots and Uses

 Applications of Reserve Modernization

 Case Studies

 Results Interpretation & Visualization
Case studies

Use Case 1 – U.S. Life Operation of Global Multi-Line Insurer – Cognitive Reserving Solution using a Random Forest Model to predict state changes impacting universal life cash flows

Use Case 2 – Multiple Cognitive Reserving Solutions leveraging Gradient Boosting and Random Forest Models to proactively identify deterioration of problematic business segments:

- Client A - U.S. Operation of Global P&C Insurer – Commercial Auto Reserving Model
- Client B – E&S Operation of a Top 5 US Insurer – Binding Casualty Policy Profitability Model
Prediction of state change

— Existing method for predicting state change was performing poorly
  - Cognitive analysis used to predict the probability of state changes
  - Response variable is probability of state change

— Possible policyholder states:
  - Stable
  - Near lapse
  - Lapsed with payment plan
  - No payment expected
Case 1

Model build considerations

— Parameter selection based on one way analysis

— Hyper parameter selection based on a grid search

— Data storage – simplified through K mean clustering analysis
Case 1

Model implementation

— Results of random forest model fed into markov chain matrix

— Markov chain monte carlo method applied recursively to obtain 80 years of cash flows

— Results summarized in 10 buckets determined by K means analysis
Generalized machine learning framework for cognitive reserving

Model Overview

Total IBNR Estimate

Known Claims Models (IBNER)

Known Claims Models

Open Claim Propensity Models

Pure IBNR Models

Future Reported Claim Models

Severity on Future Reported Claims Models

Claim coverage level

Portfolio level

Case 2 – Client A
### Case 2 – Client A

#### IBNER Approach

<table>
<thead>
<tr>
<th>Period</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4(F)</th>
<th>5(F)</th>
<th>6(F)</th>
<th>7(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Incremental Paid Losses</td>
<td>2500</td>
<td>0</td>
<td>3000</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Known Claims Model Estimate</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3000</td>
<td>4000</td>
<td>7000</td>
<td>7500</td>
</tr>
<tr>
<td>Open Claim Propensity Estimate</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>25%</td>
<td>21%</td>
<td>18%</td>
<td>15%</td>
</tr>
<tr>
<td>Conditional Probability of Open Estimate(^1)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>100%</td>
<td>84%</td>
<td>72%</td>
<td>60%</td>
</tr>
<tr>
<td>Estimated IBNER(^2)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3000</td>
<td>3360</td>
<td>5040</td>
<td>4500</td>
</tr>
</tbody>
</table>

Calculation is performed at low level of granularity (e.g. Claim) – leveraging granular data assets

---

\(^1\)Conditional probability of claim open at the beginning of each future period given that the claim is open at the beginning of period 4. (e.g. Conditional Probability of Open for Period 5 = 0.21 / 0.25 = 0.84)

\(^2\)Estimated IBNER = Known Claims Model Estimate * Conditional Probability of Open

---

Example – not derived from any company sources
### Case 2 – Client A

**Using Generalized Approach for Pure IBNR Estimation**

<table>
<thead>
<tr>
<th>Period</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4(F)</th>
<th>5(F)</th>
<th>6(F)</th>
<th>7(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Newly Reported Claims</td>
<td>1000</td>
<td>300</td>
<td>100</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Future Reported Claims Model Estimate</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>30</td>
<td>20</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Severity on Future Reported Claims Model Estimate</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>35,000</td>
<td>38,000</td>
<td>42,000</td>
<td>47,000</td>
</tr>
<tr>
<td>Estimated Pure IBNR</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1.05M</td>
<td>760k</td>
<td>504k</td>
<td>235k</td>
</tr>
</tbody>
</table>

Total IBNR = IBNER + Pure IBNR

*Example – not derived from any company sources*
Case 2 – Client B

Detailed Plan for Profitability Analysis

As currently envisioned, we will create a Known Claims Model to estimate IBNER at a granular level and conduct an analysis of Frequency and Severity to estimate pure IBNR:

