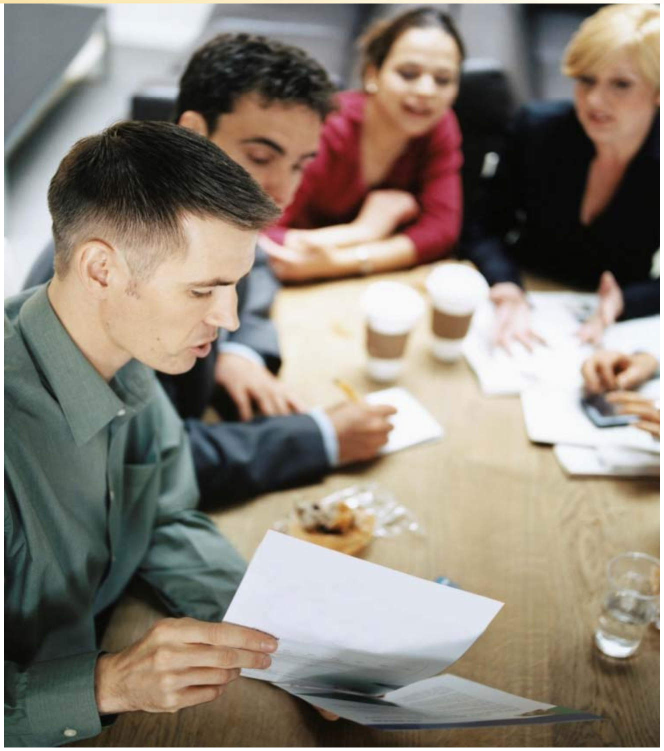
Antitrust Notice

- **The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.**
- **Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.**
- **It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.**



EagleEye Analytics

Expanding Analytics through the Use of Machine Learning

BACE Meeting 10 April 2012 **Christopher Cooksey, FCAS, MAAA**

EagleEye Analytics

Agenda…

1. **What is Machine Learning?**
2. **How can Machine Learning apply to insurance?**
3. **Model Validation**
4. **Non-rating Uses for Machine Learning**
5. **Rating Applications of Machine Learning**
6. **Analysis of high dimensional variables**

1. *What is Machine Learning?*

What is Machine Learning?

Machine Learning is a broad field concerned with the study of computer algorithms that automatically improve with experience.

A computer is said to "learn" from experience if…

… its performance on some set of tasks improves as experience increases.

This entire section draws heavily from Machine Learning, Tom M. Mitchell, McGraw-Hill, 1997.

What is Machine Learning?

"Machine Learning is a broad field concerned with the study of computer algorithms that automatically improve with experience."

Machine Learning, Tom M. Mitchell, McGraw Hill, 1997

"With algorithmic methods, there is no statistical model in the usual sense; no effort made to represent how the data were generated. And no apologies are offered for the absence of a model. There is a practical data analysis problem to solve that is attacked directly…"

"An Introduction to Ensemble Methods for Data Analysis", Richard A. Berk, UCLA, 2004

What is Machine Learning?

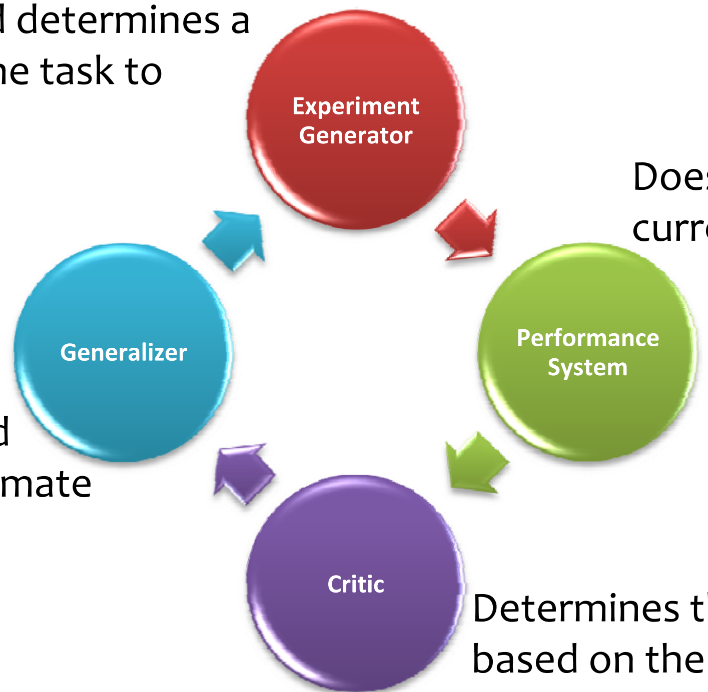
Applications of Machine Learning include…

- **Recognizing speech**
- **Driving an autonomous vehicle**
- **Predicting recovery rates of pneumonia patients**
- **Playing world-class backgammon**
- **Extracting valuable knowledge from large commercial databases**
- **Many, many, others…**

What is Machine Learning?

The general design of a machine learning approach can include…

Takes as input the currently learned best approach and determines a new example of the task to perform.



Does the “task” by using the currently learned best approach.

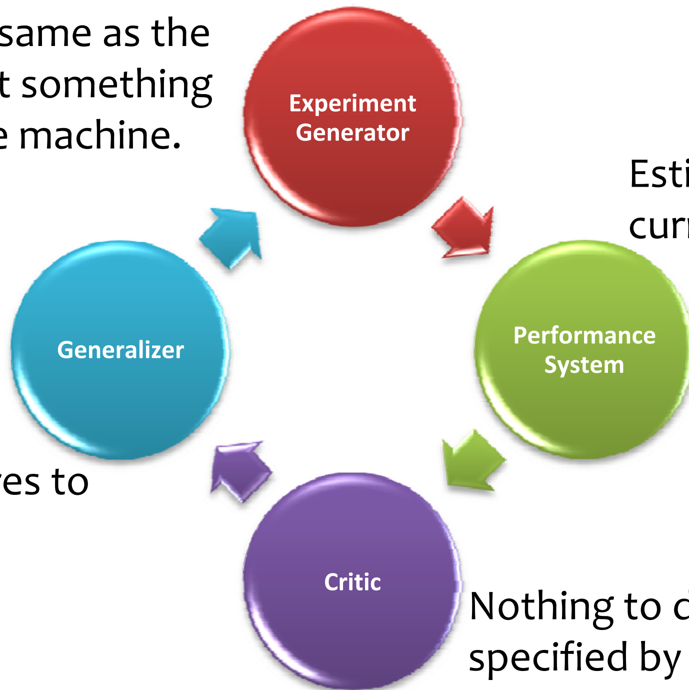
Determines the best way to train based on the output of the performance system.

Examines training examples and determines the best way to estimate the target function.

What is Machine Learning?

Assume you estimate trends using a weighted average of state trends, countrywide trends, and industry trends. What is the best set of weights?

Nothing to do here. The data to be estimated is the same as the training data, not something generated by the machine.



Estimates the trend using the current weights.

Uses the current experience period and least mean squares to estimate the weights.

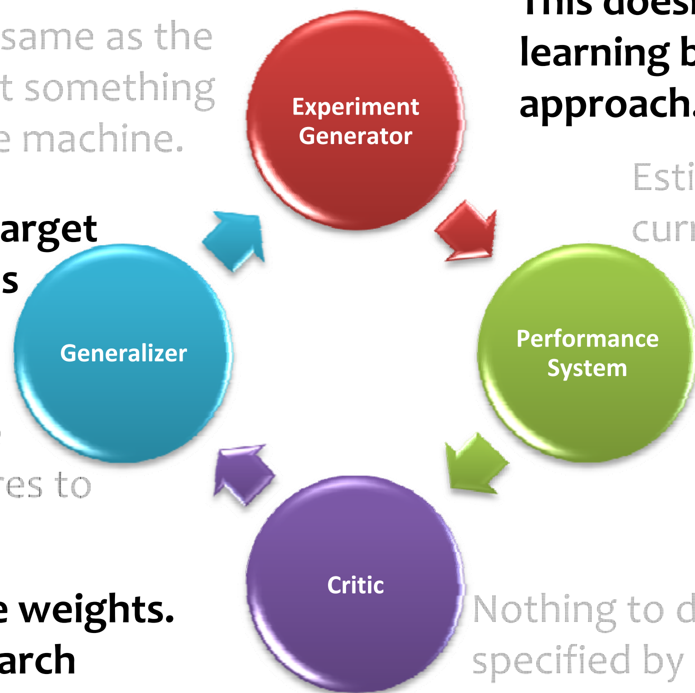
Nothing to do here. Training data is specified by the user, not the machine, and doesn't change based on system performance.

What is Machine Learning?

Assume you estimate trends using a weighted average of state trends, countrywide trends, and industry trends. What is the best set of weights?

Nothing to do here. The data to be estimated is the same as the training data, not something generated by the machine.

This doesn't "feel" like machine learning because of our traditional approach.



Estimates the trend using the current weights.

We look at the data as one group of data. Machine learning sees each policy as another training example.

Nothing to do here. Training data is specified by the user, not the machine, and doesn't change based on system performance.

Uses the current experience period and least mean squares to estimate the weights.

We see one estimate of the weights. Machine learning sees a search problem among all possible weights.

Machine learning asks explicit questions regarding how the target is estimated, how we know it is good, and how it might be improved.

