Artificial Intelligence in Insurance

Actuarial analysis in the digital age
Actuarial Society of Greater New York – 2017 Annual Meeting
November 13, 2017
Motivation

— Many new technological innovations are available today

— A wide berth of industries will be affected by artificial intelligence

— Question:
  - How about Insurance industry?
Agenda

Insurance modernization – A new vision to drive enterprise performance

Introduction to artificial intelligence applications in actuarial analysis

Robotic process automation

Introduction to cognitive methods and applications

Interpretation and explanation of results

Enterprise response
Insurance modernization

— Recent changes
  - Underwriting
  - Claim handling

— How actuarial analysis is performed has changed little over the last century
Emerging technology

**Illustrative packages**

The emergence of new technology, coupled with enhanced computing power, has the potential to radically disrupt this historic approach.

- **Data preparation**
  - R, HIVE, python, hadoop

- **Cognitive – machine learning**
  - sas, R, python

- **Visualization**
  - tableau, Qlik

- **Robotic process automation**
  - AUTOMATION ANYWHERE, blueprism

---

**Computing power has increased significantly over time**

We have seen a 1 trillion-fold increase in computer processing capabilities over the past 60 years(1)

**Today’s smartphone has more computing power than the Apollo 11 Guidance Computer**

CPU speed in GHz (2)

Source: (1) Experts Exchange, “Processing Power Compared”

Source: (2) Frost & Sullivan, “Addressing Mobile Cybersecurity”
A new vision

Insight driven performance

We envision a new way of doing actuarial work: broader more granular data feed sophisticated actuarial software that automatically determines correlations and predicts for future. It enables flexible real-time analyses to identify trends faster and take coordinated actions sooner across all key departments to strengthen performance.

Gather and Prepare Data
- Structured files
- Manual cleansing
- Aggregate-level

Produce analysis
- Spreadsheets calculations
- Past experience
- Quarterly

Review and approve
- Judgment-based
- Large-change focus
- Summary-level

Analyze and communicate
- Manual investigation into cause/effects
- Shadow department data and analyses
- Aggregate-level
- Std. Powerpoint reports

Take action
- Silo-ed departmental responses
- Delayed insight – delayed action

Today
- Structured files
- Manual cleansing
- Aggregate-level
- Spreadsheets calculations
- Past experience
- Quarterly
- Judgment-based
- Large-change focus
- Summary-level
- Manual investigation into cause/effects
- Shadow department data and analyses
- Aggregate-level
- Std. Powerpoint reports

Future
- Structured files + unstructured data
- Granular accident/coverage-level information
- Statistical software
- Automated selection
- Cognitive continual learning
- Any time update
- Initial test for fit, then increasing reliance
- Redirect effort on trend detection
- Focus on core causes/change drivers
- Fast trend identification, granular level insights
- Internal users enabled to conduct own research
- Visualization software: standard, tailored, ad hoc
- One common data source
- New operating model to quickly translate insight into action
- Coordination across all stakeholders
- Timely response/early corrective actions
Potential benefits

Granular, flexible, fast, actionable

1. **Granular data**
   - Information captured at accident/coverage level
   - Combination of structured and unstructured data
   - Flexibility to aggregate and analyze as desired

2. **Deeper and quicker insight**
   - A more precise analysis production processes
   - Ability to identify trends and other business insight faster
   - Selections based on innate risk and claims characteristics
   - Analysis reflect the detailed risks

3. **Faster reaction**
   - Realize changes in the environment more quickly, and react
   - Techniques that respond as claims are reported
   - Analysis re-parameterized regularly using machine learning techniques

4. **Frequency of review**
   - Run actuarial analyses at any valuation date for which data is available (e.g., automatically run weekly)
   - For example, an analysis could easily be run a few weeks before close

5. **Increased efficiency**
   - Robotic process automation can lead to increased speed to close
   - Actuarial analysts are freed up to digest the trends and communicate them to the organization, for timely actions