**IBNER**

- Known Claims Model
  - Remove records for all claims open as of 2Q 2018
  - Manipulate claim-level data to aggregate multiple claimants into single claims
  - Append Final Incurred at Close Back to Earlier Claims
  - Use Evaluation Date Splits and/or Random Sampling to Determine Train / Test / Validate Data
  - Isolate records for all claims open as of 2Q 2018
  - Apply Gradient Boosting Model to predict ultimate loss (and tail adjustment)
  - Validate against traditional approaches

**Pure IBNR**

- Frequency Model
  - Aggregate exposure data and historical claims reporting lags
  - Random Forest Model
  - Predict newly reported claim in each future period (separately for BI / PD)
  - Estimate Pure IBNR Loss
  - Allocate to policy level using Earned Premium

- Severity Model
  - Aggregate results from Known Claims Model to calculate ultimate severity for each known claim
  - Random Forest Model fit with ultimate loss from known claims model as target variable
  - Organize predicted ultimate severities by accident period and reporting lag
  - Aggregate from Known Claims Model to calculate ultimate severity for each known claim
Case 2 – Other Considerations

Test and Validation Approaches

Validation approach used for selections of

- Model methods
- Hyper parameters
- Predictive variables

Test approach used to evaluate the model performance

- Do the selected methods with hyper parameters and variables generalize well in different time periods?

Two approaches test

- Actual vs predicted claim emergence / ultimate at close
  - Using data test set
- Model predicted IBNER / Pure IBNR vs. traditional methods indications

Approach #1: Longitudinal Holdout

Consider Entire Population of Closed Claims

Apply Random Sampling to divide into 3 buckets

Train Dataset 70%

Validate Dataset 20%

Test Dataset 10%

Limiting the train dataset to claims closed by 1Q2016 (18,000) reduces total closed claims available by 40% versus 2Q2018 (30,000)

Approach #2: Based on Expected Distribution of Closed Claims

Consider Entire Population of Closed Claims

Apply Random Sampling to divide into 3 buckets

Train Dataset 70%

Validate Dataset 20%

Test Dataset 10%
### Case 2

Granular Results = Flexible Insights

Levels with largest paid deviations from expected

Commercial auto liability

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>Actual – expected paid losses</th>
<th>Primary driver</th>
<th>Secondary driver</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>Texas</td>
<td>$32,000,000</td>
<td>Closed claim count (Higher than expected)</td>
<td>Paid severity (Higher than expected)</td>
</tr>
<tr>
<td>Segment</td>
<td>Construction large account</td>
<td>$20,000,000</td>
<td>Paid severity (Higher than expected)</td>
<td>Closed claim count (Higher than expected)</td>
</tr>
<tr>
<td>Vehicle weight</td>
<td>Heavy weight truck</td>
<td>- $19,000,000</td>
<td>Paid severity (Lower than expected)</td>
<td>N/a</td>
</tr>
<tr>
<td>Unbundled indicator</td>
<td>TPA Handled</td>
<td>- $55,000,000</td>
<td>Newly reported claim count (lower than expected)</td>
<td>Paid severity (Lower than expected)</td>
</tr>
</tbody>
</table>

**Illustrative**

This Cognitive Actual vs. Expected tool will also provide fast insights about trends which benefit risk selection, pricing, claim handling, etc.

Is TPA data properly in our systems?
### Profitable Classes

**"A" Classes**

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Earned Premium</th>
<th>Ultimate Loss Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>37,006,340</td>
<td>34.3%</td>
</tr>
<tr>
<td>2012</td>
<td>36,407,096</td>
<td>32.0%</td>
</tr>
<tr>
<td>2013</td>
<td>39,034,261</td>
<td>39.2%</td>
</tr>
<tr>
<td>2014</td>
<td>40,261,749</td>
<td>42.3%</td>
</tr>
<tr>
<td>2015</td>
<td>41,718,413</td>
<td>26.9%</td>
</tr>
<tr>
<td>2016</td>
<td>42,087,124</td>
<td>22.8%</td>
</tr>
<tr>
<td>2017</td>
<td>42,267,347</td>
<td>27.0%</td>
</tr>
</tbody>
</table>