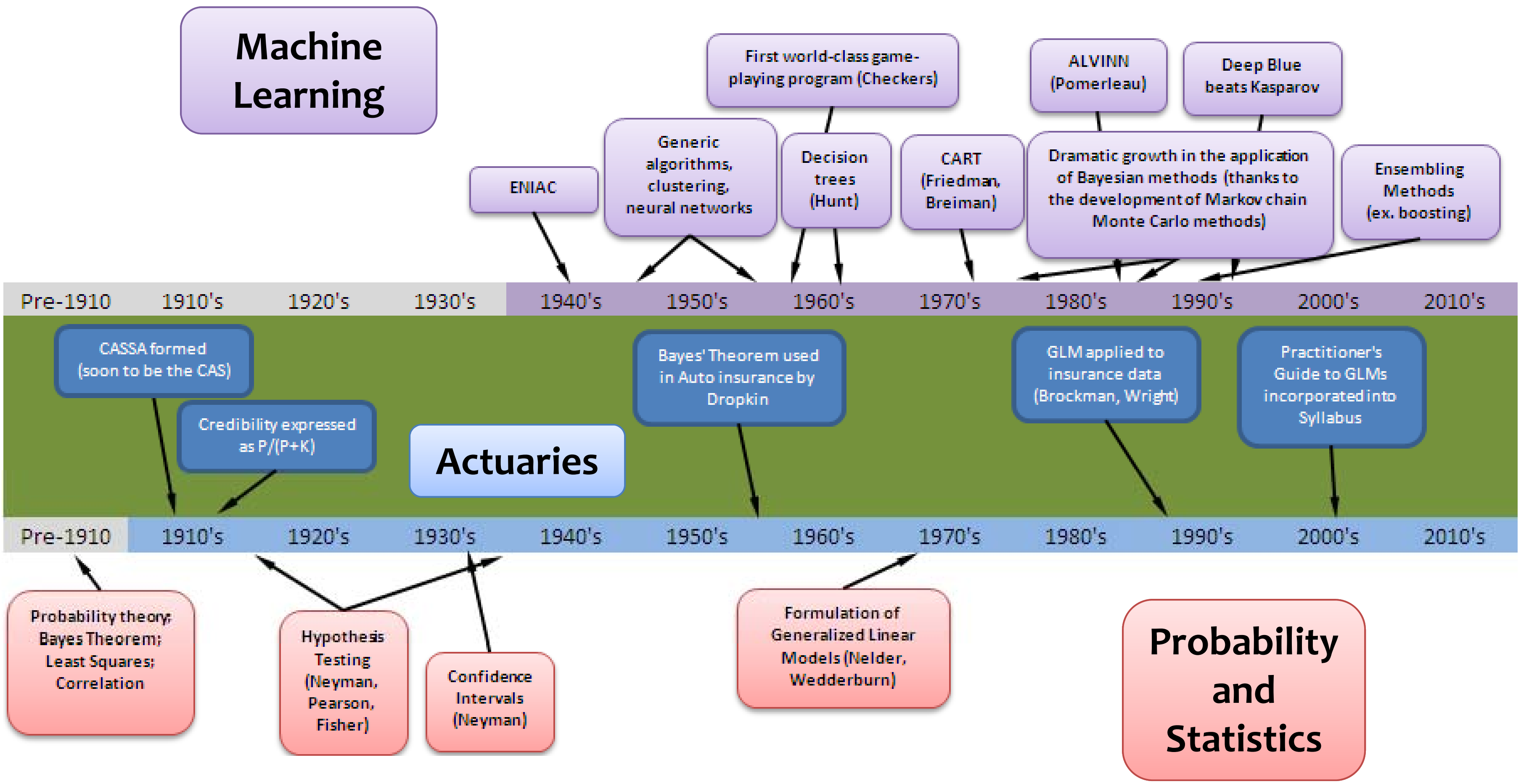
What is Machine Learning?

"Solving" a System of Equations	Gradient Descent
Predictive model with unknown parameters	Predictive model with unknown parameters
Define error in terms of unknown parameters	Define error in terms of unknown parameters
Take partial derivative of error equation with respect to each unknown	Take partial derivative of error equation with respect to each unknown
Set equations equal to zero and find the parameters which solve this system of equations	Give unknown parameters starting values – determine the change in values which moves the error lower
When derivatives are zero, you have a min (or max) error	Searches the error space by iteratively moving towards the lowest error

Limited to only those models which ***can*** be solved.

More general approach, but must worry about local minima.

What is Machine Learning?



2.

How can Machine Learning apply to insurance?

How can Machine Learning apply to insurance?

Machine Learning includes many different approaches…

- **Neural networks**
- **Decision trees**
- **Genetic algorithms**
- **Instance-based learning**
- **Others**

…and many different approaches for improving results

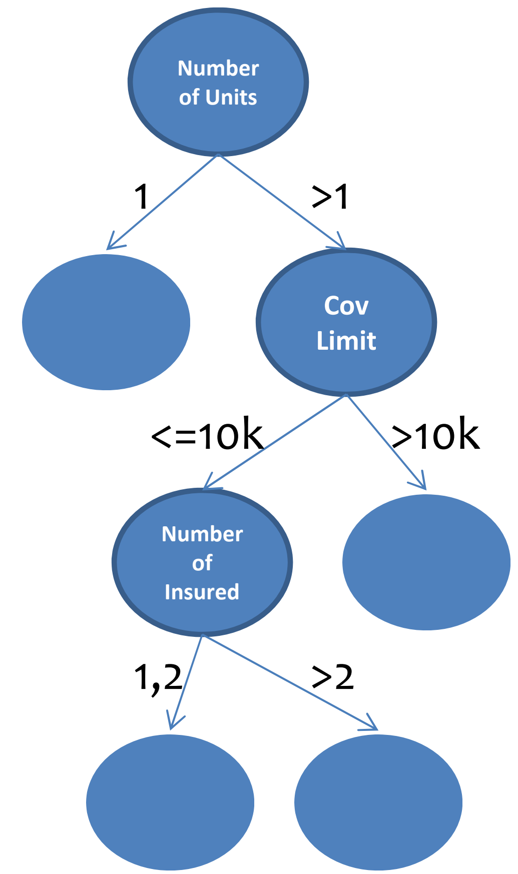
- **Ensembling**
- **Boosting**
- **Bagging**
- **Bayesian learning**
- **Others**

Focus here on decision trees – applicable to insurance & accessible

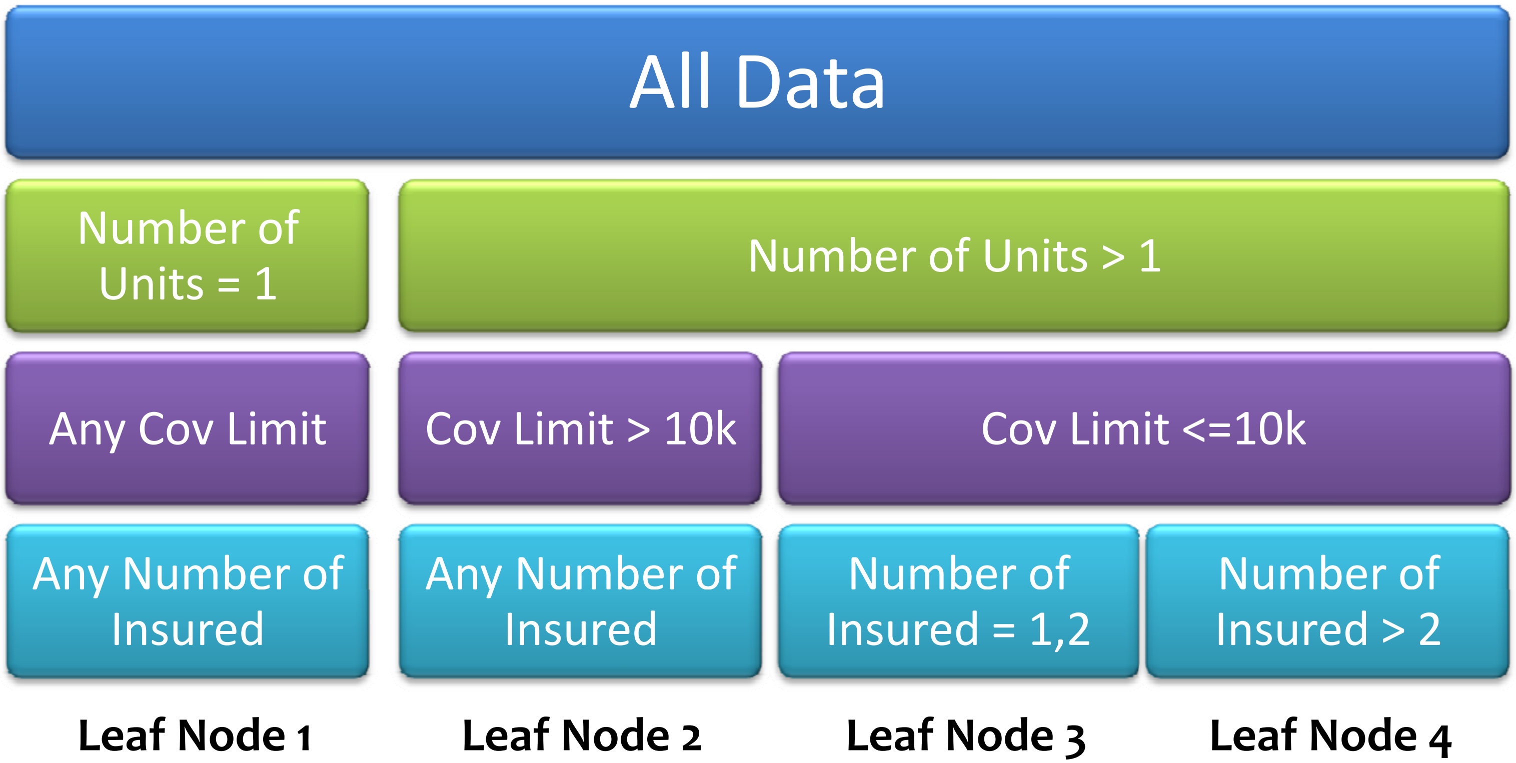
How can Machine Learning apply to insurance?

Basic Approach of Decision Trees

- **Data split based on some target and criterion**
  - Target: entropy, frequency, severity, loss ratio, loss cost, etc.
  - Criteria: maximize the difference, maximize the Gini coefficient, minimize the entropy, etc.
- **Each path is split again until some ending criterion is met**
  - Statistical tests on the utility of further splitting
  - No further improvement possible
  - Others
- **The tree may include some pruning criteria**
  - Performance on a validation set of data (i.e. reduced error pruning)
  - Rule post-pruning
  - Others