6. **Seamless communication**
   - A modern reserving process produces an output ready made for deriving insights using visualization tools
   - Others can be given views of the data appropriate to their access requirements, to derive their own insights for their business segments
Agenda

Insurance modernization – A new vision to drive enterprise performance

Introduction to artificial intelligence applications in actuarial analysis

Robotic process automation

Introduction to cognitive methods and applications

Interpretation and explanation of results

Enterprise response
Future analysis

Four core components

With cognitive analysis at the core, results interpretation and presentation surround this automated analytical engine to enable easy and fast consumption of claim insights. The enterprise response component coordinates and mobilizes actions across the stakeholder departments, while RPA ultimately aims to further streamline the underlying processes.
The components are interconnected and build on each other

<table>
<thead>
<tr>
<th>Description</th>
<th>Approach</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robotic process automation</td>
<td>Cognitive Analysis</td>
<td>Results Presentation</td>
</tr>
<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>
</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>
</tr>
<tr>
<td></td>
<td></td>
<td></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>
</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>
</tr>
<tr>
<td></td>
<td></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>
</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>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Agenda

Insurance modernization – A new vision to drive enterprise performance

Introduction to artificial intelligence applications in actuarial analysis

Robotic process automation

Introduction to cognitive methods and applications

Interpretation and explanation of results

Enterprise response
Robotic process automation has the ability to improve operational efficiencies across the entire claims operation. The reserving process is particularly ripe for automation.

### The path to cognitive automation

**Class 1: Basic process automation**
- Learning
- Rules

**Class 2: Enhanced process automation**
- Reasoning

**Class 3: Cognitive automation**
- Limited RPA opportunities or not enough information

---

**Legend:**
- CLASS 1: Basic Process Automation
- CLASS 2: Enhanced Process Automation
- CLASS 3: Autonomic/Cognitive
- Limited RPA opportunities or not enough information
## Use case

### RPA for analysis

Using RPA bot process automation, we have configured a bot to complete 8 of the 18 high-level manual tasks in the analyst’s analysis process.

<table>
<thead>
<tr>
<th>Task</th>
<th>Tool(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create folder structure and copy prior analysis files</td>
<td>Share folder</td>
</tr>
<tr>
<td>Reconcile input data</td>
<td>Analysis Software; Excel pivot tables</td>
</tr>
<tr>
<td>Pull large loss data</td>
<td>Large loss system</td>
</tr>
<tr>
<td>Pull rate data</td>
<td>SharePoint; Excel</td>
</tr>
<tr>
<td>Pull trend data</td>
<td>Various reports</td>
</tr>
<tr>
<td>Replicate reports</td>
<td>Analysis Software</td>
</tr>
<tr>
<td>Review prior study memos</td>
<td>Share folder</td>
</tr>
<tr>
<td>Selection of parameters</td>
<td>Analysis Software</td>
</tr>
<tr>
<td>Calculate impact from parameter selection</td>
<td>Analysis Software</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task</th>
<th>Tool(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select a method</td>
<td>Analysis Software</td>
</tr>
<tr>
<td>Calculate selection impact</td>
<td>Analysis Software</td>
</tr>
<tr>
<td>Ad-Hoc analysis</td>
<td>Analysis Software</td>
</tr>
<tr>
<td>Produce memo, checklist, etc.</td>
<td>Word, Excel</td>
</tr>
<tr>
<td>Compile PDF outputs</td>
<td>Acrobat</td>
</tr>
<tr>
<td>Peer review data checks</td>
<td>Share folder</td>
</tr>
<tr>
<td>Peer review judgment checks</td>
<td>Share folder</td>
</tr>
<tr>
<td>Manager review</td>
<td>Share folder</td>
</tr>
<tr>
<td>Update booked estimates in financial system</td>
<td>Share folder; Excel</td>
</tr>
</tbody>
</table>

— In 10 weeks, we automated 18% of analyst effort in analysis
— We also identified process re-engineering opportunities (incl. RPA) that are expected to reduce analyst effort approximately 50%
Agenda

Topic

Insurance modernization – A new vision to drive enterprise performance

Introduction to artificial intelligence applications in actuarial analysis

Robotic process automation

Introduction to cognitive methods and applications

Interpretation and explanation of results

Enterprise response
Cognitive computing/analysis

Cognitive computing – Systems which mimic the functioning of the human brain such as the ability to learn, understand, reason, and interact.¹

Cognitive analysis – refers to leveraging cognitive computing to make the actuarial analysis more efficient and more insightful.