**Total** 278,782,331 30.7%

**"B" Classes**

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Earned Premium</th>
<th>Ultimate Loss Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>217,759,676</td>
<td>48.9%</td>
</tr>
<tr>
<td>2012</td>
<td>226,605,359</td>
<td>48.0%</td>
</tr>
<tr>
<td>2013</td>
<td>246,685,981</td>
<td>48.3%</td>
</tr>
<tr>
<td>2014</td>
<td>276,321,007</td>
<td>56.4%</td>
</tr>
<tr>
<td>2015</td>
<td>293,284,212</td>
<td>54.2%</td>
</tr>
<tr>
<td>2016</td>
<td>302,186,749</td>
<td>56.9%</td>
</tr>
<tr>
<td>2017</td>
<td>308,008,582</td>
<td>58.6%</td>
</tr>
</tbody>
</table>

**Total** 1,870,851,567 53.6%

Consists of Derived Program Codes for:
- Selected Mercantile
- Alarm Installation
- Landowners
- Offices
- Welding
- Landscapers
- Home Health Care
- Exercise and Health

8.4% decline in already low loss ratios from older years to more recent

### Unprofitable Classes

**"C" Classes**

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Earned Premium</th>
<th>Ultimate Loss Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>18,434,129</td>
<td>100.9%</td>
</tr>
<tr>
<td>2012</td>
<td>19,897,639</td>
<td>109.9%</td>
</tr>
<tr>
<td>2013</td>
<td>22,863,494</td>
<td>103.7%</td>
</tr>
<tr>
<td>2014</td>
<td>28,277,897</td>
<td>130.1%</td>
</tr>
<tr>
<td>2015</td>
<td>31,133,767</td>
<td>138.1%</td>
</tr>
<tr>
<td>2016</td>
<td>32,026,072</td>
<td>127.2%</td>
</tr>
<tr>
<td>2017</td>
<td>33,078,866</td>
<td>154.9%</td>
</tr>
</tbody>
</table>

**Total** 185,711,865 127.1%

**"B" Classes**

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Earned Premium</th>
<th>Ultimate Loss Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>217,759,676</td>
<td>48.9%</td>
</tr>
<tr>
<td>2012</td>
<td>226,605,359</td>
<td>48.0%</td>
</tr>
<tr>
<td>2013</td>
<td>246,685,981</td>
<td>48.3%</td>
</tr>
<tr>
<td>2014</td>
<td>276,321,007</td>
<td>56.4%</td>
</tr>
<tr>
<td>2015</td>
<td>293,284,212</td>
<td>54.2%</td>
</tr>
<tr>
<td>2016</td>
<td>302,186,749</td>
<td>56.9%</td>
</tr>
<tr>
<td>2017</td>
<td>308,008,582</td>
<td>58.6%</td>
</tr>
</tbody>
</table>

**Total** 1,870,851,567 53.6%

### Remaining Classes

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Earned Premium</th>
<th>Ultimate Loss Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>61,195,262</td>
<td>105.1%</td>
</tr>
<tr>
<td>2012</td>
<td>124,516,603</td>
<td>137.9%</td>
</tr>
</tbody>
</table>

**Total** 155,711,865 127.1%

### New York Contractors

**NY Contractors**

<table>
<thead>
<tr>
<th>Accident Year</th>
<th>Earned Premium</th>
<th>Ultimate Loss Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>13,745,120</td>
<td>105.2%</td>
</tr>
<tr>
<td>2012</td>
<td>14,482,344</td>
<td>117.6%</td>
</tr>
<tr>
<td>2013</td>
<td>16,956,029</td>
<td>106.3%</td>
</tr>
<tr>
<td>2014</td>
<td>22,427,473</td>
<td>141.6%</td>
</tr>
<tr>
<td>2015</td>
<td>29,477,192</td>
<td>108.9%</td>
</tr>
<tr>
<td>2016</td>
<td>19,630,336</td>
<td>74.6%</td>
</tr>
<tr>
<td>2017</td>
<td>9,162,464</td>
<td>25.6%</td>
</tr>
</tbody>
</table>