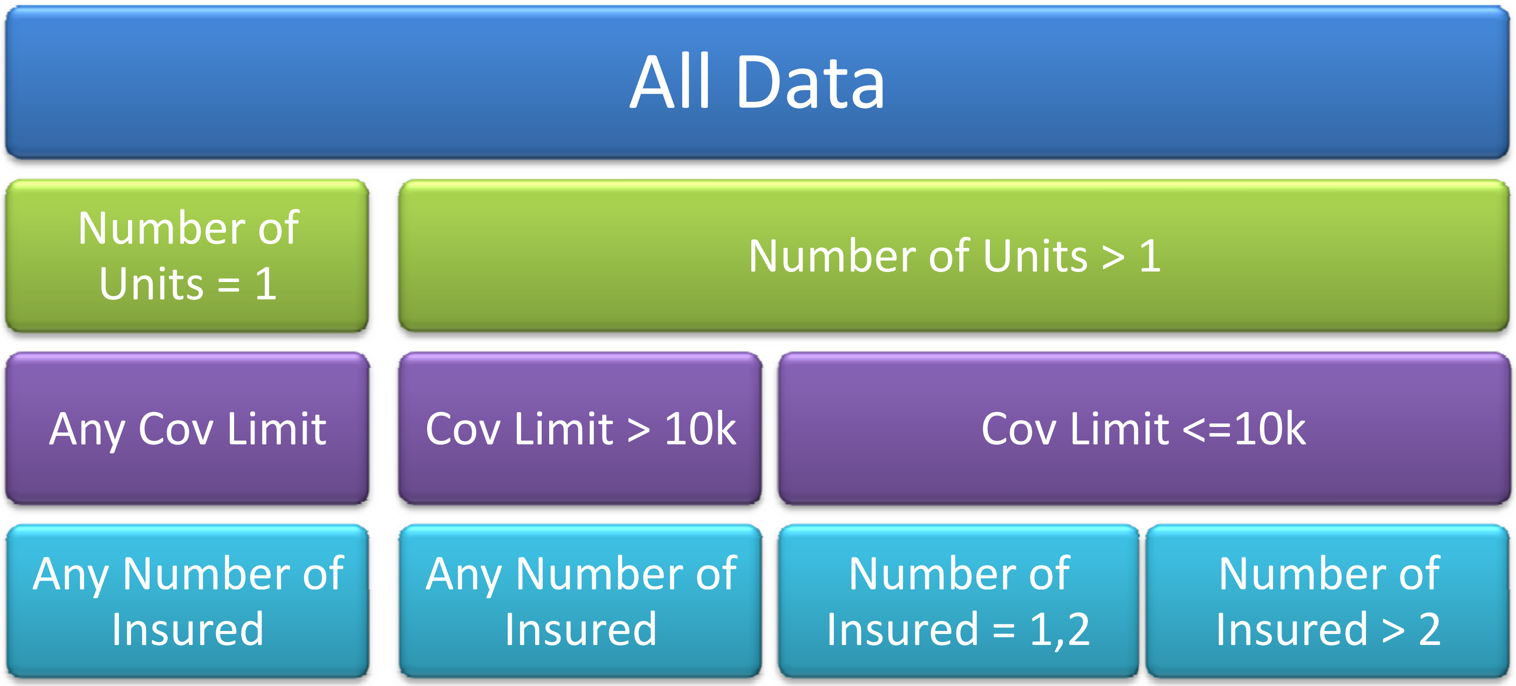


How can Machine Learning apply to insurance?



- **In decision trees all the data is assigned to one leaf node only**
- **Not all attributes are used in each path –**
  for example, Leaf Node 2 does not use Number of Insured

How can Machine Learning apply to insurance?



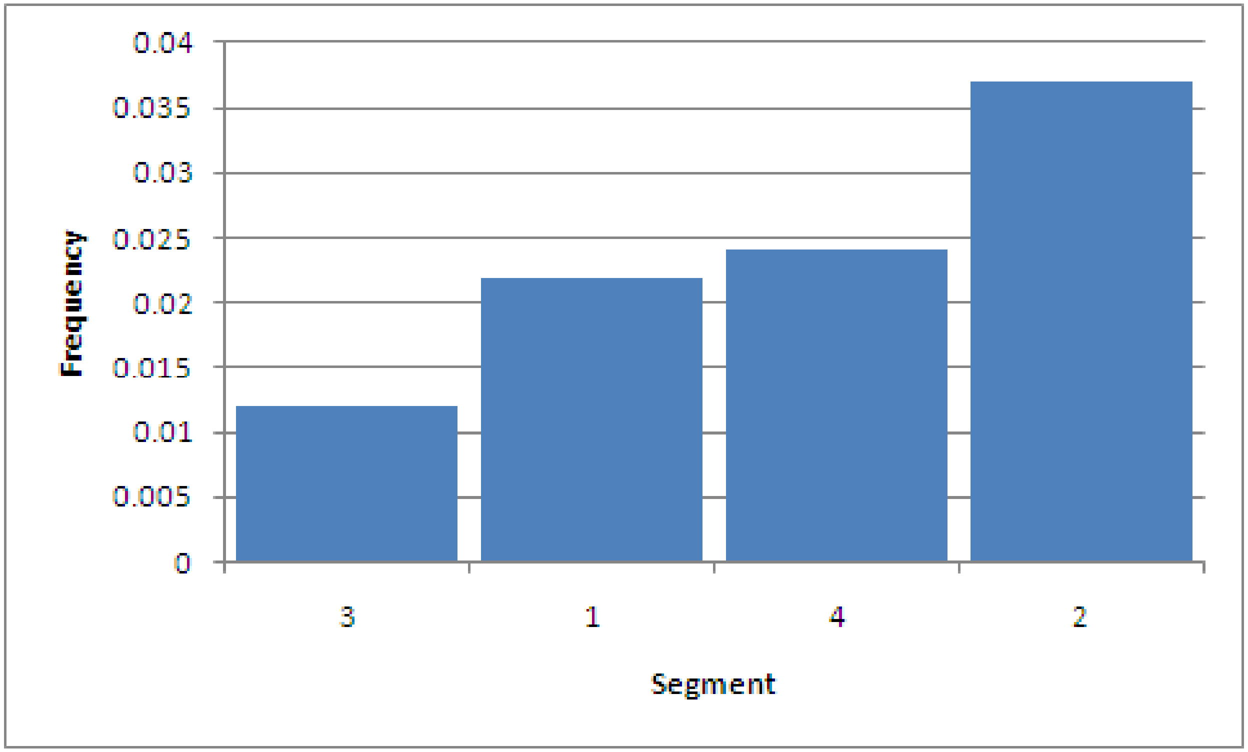
Freq = 0.022	**Freq = 0.037**	**Freq = 0.012**	**Freq = 0.024**
Segment 1	**Segment 2**	**Segment 3**	**Segment 4**

- **Decision trees are easily expressed as lift curves**
- **Segments are relatively easily described**

How can Machine Learning apply to insurance?

Who are my highest frequency customers?

- Policies with higher coverage limits (>10k) and multiple units (>1)

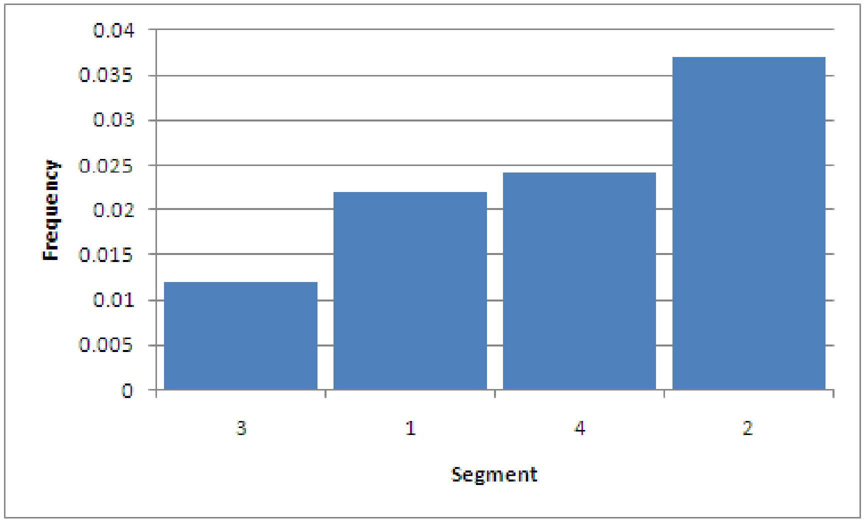


Who are my lowest frequency customers?

- Policies with lower coverage limts (<=10k), multiple units (>1), but lower numbers of insureds (1 or 2)

How can Machine Learning apply to insurance?

This approach can be used on different types of data
- Pricing
- Underwriting
- Claims
- Marketing
- Etc.



This approach can be used to target different criteria
- Frequency
- Severity
- Loss Ratio
- Retention
- Etc.

This approach can be used at different levels
- Vehicle/Coverage or Peril
- Vehicle
- Unit/building
- Policy
- Etc.

3. *Model Validation*

Why validate models?

Because you have to…

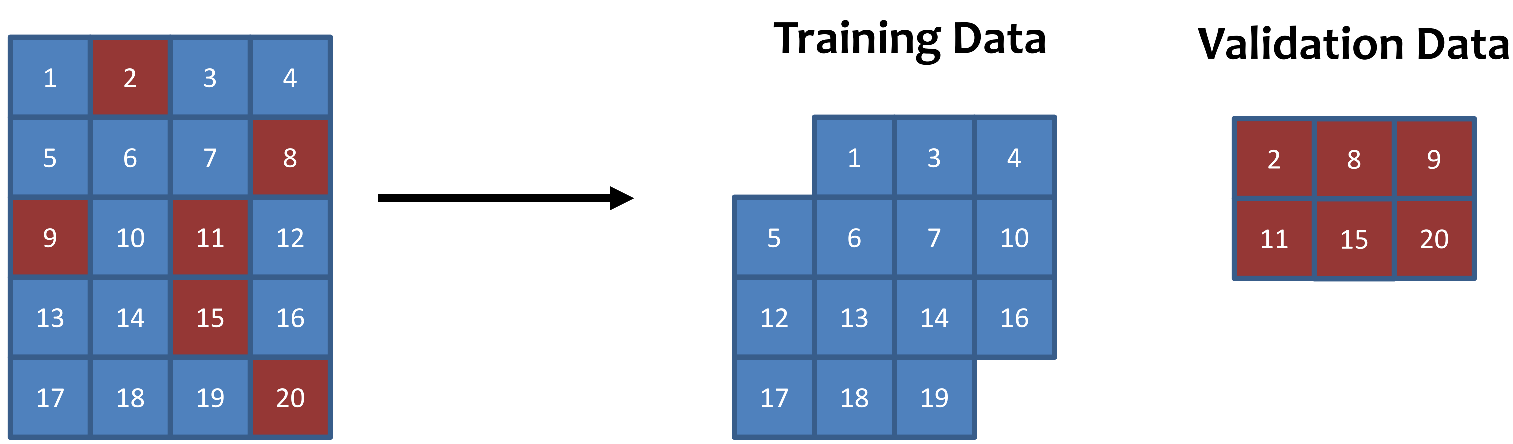
…and because you should.

Model Validation

Hold-out datasets

Used two methods –

- Out of sample: randomly trained on 70% of data; validated against remaining 30% of data.