¹“Computing, cognition and the future of knowing – How humans and machines are forging a new age of understanding”, John Kelly, IBM Research and Solutions Portfolio
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.
Decision tree heuristics

One decision tree

*via a very simplified illustrative example*
Decision tree heuristics (continued)

One decision tree

<table>
<thead>
<tr>
<th>Initial Claim Amount</th>
<th>Procedure Reviewed</th>
<th>Incremental Loss</th>
<th>Predicted</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>150,000</td>
<td>Yes</td>
<td>100,000</td>
<td>150,000</td>
<td>(50,000)</td>
</tr>
<tr>
<td>10,000</td>
<td>No</td>
<td>15,000</td>
<td>5,000</td>
<td>10,000</td>
</tr>
<tr>
<td>5,000</td>
<td>Yes</td>
<td>3,000</td>
<td>5,000</td>
<td>(2,000)</td>
</tr>
<tr>
<td>200,000</td>
<td>Yes</td>
<td>200,000</td>
<td>150,000</td>
<td>50,000</td>
</tr>
<tr>
<td>100,000</td>
<td>No</td>
<td>10,000</td>
<td>5,000</td>
<td>5,000</td>
</tr>
<tr>
<td>125,000</td>
<td>No</td>
<td>50,000</td>
<td>75,000</td>
<td>(25,000)</td>
</tr>
<tr>
<td>50,000</td>
<td>No</td>
<td>8,000</td>
<td>5,000</td>
<td>3,000</td>
</tr>
<tr>
<td>1,000</td>
<td>No</td>
<td>2,000</td>
<td>5,000</td>
<td>(3,000)</td>
</tr>
<tr>
<td>250,000</td>
<td>Yes</td>
<td>125,000</td>
<td>150,000</td>
<td>(25,000)</td>
</tr>
<tr>
<td>25,000</td>
<td>Yes</td>
<td>4,000</td>
<td>5,000</td>
<td>(1,000)</td>
</tr>
</tbody>
</table>
Decision tree heuristics (continued)

Boosting (more decision trees…)

<table>
<thead>
<tr>
<th>Initial Claim Amount</th>
<th>Procedure Reviewed?</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>150,000 Yes</td>
<td>(50,000)</td>
<td></td>
</tr>
<tr>
<td>10,000 No</td>
<td>10,000</td>
<td></td>
</tr>
<tr>
<td>5,000 Yes</td>
<td>(2,000)</td>
<td></td>
</tr>
<tr>
<td>200,000 Yes</td>
<td>50,000</td>
<td></td>
</tr>
<tr>
<td>100,000 No</td>
<td>5,000</td>
<td></td>
</tr>
<tr>
<td>125,000 No</td>
<td>(25,000)</td>
<td></td>
</tr>
<tr>
<td>50,000 No</td>
<td>3,000</td>
<td></td>
</tr>
<tr>
<td>1,000 No</td>
<td>(3,000)</td>
<td></td>
</tr>
<tr>
<td>250,000 Yes</td>
<td>(25,000)</td>
<td></td>
</tr>
<tr>
<td>25,000 Yes</td>
<td>(1,000)</td>
<td></td>
</tr>
</tbody>
</table>

Initial Claim Amount >= 100K?
- Yes
- No

Procedure Reviewed?
- Yes
- No

Build another tree to fit to the residuals

1,400
Decision tree heuristics (continued)