**Total** 125,880,958 103.5%

Consists of all other business excluding New York Labor Law (NYLL). NYLL classes were excluded as Nationwide E&S has exited that business

- Increased 7.1% in more recent years
- There likely remains significant opportunity to further segment this business

Case 2 – Client B

Tiering – Illustrative Example of Segmentation

© 2017 KPMG LLP, a Delaware limited liability partnership and the U.S. member firm of the KPMG network of independent member firms affiliated with KPMG International Cooperative (“KPMG International”), a Swiss entity. All rights reserved.
Agenda

- Topic

  Reserve Modernization – Refresh

  Industry Overview – Current Pilots and Uses

  Applications of Reserve Modernization

Case Studies

Results Interpretation & Visualization
Model lift

Coverage one incremental paid loss model

Lift charts show the performance of the model on the test dataset (14Q1 – 16Q4).

- Predictions are binned from low to high into deciles.
- The red line (Predicted) tracks well with the blue line (Actual), except for some under fitting.
Relative variable importance

— Method of ranking the variables in the model in terms of their “importance”

— Importance of a variable calculated by crediting it with the reduction in the sum of squares

— Scaling done so that the variable with the largest reduction in sum of squares is one
Partial dependency plots

— Tool for visualizing the relationship of variables with target variable.

— Helpful with machine learning methods to provide more insight into the models.

— Partial dependence represents the effect of a predictor(s) on target variable after accounting for the average effects of the other predictors.

— Use caution if the variable whose partial dependence you are calculating has interactions with the remaining variables.

Note:
All else being equal, Group 1 has 70% greater incremental payments per quarter than Group 3.

*Example – not derived from any company sources*
POJO review

- Final tree structure can be viewed in “Plain Old Java Object” format
- Interpretation of variable usage in tree structure
  - Variable type
  - Tree split points

```java
class model_gbm.validation.Tree_0_class_0 {
    static final double score0(double[] data) {
        double pred =
            (Double.isNaN(data[0]) || data[8] /* UnderwritingOffice */ < 20.5f
            (data[7] /* CededIndicator */ < 1.0002446f
            (Double.isNaN(data[10]) || data[10] /* InitialExpectedLoss */ < 14186.336f
                78.54965f:
                25125.646f:
                (data[10] /* InitialExpectedLoss */ < -12199.999f
                    -21845.748f:
                    1463.4309f)) :
            (Double.isNaN(data[10]) || data[10] /* InitialExpectedLoss */ < 120323.08f
                (Double.isNaN(data[0] /* ClaimType */) && (GenModel.bitSetIsInRange(GRPSSPLIT0, 0, data[0]) && !GenModel.bitSetContains(GRPSSPLIT0, 0, data[0])))
                    -878.5838f:
                    771.9814f:
                    22522.334f));
        return pred;
    }  // constant pool size = 318, number of visited nodes = 6, static init size = 308
    // {01000100 00101010 10000000 00011100}
    public static final byte[] GRPSSPLIT0 = new byte[] {34, 84, 1, 56};
};
```
Data Visualization – Granular Investigation

Ultimate and Ultimate Loss Ratio by Accident Year

Loss Ratios by Accident Year and Industry

Box & Whiskers Breakout

Industry

© 2017 KPMG LLP, a Delaware limited liability partnership and the U.S. member firm of the KPMG network of independent member firms affiliated with KPMG International Cooperative (“KPMG International”), a Swiss entity. All rights reserved.
Thank you

<table>
<thead>
<tr>
<th></th>
<th>Nate Loughin</th>
<th>Frankie Logan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Office</strong></td>
<td>610-230-2068</td>
<td>610-263-2901</td>
</tr>
<tr>
<td><strong>Cell</strong></td>
<td>610-348-5126</td>
<td>307-575-0945</td>
</tr>
<tr>
<td><strong>Email</strong></td>
<td><a href="mailto:nloughin@kpmg.com">nloughin@kpmg.com</a></td>
<td><a href="mailto:flogan@kpmg.com">flogan@kpmg.com</a></td>
</tr>
</tbody>
</table>
Appendix:
Machine Learning Basics
A tree is a simple set of splitting rules on the data, what we call a “weak learner”

A group of “weak learners” can come together to form a “strong learner”

**Random Forest** is a collection of “weak learners” (trees) built using bootstrap sample of training data. The prediction is a combination of predictions over the individual trees.