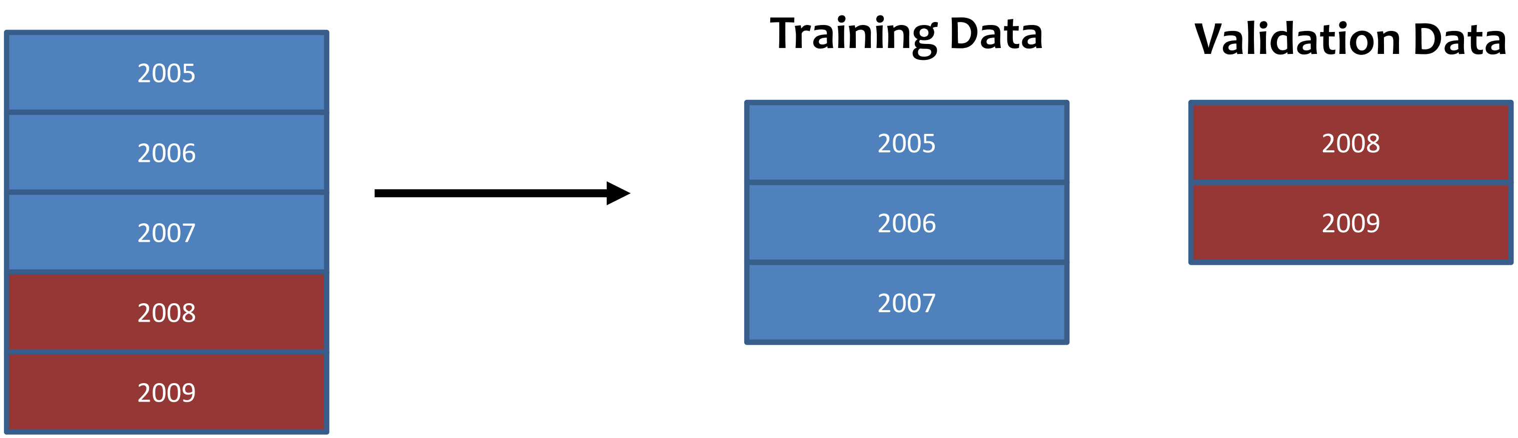


Model Validation

Hold-out datasets

Used two methods –

- Out of sample: randomly trained on 70% of data; validated against remaining 30% of data.
- Out of time: trained against older years of data; validated against newest years of data.



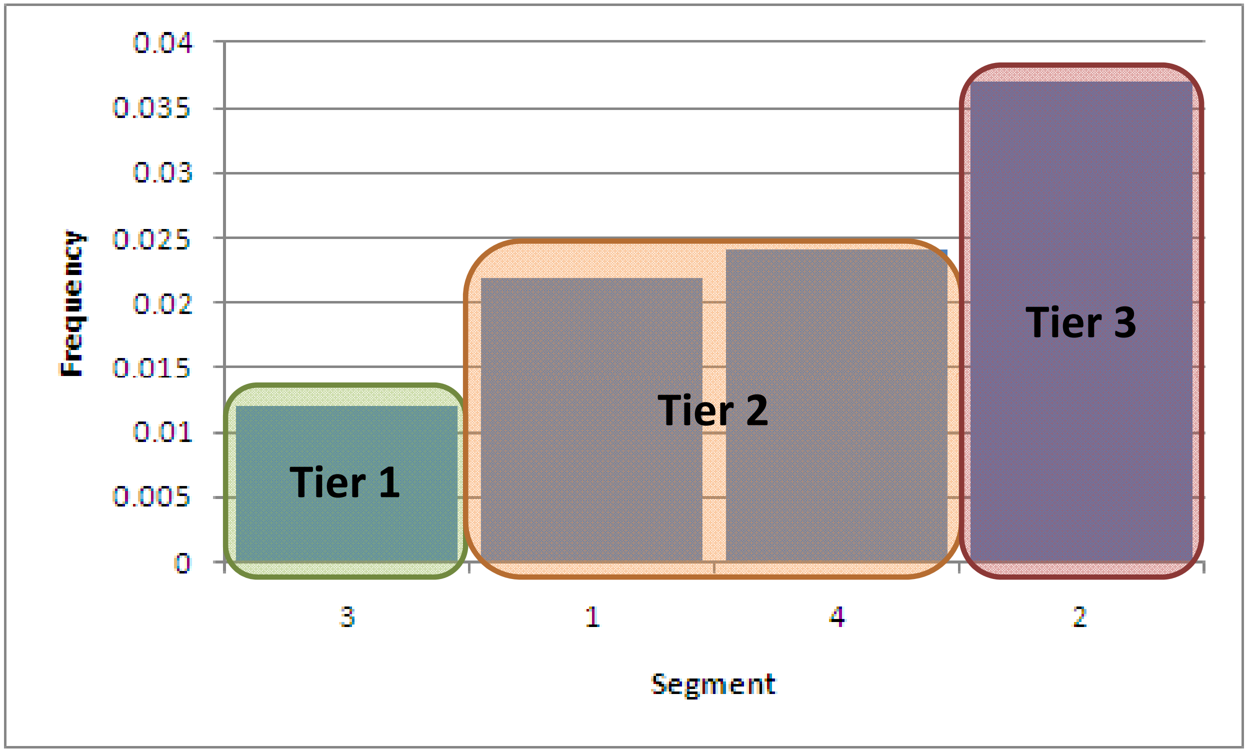
4. *Non-rating Uses for Machine Learning*

Non-rating Uses for Machine Learning

Underwriting Tiers and Company Placement

Target frequency at the policy level

Define tiers based on similar frequency characteristics.



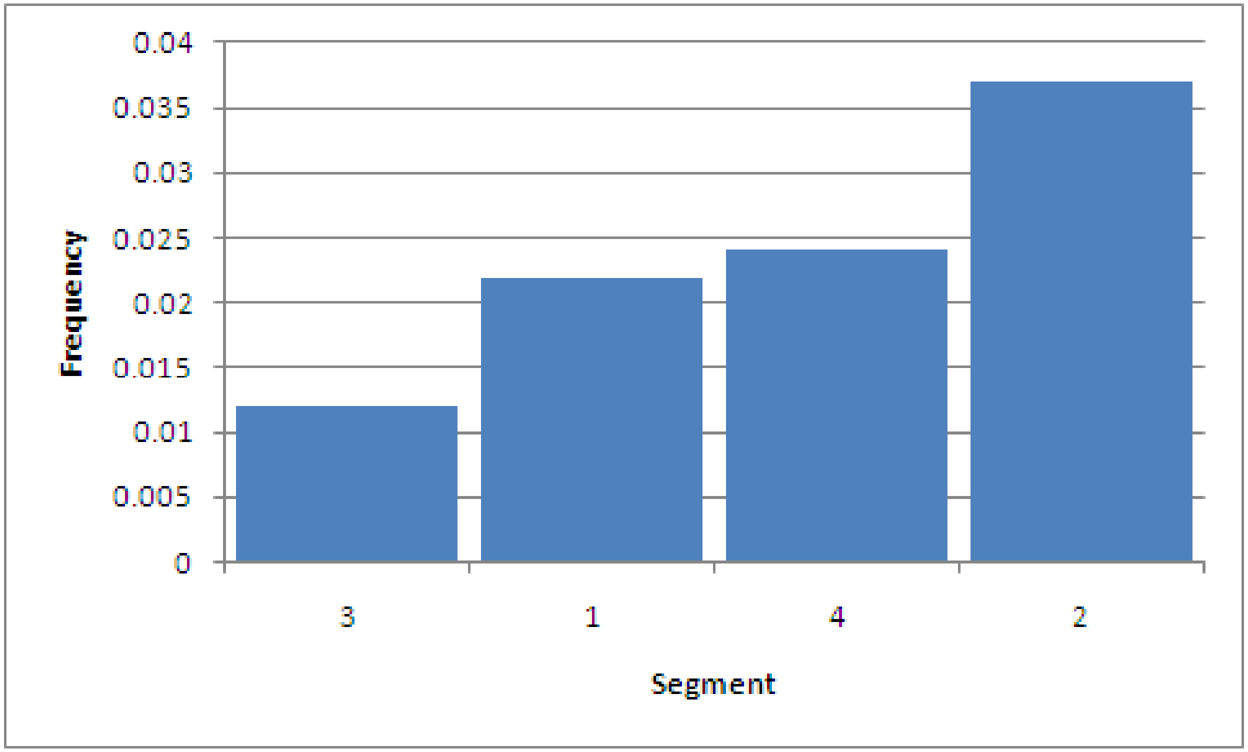
Note that a project like this would need to be done in conjunction with pricing. This sorting of data occurs prior to rating and would need to be accounted for.

Non-rating Uses for Machine Learning

Straight-thru
versus
Expert UW

Target frequency or loss ratio at the policy level

Consider policy performance versus current level of UW scrutiny.



Do not forget that current practices affect the frequency and loss ratio of your historical business. Results like this may indicate modifications to current practices.

Non-rating Uses for Machine Learning

"I have the budget to re-underwrite 10% of my book. I just need to know which 10% to look at!"

With any project of this sort, the level of the analysis should reflect the level at which the decision is made, and the target should reflect the basis of your decision.

In this case, we are making the decision to re-underwrite a given *POLICY*. Do the analysis at the policy level. *(Re-inspection of buildings may be done at the unit level.)*

To re-underwrite unprofitable policies, use loss ratio as the target.

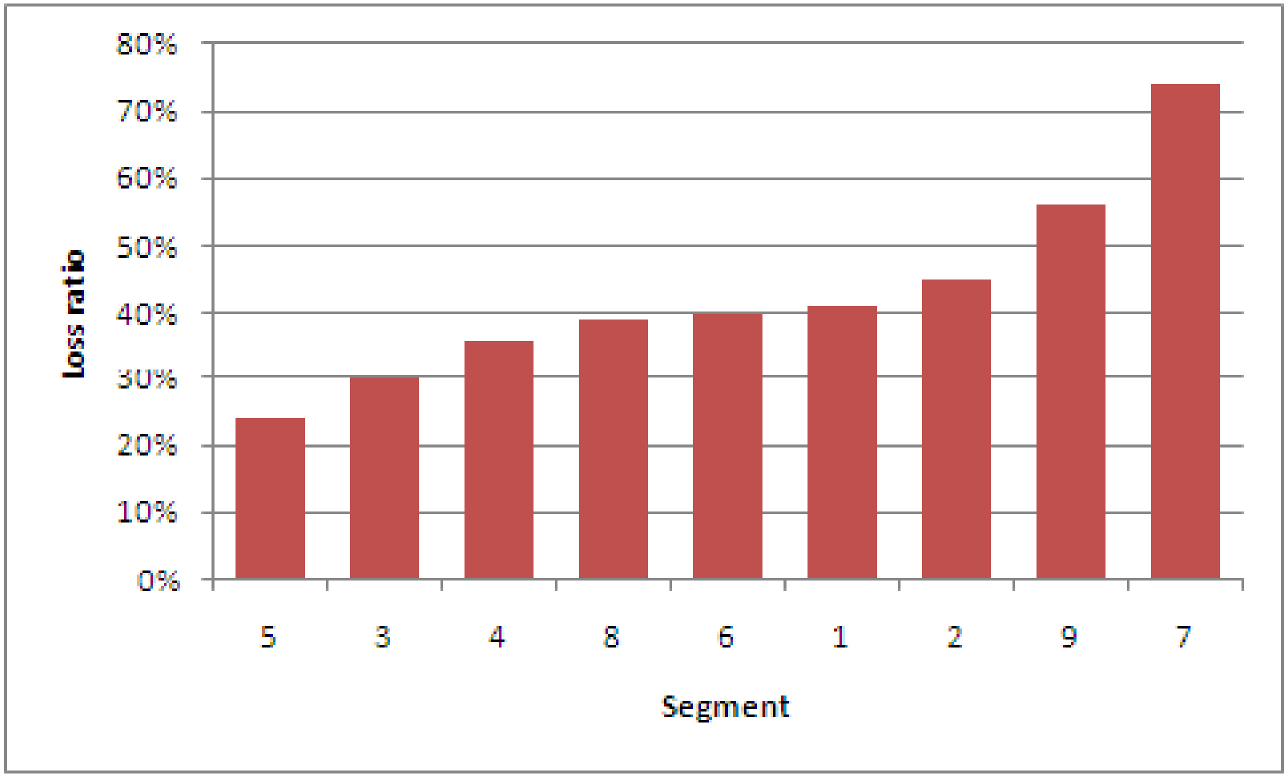
Note: when using loss ratio, be sure to current-level premium at the policy level (not in aggregate).