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>Initial Claim Amount</th>
<th>Procedure Reviewed?</th>
<th>Residual (1st tree)</th>
<th>Prediction (2nd tree)</th>
<th>Residual (2nd tree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>150,000 Yes</td>
<td></td>
<td>(50,000)</td>
<td>(8,333)</td>
<td>(41,667)</td>
</tr>
<tr>
<td>10,000 No</td>
<td></td>
<td>10,000</td>
<td>1,400</td>
<td>8,600</td>
</tr>
<tr>
<td>5,000 Yes</td>
<td></td>
<td>(2,000)</td>
<td>1,400</td>
<td>(3,400)</td>
</tr>
<tr>
<td>200,000 Yes</td>
<td></td>
<td>50,000</td>
<td>(8,333)</td>
<td>58,333</td>
</tr>
<tr>
<td>100,000 No</td>
<td></td>
<td>5,000</td>
<td>1,400</td>
<td>3,600</td>
</tr>
<tr>
<td>125,000 No</td>
<td></td>
<td>(25,000)</td>
<td>(10,000)</td>
<td>(15,000)</td>
</tr>
<tr>
<td>50,000 No</td>
<td></td>
<td>3,000</td>
<td>1,400</td>
<td>1,600</td>
</tr>
<tr>
<td>1,000 No</td>
<td></td>
<td>(3,000)</td>
<td>1,400</td>
<td>(4,400)</td>
</tr>
<tr>
<td>250,000 Yes</td>
<td></td>
<td>(25,000)</td>
<td>(8,333)</td>
<td>(16,667)</td>
</tr>
<tr>
<td>25,000 Yes</td>
<td></td>
<td>(1,000)</td>
<td>1,400</td>
<td>(2,400)</td>
</tr>
</tbody>
</table>

$10,000 - 1,400 = 8,600$
Decision tree hyperparameters

There are many ways to specify a decision tree algorithm; for example:

- Number of trees

- Depth of trees

- Learning rate

\[ \gamma \cdot \text{Decision Tree} + \gamma \cdot \text{Decision Tree} + \gamma \cdot \text{Decision Tree} + \ldots \]
<table>
<thead>
<tr>
<th>Models</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
</table>
| Random Forest                 | — Modeling non-linear and complex relationships  
— Fast algorithms – can leverage parallel computing                                                                                                                                         | — Can be difficult to explain  
— Often “beat” by well-tuned GBMs                                                                                                                                                                      |
| Gradient Boosting Machine (GBM) | — Modeling non-linear and complex relationships  
— Algorithm has more “levers” in terms of hyperparameters                                                                                                                                 | — Can be difficult to explain  
— Can be difficult to tune due to large number of hyperparameters                                                                                                                                 |
| Generalized Linear Models (GLM) | — Easy to explain  
— Well established in Actuarial community                                                                                                                                                    | — Need to be more explicit about interactions and non-linear relationships                                                                                                                                 |

Evaluating different model methods
Agenda

Topic

Insurance modernization – A new vision to drive enterprise performance

Introduction to artificial intelligence applications in actuarial analysis

Robotic process automation

Introduction to cognitive methods and applications

Interpretation and explanation of results

Enterprise response
Model validation

Model dataset (AY 2005 – 2013)

Accident period

2005

Training set: CalDr Qtr upto 10Q4

Validation set: CalDr Qtr 11Q1 – 13Q4

2010

Validation approach used for the selections of
— Model methods
— Hyper parameters
— Predictive variables

To ensure the models working properly
— Model Dataset could be split into training and validation by three calendar quarter cuts by:
- 10Q4
- 11Q2
- 11Q4
Model testing