**Gradient Boosting** is a collection of “weak learners” (trees) used sequentially, with each tree focused on improving the prediction of the previous tree. In each step a bootstrap sample of data is taken. A tree is fit to the “current residuals” and the residuals are updated for the next step.
Gradient Boosted Regression Trees

What's in a name?

Gradient Boosting

- Machine learning technique that combines many weak models into a stronger model (ensembling)

Regression Tree

- Predictive model that can be represented using a tree

First described by Friedman (2001).
Some Heuristics*

One decision tree

*via a very simplified illustrative example
Some Heuristics

One decision tree

Not so great performance

<table>
<thead>
<tr>
<th>AY</th>
<th>Dev</th>
<th>Increm loss</th>
<th>Predicted</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>1</td>
<td>5,000</td>
<td>5,000</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>2</td>
<td>7,000</td>
<td>5,000</td>
<td>2,000</td>
</tr>
<tr>
<td>2000</td>
<td>3</td>
<td>4,000</td>
<td>3,000</td>
<td>1,000</td>
</tr>
<tr>
<td>2000</td>
<td>4</td>
<td>2,000</td>
<td>3,000</td>
<td>-1,000</td>
</tr>
<tr>
<td>2001</td>
<td>1</td>
<td>4,500</td>
<td>5,000</td>
<td>-500</td>
</tr>
<tr>
<td>2001</td>
<td>2</td>
<td>6,500</td>
<td>5,000</td>
<td>1,500</td>
</tr>
<tr>
<td>2001</td>
<td>3</td>
<td>3,500</td>
<td>3,000</td>
<td>500</td>
</tr>
<tr>
<td>2002</td>
<td>1</td>
<td>5,500</td>
<td>6,500</td>
<td>-1,000</td>
</tr>
<tr>
<td>2002</td>
<td>2</td>
<td>7,500</td>
<td>6,500</td>
<td>1,000</td>
</tr>
<tr>
<td>2003</td>
<td>1</td>
<td>6,000</td>
<td>6,500</td>
<td>-500</td>
</tr>
</tbody>
</table>
Some Heuristics

Boosting (more decision trees…)

<table>
<thead>
<tr>
<th>AY</th>
<th>Dev</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>2</td>
<td>2,000</td>
</tr>
<tr>
<td>2000</td>
<td>3</td>
<td>1,000</td>
</tr>
<tr>
<td>2000</td>
<td>4</td>
<td>-1,000</td>
</tr>
<tr>
<td>2001</td>
<td>1</td>
<td>-500</td>
</tr>
<tr>
<td>2001</td>
<td>2</td>
<td>1,500</td>
</tr>
<tr>
<td>2001</td>
<td>3</td>
<td>500</td>
</tr>
<tr>
<td>2002</td>
<td>1</td>
<td>-1,000</td>
</tr>
<tr>
<td>2002</td>
<td>2</td>
<td>1,000</td>
</tr>
<tr>
<td>2003</td>
<td>1</td>
<td>-500</td>
</tr>
</tbody>
</table>

Build another tree to fit to the residuals
Some Heuristics

Boosting (more decision trees…)

Each tree tries to correct the error of the previous trees. By constructing a sequence of many trees we’ll have ourselves a decent model.