Non-rating Uses for Machine Learning

Re-underwrite
or
Re-inspect

Target loss ratio at the policy level

Depending on the size of the program, target segments 7 & 9 as unprofitable.



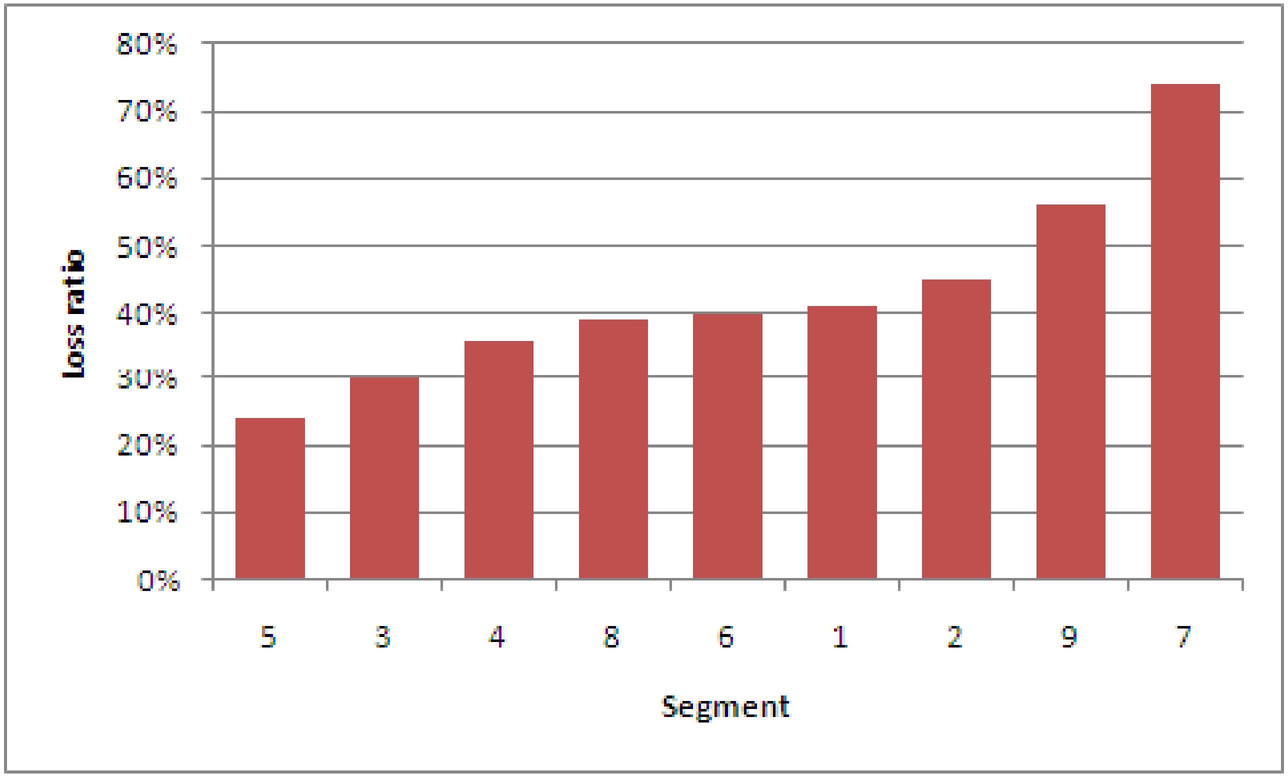
If the analysis data is current enough, and if in-force policies can be identified, this kind of analysis can result in a list of policies to target rather than just the attributes that correspond with unprofitable policies (segments 7 & 9).

Non-rating Uses for Machine Learning

<u>Profitability – reduce the bad</u>

Target loss ratio at the policy level

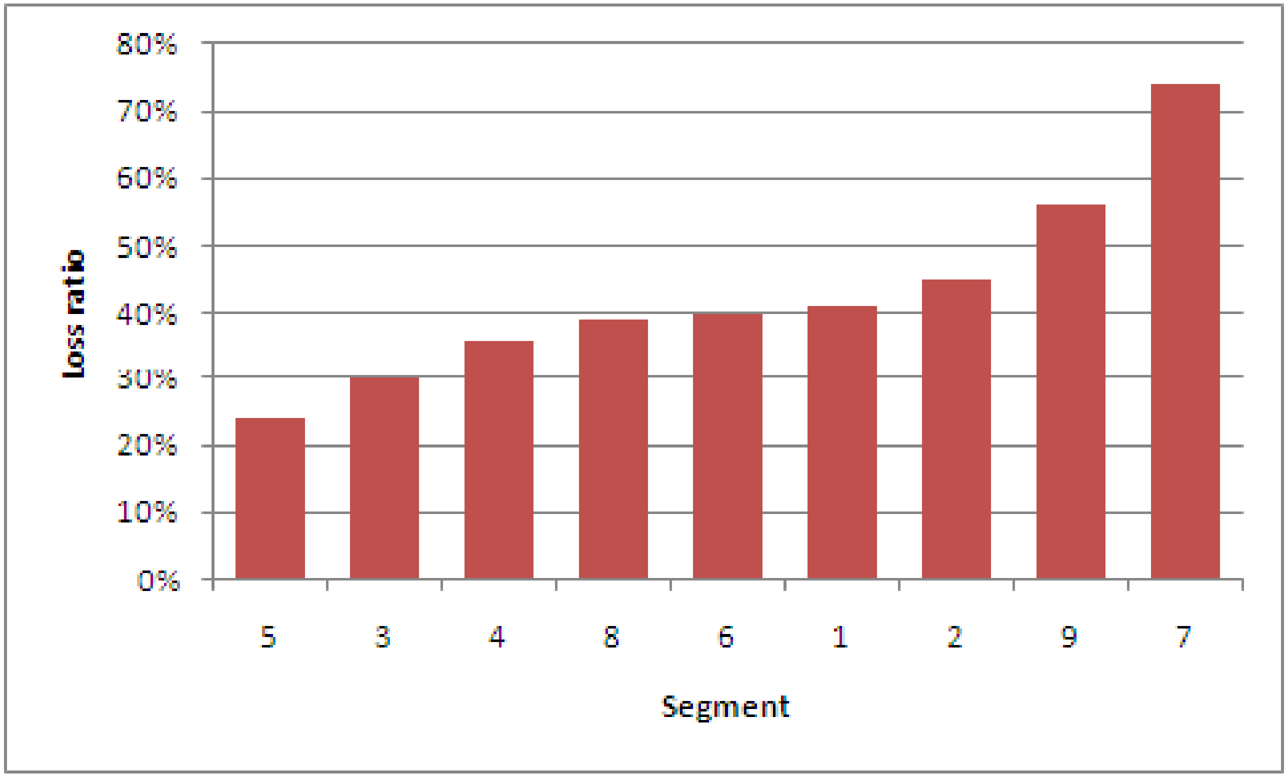
Reduce the size of segment 7 – consider non-renewals and/or the amount of new business.



There is a range of aggressiveness here which may also be affected by the regulatory environment.

Non-rating Uses for Machine Learning

Profitability – increase the good (target marketing)



Target loss ratio at the policy level

If the attributes of segment 5 define profit-able business, get more of it.

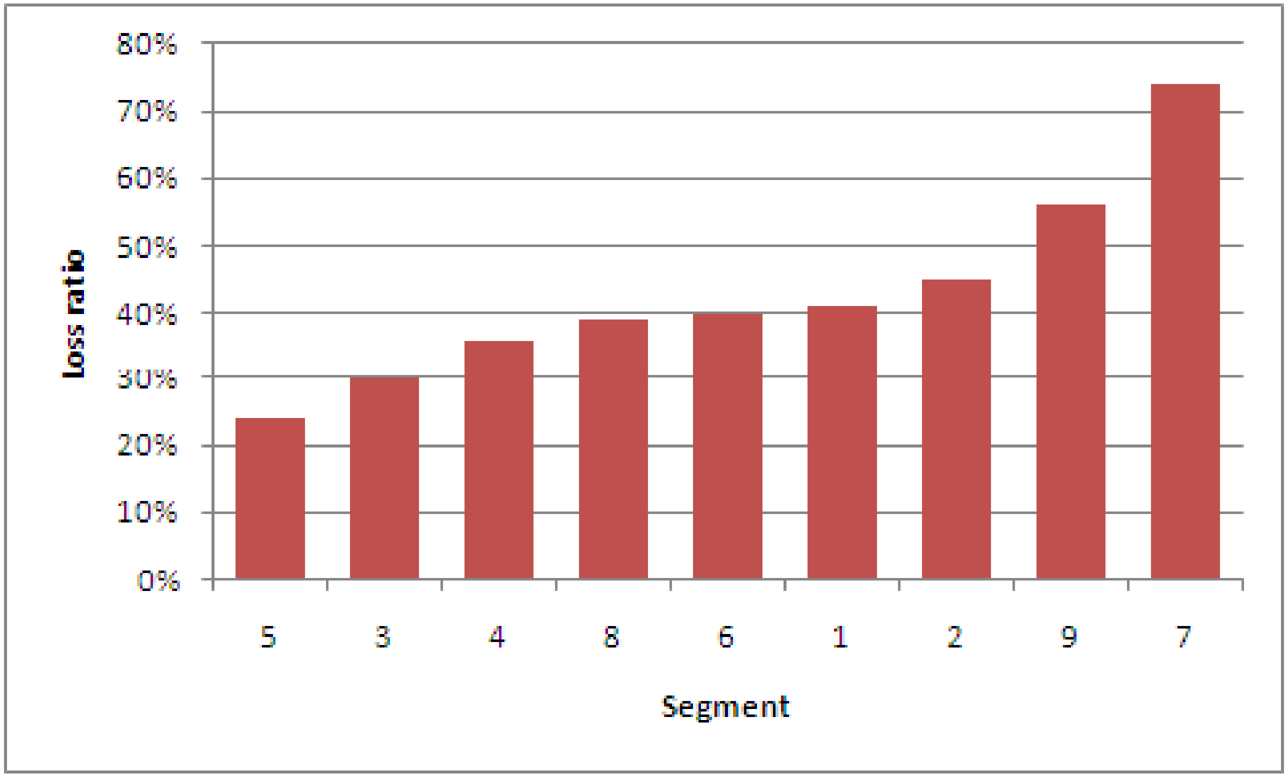
This kind of analysis defines the kind of business you write profitably. This needs to be combined with marketing/demographic data to identify areas rich in this kind of business. Results may drive agent placement or marketing.