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 emergence
  - Using data test set – part 1
— Model predicted ultimates vs. traditional methods predicted ultimates
  - Using data test set – part 1 + part 2
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 = 0;
        if (Double.isNaN(data[0]) || data[10] < 20.5f)
            pred = 1;
        if (Double.isNaN(data[10]) || data[10] < 1.0002446f)
            pred = 2;
        if (Double.isNaN(data[10]) || data[10] < 78.54965f)
            pred = 3;
        if (Double.isNaN(data[10]) || data[10] < 25125.646f)
            pred = 4;
        if (Double.isNaN(data[10]) || data[10] < -12199.999f)
            pred = 5;
        if (Double.isNaN(data[10]) || data[10] < -21045.748f)
            pred = 6;
        if (Double.isNaN(data[10]) || data[10] < -1463.4309f)
            pred = 7;
        if (Double.isNaN(data[10]) || data[10] < 120323.08f)
            pred = 8;
        return pred;
    }
}
```

// Column domains. The last array contains domain of response column.
public static final String[][] DOMAINS = new String[[]] {
    "ClaimType": model_gbm_validation_ColInfo_0.VALUES,
    "ClaimDescriptionCode": model_gbm_validation_ColInfo_1.VALUES,
    "ClaimState": model_gbm_validation_ColInfo_2.VALUES,
    "DevelopmentMonth": null,
    "ClaimMonth": null,
    "TPANumber": model_gbm_validation_ColInfo_5.VALUES,
    "StatutoryCo": model_gbm_validation_ColInfo_6.VALUES,
    "CededIndicator": null,
    "UnderwritingOffice": model_gbm_validation_ColInfo_8.VALUES,
    "ClassGrouping": model_gbm_validation_ColInfo_9.VALUES,
    "InitialExpectedLoss": null,
};
Data visualization

Incorporating Digital Visualization to the Analysis results enables **deep, timely, and widespread** understanding of **complex** actuarial insights and the **interaction** of those insights.

The solution drives a **faster recognition** of claim developments and **root cause analysis**, with an **intuitive interface** for actuaries and executives alike.
Data visualization

Granularity of insights

Dynamic charts enable detailed understanding of results

- Granularity of visuals matches granularity of analysis
- Supports quick investigation of outliers
Data visualization

Insights into existing methods

Comparisons of traditional and cognitive results can assist with validation of:
— Traditional procedures (e.g., reserve allocations)
— Model performance
Similar dash-boarding concepts can help evaluate performance by business segment
Case studies

Case 1 – U.S. Operation of Global P&C Insurer – Cognitive Reserving Solution leveraging a Gradient Boosting Model to proactively identify deterioration of a problematic business segment

Case 2 – 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
Case 1
Responsiveness of cognitive methods

Estimated ultimate development for the LOB

AY_1

AY_2

More accurate early
More accurate early
Stable – No reversals

More accurate early
Case 1

Additional insights

What is causing this anomaly?
- Severity and claim reporting patterns are normal
- High Closed Claim Counts are causing the anomaly
  - New claims are being closed faster

Cognitive insights quickly rule out some hypotheses and confirm others.

How do Cognitive Methods know what is normal?

Machine Learning methods provide prediction ranges and percentiles, not just expected values. These ranges tell us which results are normal and which are unusually low or high.

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Reported claim counts</th>
<th>Closed claim counts</th>
<th>Paid severity ($000s)</th>
<th>Paid severity (Ex. Large) ($000s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXTREMELY LOW (Lowest 1%) 0</td>
<td>&lt; 13</td>
<td>&lt; 4</td>
<td>&lt; 2</td>
<td>&lt; 2</td>
</tr>
<tr>
<td>VERY LOW</td>
<td>1 – 9</td>
<td>4 – 15</td>
<td>2 – 4</td>
<td>2 – 4</td>
</tr>
<tr>
<td>Low</td>
<td>10 – 19</td>
<td>15 – 21</td>
<td>4 – 6</td>
<td>4 – 6</td>
</tr>
<tr>
<td>Normal</td>
<td>20 – 79</td>
<td>21 – 36</td>
<td>6 – 22</td>
<td>6 – 21</td>
</tr>
<tr>
<td>High</td>
<td>80 – 89</td>
<td>36 – 43</td>
<td>22 – 38</td>
<td>21 – 35</td>
</tr>
<tr>
<td>VERY HIGH</td>
<td>90 – 98</td>
<td>43 – 61</td>
<td>38 – 91</td>
<td>35 – 50</td>
</tr>
<tr>
<td>EXTREMELY HIGH (Highest 1%) 99</td>
<td>&gt; 87</td>
<td>&gt; 61</td>
<td>&gt; 91</td>
<td>&gt; 50</td>
</tr>
</tbody>
</table>
### 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>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</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*

---