<table>
<thead>
<tr>
<th>AY</th>
<th>Dev</th>
<th>Residual (1st tree)</th>
<th>Prediction (2nd tree)</th>
<th>Residual (2nd tree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>1</td>
<td>0</td>
<td>1,500</td>
<td>-1,500</td>
</tr>
<tr>
<td>2000</td>
<td>2</td>
<td>2,000</td>
<td>1,500</td>
<td>500</td>
</tr>
<tr>
<td>2000</td>
<td>3</td>
<td>1,000</td>
<td>1,500</td>
<td>-500</td>
</tr>
<tr>
<td>2000</td>
<td>4</td>
<td>-1,000</td>
<td>-1,000</td>
<td>0</td>
</tr>
<tr>
<td>2001</td>
<td>1</td>
<td>-500</td>
<td>1,500</td>
<td>-2,000</td>
</tr>
<tr>
<td>2001</td>
<td>2</td>
<td>1,500</td>
<td>1,500</td>
<td>0</td>
</tr>
<tr>
<td>2001</td>
<td>3</td>
<td>500</td>
<td>1,500</td>
<td>-1,000</td>
</tr>
<tr>
<td>2002</td>
<td>1</td>
<td>-1,000</td>
<td>-500</td>
<td>-500</td>
</tr>
<tr>
<td>2002</td>
<td>2</td>
<td>1,000</td>
<td>-500</td>
<td>1,500</td>
</tr>
<tr>
<td>2003</td>
<td>1</td>
<td>-500</td>
<td>-500</td>
<td>0</td>
</tr>
</tbody>
</table>

2,000 – 1,500 = 500
Hyperparameters & Tuning

There are many ways to specify a GBM algorithm; as examples,

- Number of trees

- Depth of trees

- Learning rate

+ \gamma \cdot + \gamma \cdot (\gamma < 1) + \gamma \cdot + \gamma \cdot + \ldots
Hyperparameters & Tuning

There are many ways to specify a GBM algorithm

- Number of trees
- Depth of trees
- Learning rate
- Sampling rate of training data
- Sampling rate of predictors
- ...

- 50, 100, 200
- 1, 5, 20
- 0.01, 0.1
- 0.5, 0.8
- 0.5, 0.8
- ...
Hyperparameters & Tuning

How do we pick the best one(s)?

- Number of trees
- Depth of trees
- Learning rate
- Sampling rate of training data
- Sampling rate of predictors
- ...

- 50, 100, 200
- 1, 5, 20
- 0.01, 0.1
- 0.5, 0.8
- 0.5, 0.8
- ...

$3(3)(2)(2)(2) = 72$ combinations!
Hyperparameters & Tuning

“Autopilot”

\[ 3(3)(2)(2)(2) = 72 \text{ combinations!} \]

- Models are fit using each of the 72 combinations and are compared using cross-validation, the combination of hyperparameters with the lowest MSE is then fit to the total data set.
Hyperparameters & Tuning

“Autopilot”

\[3(3)(2)(2)(2) = 72 \text{ combinations!}\]

- Models are fit using each of the 72 combinations and are compared using cross-validation, the combination of hyperparameters with the lowest MSE is then fit to the total data set.
- We can feed into our funnel more than one type of algorithm. In other words, we can simultaneously test GBM, GLM, and other techniques such as Random Forests or Neural Networks, much like actuaries considering Chain Ladder and Bornhuetter-Ferguson
Hyperparameters & Tuning

“Autopilot”

3(3)(2)(2)(2) = 72 combinations!

- Models are fit using each of the 72 combinations and are compared using cross-validation, the combination of hyperparameters with the lowest MSE is then fit to the total data set.

- We can feed into our funnel more than one type of algorithm. In other words, we can simultaneously test GBM, GLM, and other techniques such as Random Forests or Neural Networks, much like actuaries considering Chain Ladder and Bornhuetter-Ferguson.

- Instead of building one model, we build a pipeline which generates a model on its own for subsequent review dates.
Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates.

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

© 2017 KPMG LLP, a Delaware limited liability partnership and the U.S. member firm of the KPMG network of independent member firms affiliated with KPMG International Cooperative (“KPMG International”), a Swiss entity. All rights reserved. NDPPS 718321

The KPMG name and logo are registered trademarks or trademarks of KPMG International.