Non-rating Uses for Machine Learning

Quality of Business

Target loss ratio at the policy level

Knowing who you write at a profit and loss, you can monitor new business as it comes in.



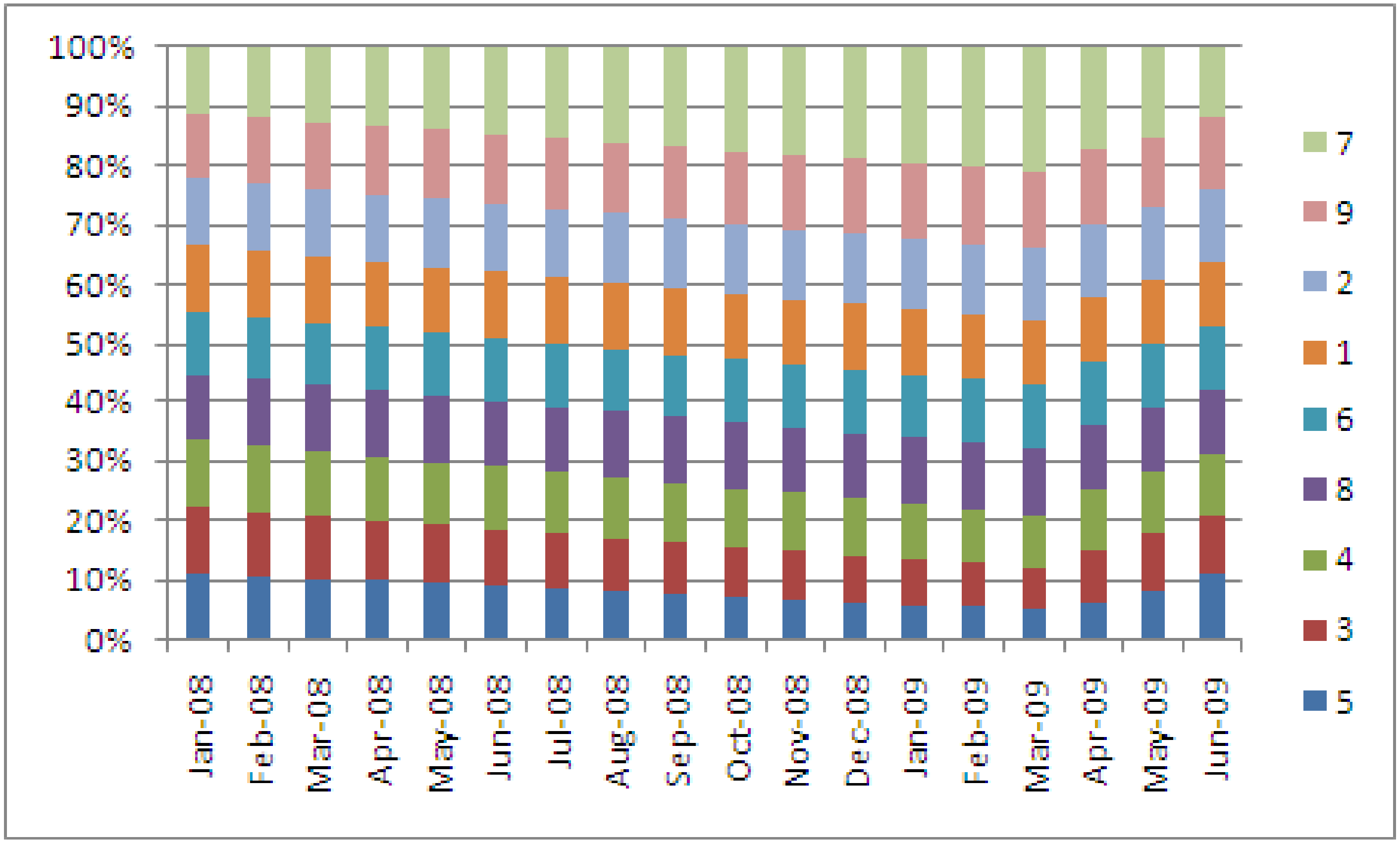
Monitor trends over time to assess the adverse selection against your company. Estimate the effectiveness of underwriting actions to change your mix of business.

Non-rating Uses for Machine Learning

Quality of Business

Here you can see adverse selection occurring through March 2009.

Company action at that point reversed the trend.



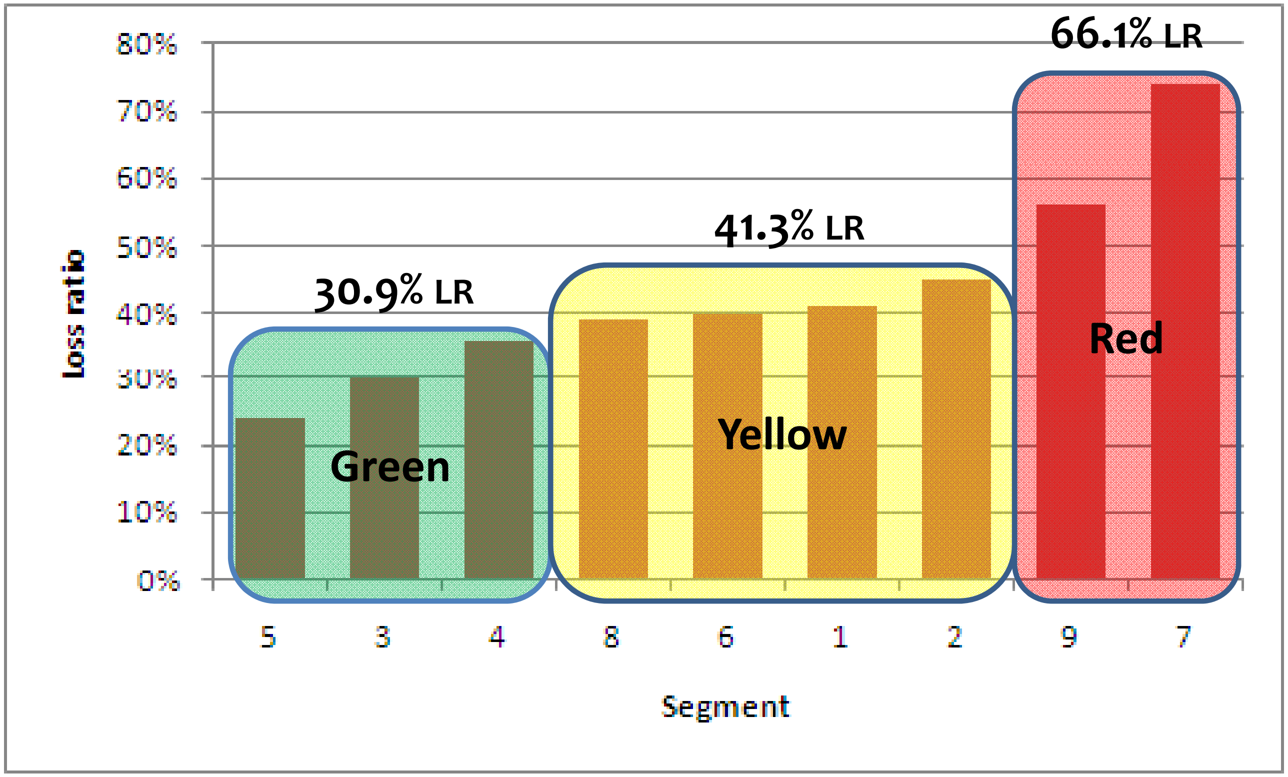
This looks at the total business of the book. Can also focus exclusively on new business.

Non-rating Uses for Machine Learning

Agent/broker Relationship

Target loss ratio at the policy level

Use this analysis to inform your understanding of agent performance.



Actual agent loss ratios are often volatile due to smaller volume. How can you reward or limit agents based on this? A loss ratio analysis can help you understand *EXPECTED* performance as well as actual.

Non-rating Uses for Machine Learning

Agent/broker Relationship

More profitable than expected…

This agent writes yellow and red business better than expected.

<u>Best practices</u> – is there something this agent does that others should be doing?

Agent xxxxx				
Group	Exposures	Earned Premium	Actual Loss Ratio	Expected Loss Ratio
Green	1,644	1,395,788	31.1%	30.9%
Yellow	3,381	2,763,714	34.5%	41.3%
Red	3,085	2,559,968	42.0%	66.1%
			36.7%	47.0%

<u>Getting lucky</u> – is this agent living on borrowed time? Have the conversation to share this info with the agent.

Non-rating Uses for Machine Learning

Agent/broker Relationship

Less profitable than expected…

This agent writes all business worse than expected.

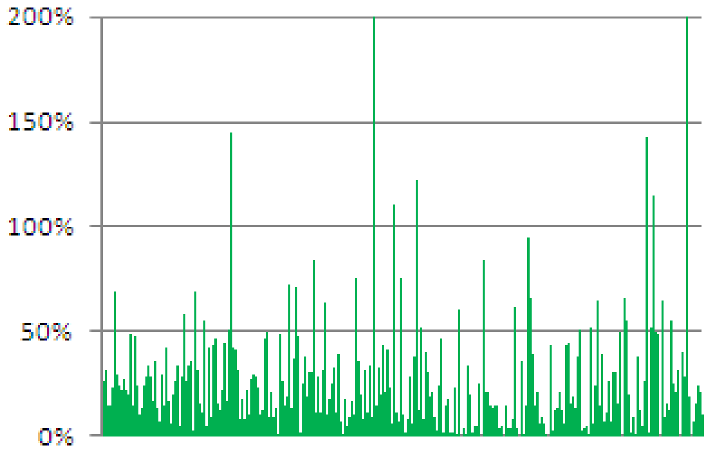
Worst practices – is this agent skipping inspections or not following UW rules?