**Case 1**

**Granularity of results**

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?
Case 2
Prediction of state change

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

— 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
Case 2

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 2

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
Agenda

Topic

Insurance modernization – A new vision to drive enterprise performance

Introduction to artificial intelligence applications in actuarial analysis

Robotic process automation

Introduction to cognitive methods and applications

Interpretation and explanation of results

Enterprise response
Enterprise response

1. Having the ability to deliver the insights broadly to management

2. Being positioned to make timely decisions/changes/reactions

3. Having the ability to put those into actions effectively
**Digital trust in AI analysis**

In its common and broadest form, digital trust involves customers, data and errors, and misuse or unintended consequences of related analytics. The pillars of digital trust are equally applicable in the context of loss reserve analysis and provide a framework for management and regulators to assess the actuary’s analysis.

### Quality

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are the data management practices appropriate?</td>
</tr>
<tr>
<td>Is the data timely, internally consistent, and complete?</td>
</tr>
<tr>
<td>Data quality assurance for first-generation machine learning approaches that build on existing actuarial data should not be significantly different from current quality requirements for actuarial data formats and segmentations.</td>
</tr>
</tbody>
</table>

### Accepted use

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confirmation that the estimation methods being developed are fit for their intended purpose will take on heightened importance.</td>
</tr>
<tr>
<td>The use, segmentation, and manipulation of data will have to be appropriate, documented, suitable for its intended purpose, and defensible.</td>
</tr>
</tbody>
</table>

### Accuracy

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictions and insights must provide timely actionable information that reflects reality.</td>
</tr>
<tr>
<td>We must also consider that models may be held to higher standards of precision than models used for purposes where directional indications are sufficient.</td>
</tr>
<tr>
<td>Increased frequency of analysis (e.g., from quarterly to weekly) is likely to be one factor in monitoring accuracy.</td>
</tr>
</tbody>
</table>

### Integrity

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data, models, and resulting predictions must be managed ethically and with the utmost attention to the veracity of the estimates.</td>
</tr>
<tr>
<td>Methods that rely upon actuarial judgment or are prone to manipulation could be compromised by perception of bias.</td>
</tr>
</tbody>
</table>
Preparing for evolving actuarial roles

—Learn R, Python, Blue Prism, etc.

—Study Machine Learning Methods and Output

—Creatively Assess Potential Benefits
Potential benefits of AI in actuarial analysis

<table>
<thead>
<tr>
<th>General</th>
<th>Life</th>
<th>Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserve Analysis</td>
<td>Cash Flow Projection</td>
<td>Health Plan Enrollment</td>
</tr>
<tr>
<td>Transformation</td>
<td>Mortality Table Modeling</td>
<td>Reconciliation</td>
</tr>
<tr>
<td>Actuarial Pricing</td>
<td>Lapse and Surrender</td>
<td>Medical Cost Forecasting</td>
</tr>
<tr>
<td>Transformation</td>
<td>Analysis</td>
<td>Automation of Clinical</td>
</tr>
<tr>
<td>Underwriting Analysis</td>
<td></td>
<td>Quality Reporting</td>
</tr>
<tr>
<td>Transformation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- These tools can be used to simplify and streamline most data processing and reconciliation
- New tools present opportunities for actuaries to expand our role as business experts with advanced analytical and statistical capabilities
- No reason that we cannot apply these techniques throughout the insurance space to:
  - Model consumer behavior to optimize health care outcomes and provide superior quality of service to policy holders
  - Apply modeling and automation to improve the sales, underwriting, and claims handling processes
Thank you

Nate Loughin
Office: 610-230-2068
Cell: 610-348-5126
nloughin@kpmg.com

Sophia Song
Office: 212-954-4336
Cell: 917-455-7785
chisong@KPMG.com
Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates.

kpmg.com/socialmedia

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.