Agent xxxxx				
Group	Exposures	Earned Premium	Actual Loss Ratio	Expected Loss Ratio
Green	1,888	1,211,599	47.8%	30.9%
Yellow	1,628	1,144,790	55.7%	41.3%
Red	478	355,295	82.5%	66.1%
			55.7%	**47.0%**

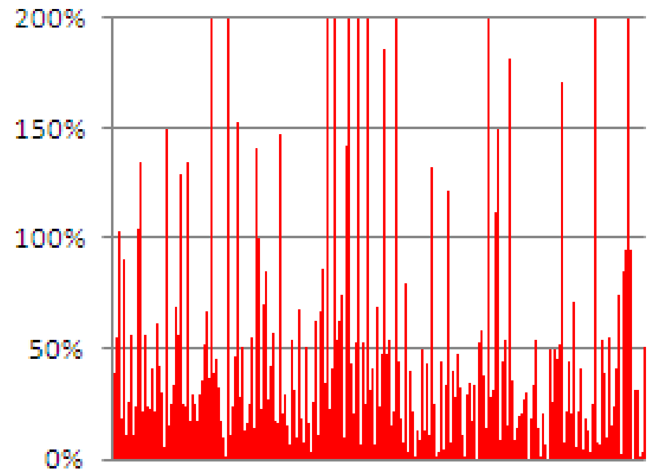
Getting unlucky – This agent doesn't write much red business. Maybe they are given more time because their mix of business should give good results over time.

Non-rating Uses for Machine Learning

Agent/broker Relationship



Agents with the most Red Business
Not only is the underlying loss ratio higher, but the odds of that big loss is much higher too.



Agents with the most Green Business
Some of these agents who write large amounts of low-risk business get unlucky, but the odds are good that they'll be profitable.

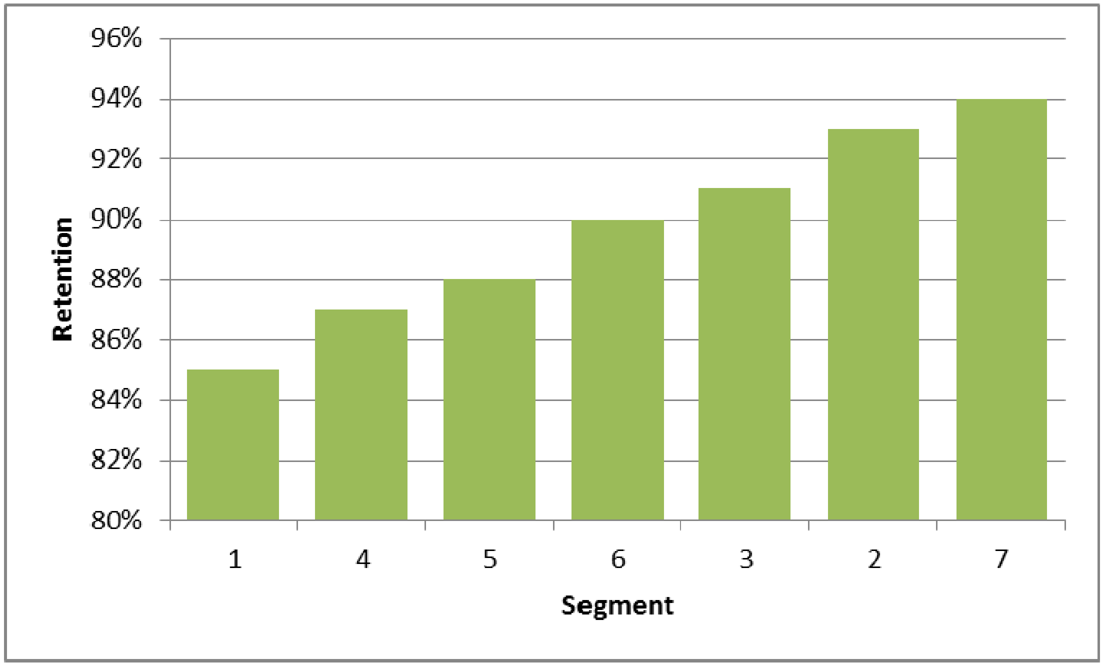
Non-rating Uses for Machine Learning

Retention Analyses

Target retention at the policy level

What are the common characteristics of those with high retention (segment 7)?

This information can be used in a variety of ways…



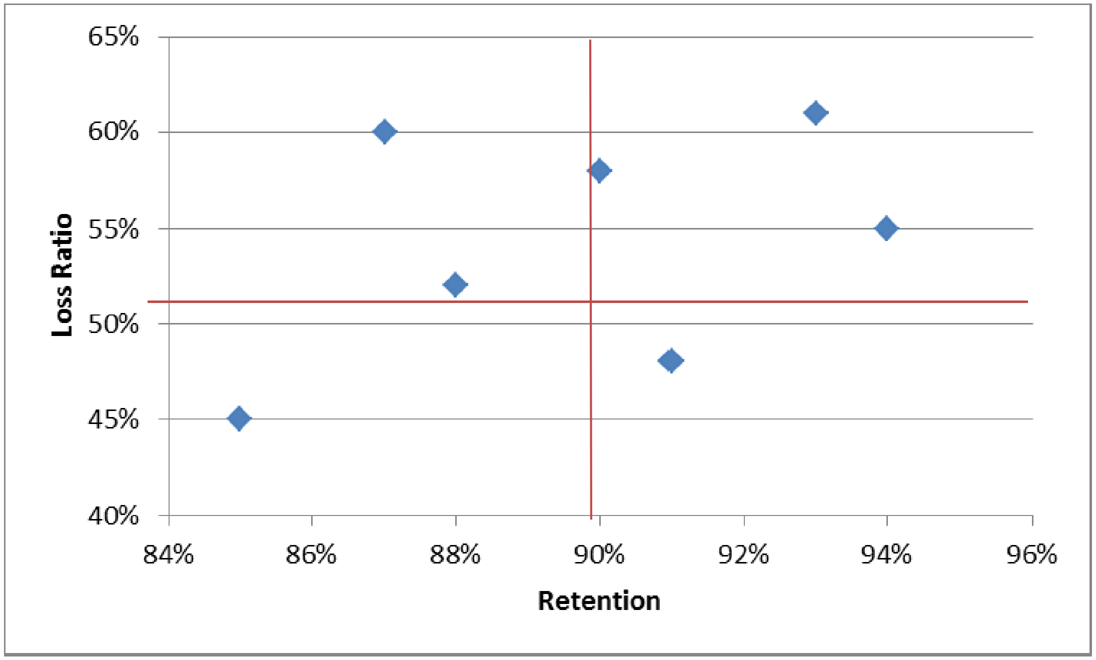
- Guide marketing & sales towards customers with higher retention
- Form the basis of a more formal lifetime value analysis
- Cross-reference retention and loss ratio to get a more useful look…

Non-rating Uses for Machine Learning

Retention Analyses

Simple looks at retention can be even more useful when cross-referenced with loss ratio.

Is a segment of business above or below average retention? Above or below the target loss ratio?



Note: retention is essentially a static look at your book. What kinds of customers retained? What kinds didn't? There is no consideration of the choice customers had at renewal. Were they facing a rate change and renewed anyway?

5.

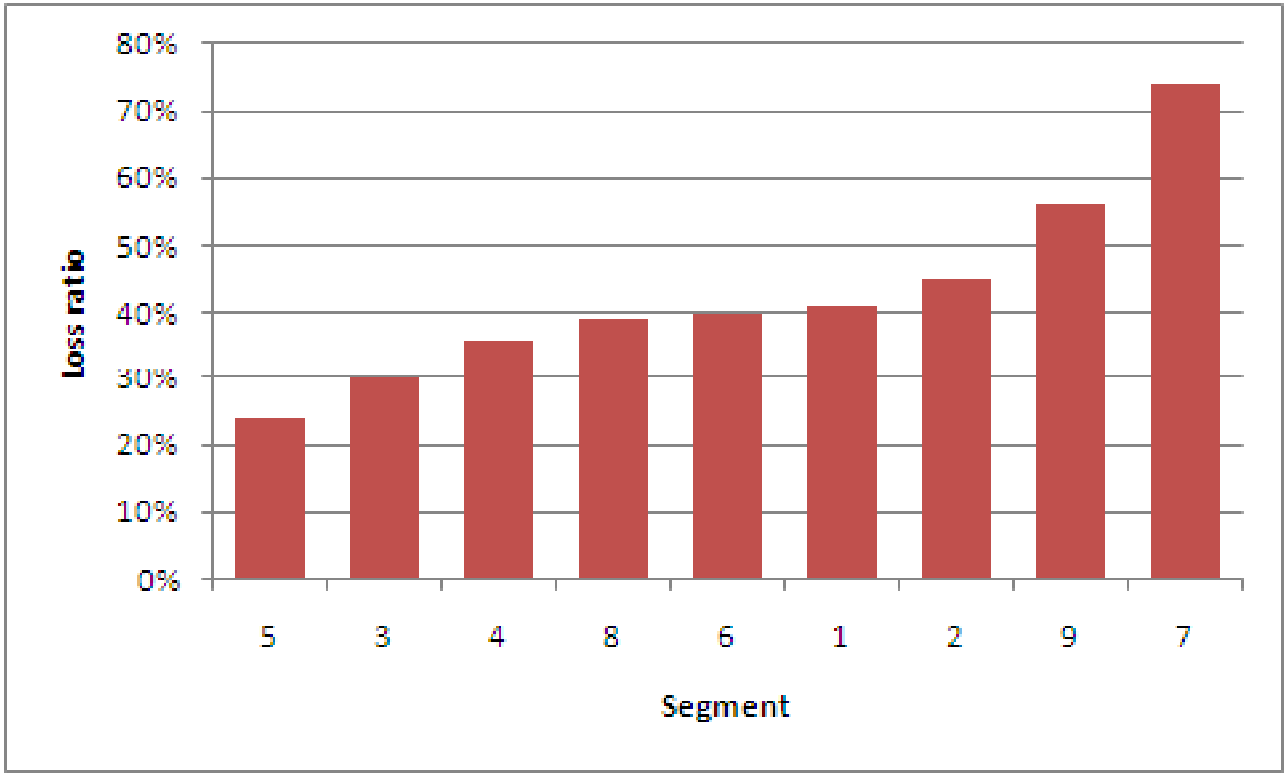
Rating Applications of Machine Learning

Rating Applications of Machine Learning

<u>**The Quick Fix**</u>

Target loss ratio at the coverage level

The lift curve is easily translated into relativities which can even out your rating.



Note that the quickest fix to profitability is taking underwriting action. But the quickest fix for rating is to add a correction to existing rates. This can be done because loss ratio shows results ***given the current rating plan.***

The Quick Fix

Segments	Exposures	Premium	Loss Ratio	Relativity	Rel (base 6)
5	9,320	1,043,894	24.0%	0.513	0.600
3	12,042	1,709,934	30.0%	0.641	0.750
4	14,763	1,446,784	36.0%	0.769	0.900
8	17,484	1,643,534	39.0%	0.833	0.975
6	17,484	1,835,863	40.0%	0.855	1.000
1	17,484	1,923,285	41.0%	0.876	1.025
2	18,845	2,336,788	45.0%	0.962	1.125
9	20,206	1,818,514	57.0%	1.218	1.425
7	31,114	3,578,067	72.0%	1.539	1.800
Total	**158,743**	**17,336,663**	**46.8%**	**1.000**	

# of Units	Cov Limit	# of Insured	...	Relativity
1	na	na	...	1.025
>1	>10000	na	...	1.125
>1	<=10000	1,2	...	0.750
>1	<=10000	>2	...	0.900
...	...	...	...	...

First determine relativities based on the analysis loss ratios.

Then create a table which assigns relativities.

Note that this can be one table as shown, or it can be two tables: one which assigns the segments and one which connects segments to relativities. The exact form will depend on your system.

Creating a class plan from scratch

Machine Learning algorithms, such as decision trees, can be used to create class plans rather than just to modify them. However, they will not look like any class plan we are used to using.

"An 18 year old driver in a 2004 Honda Civic, that qualifies for defensive driver, has no violations but one accident, with a credit score of 652, who lives in territory 5 and has been with the company for 1 year, who has no other vehicles on the policy nor has a homeowners policy, who uses the vehicle for work, is unmarried and female, and has chosen BI limits of 25/50 falls in segment 195 which has a rate of $215.50."

Traditional statistical techniques, such as Generalized Linear Models, are more appropriate for this task. However, the process of creating a GLM model can be supplemented using decision trees or other Machine Learning techniques.

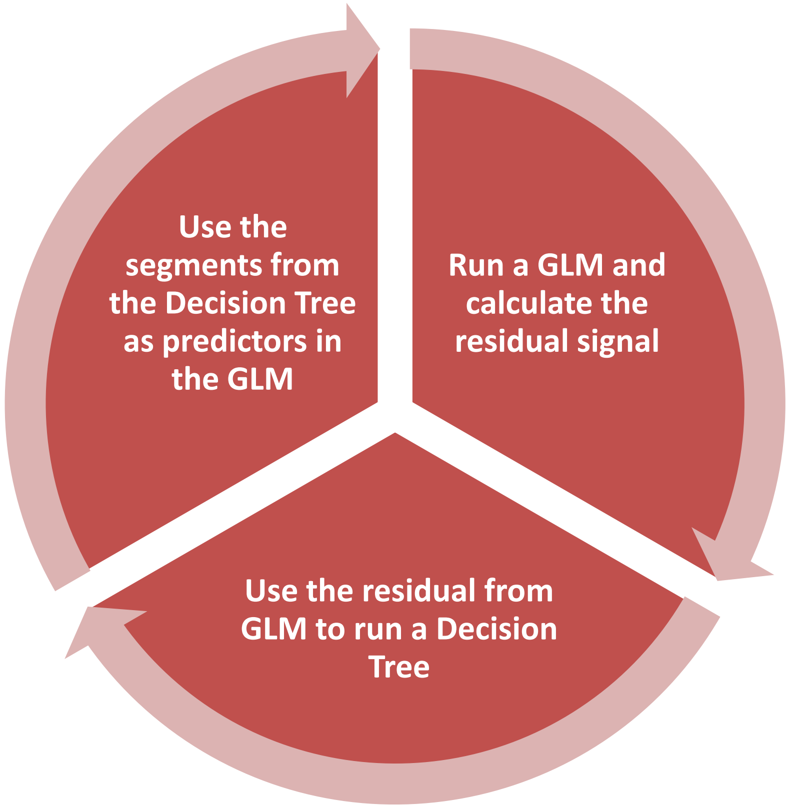
Rating Applications of Machine Learning

Creating a class plan from scratch

Disadvantages of GLMs alone	Advantages of combining GLMs and Machine Learning
Linear by definition	Machine Learning can explore the non-linear effects
Parametric – requires the assumption of error functions	Supplements with an alternate approach which make no such assumption
Interactions are "global" – they apply to all the data if used	Decision trees find "local" interactions by definition
Trial and error approach to evaluating predictors – only a small portion of all possible interactions can be explored, given real-world resources and time constraints	Machine Learning explores interactive, non-linear parts of the signal in an automated, fast manner

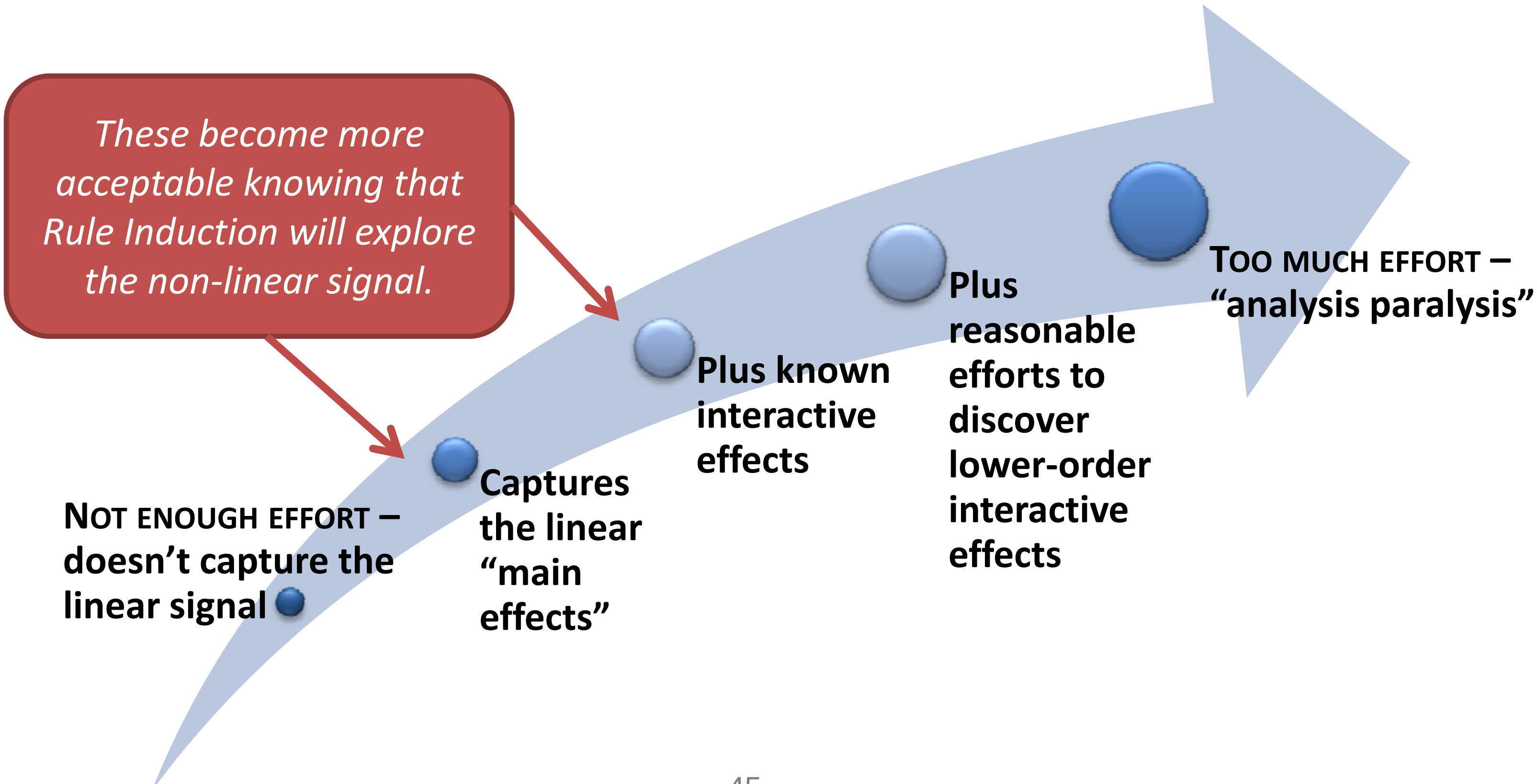
Creating a class plan from scratch

Using Machine Learning and GLMs together…



Second way to "enhance" GLMs – rebalance the workload

The first place to look is in how much effort is put into building the initial GLM.



6. *Analysis of high dimensional variables*

Analysis of high dimensional variables

<u>High Dimensional Variables</u>

Geographic and vehicle information are classic examples of predictors with many, many levels.

- Geographic building blocks of Territories are usually county/zip code combinations, zip code, census track, or lat/long.
- Vehicle building blocks of Rate Symbols are usually VINs.

In both cases, you cannot simply plug the building blocks into a GLM; the data is too sparse. You need to group "like" levels in order to reduce the total number of levels. In other words, you need to find Territory Groups or Rate Symbol Groups.

Note: once grouped, you should use a GLM to determine rate relativities. This ensures that these parts of the class plan are in sync with the others.

Analysis of high dimensional variables

High Dimensional Variables

Current analytical approaches for geography use some form of distance in order to smooth the data, providing estimates of risk for levels with little to no data.

Once each building block has a credible estimate of risk, levels with similar risk are clustered together into groups.

Issues with this approach:

- What is the measure of risk to be smoothed?
- What distance measure should be used?
- What smoothing process & how much smoothing?
- What clustering process & how many clusters?

Analysis of high dimensional variables

High Dimensional Variables

Tree-based approaches, a form of **rule induction**, provide a simpler alternative.

Geographic proxies are attached to the data.

- Census/demographic data
- Weather data
- Retail data
- Etc.

Branches of the tree define territories…

Segment 1 = Territory 1 = all zip codes where rainfall > 0.1 and popdensity < 0.5

Zip codes with little data will not drive the analysis, but will get assigned to groups. No need for smoothing.

Analysis of high dimensional variables

High Dimensional Variables

Eliade Micu presented a direct comparison between these two approaches: smoothing/clustering versus rule induction.

He found quite similar results, though his version of rule induction did outperform his version of smoothing/clustering.

This presentation can be found on-line at the CAS Website:

Seminar Presentations of the 2011 RPM Seminar
Session PM-10: Territorial Ratemaking (Presentation 2)
http://www.casact.org/education/rpm/2011/handouts/PM10-Micu.pdf

Extension of smoothing/clustering to vehicle information can be problematic. What is "distance"? What are "like" VINs? However rule induction can be applied to vehicle information in an exactly analogous manner.

Summary

- **The more accessible Machine Learning techniques, such as decision trees, can be used today to enhance insurance operations.**

- **Machine Learning results are not too complicated to use in insurance.**

- **Non-rating applications of Machine Learning span underwriting, marketing, product management, and executive-level functions.**

- **Actuaries should pursue the business goal most beneficial to the company – this may include some of these non-rating applications.**

- **Rating applications of Machine Learning include both quick fixes and fundamental restructuring of rating algorithms.**

- **Rule induction has intriguing applications to analyzing high dimensional variables.**

Questions?

Contact Info

Christopher Cooksey, FCAS, MAAA
EagleEye Analytics
ccooksey@eeanalytics.com
www.eeanalytics.com