Practical Applications of Predictive Analytics

Case Study of Claim Termination Assumption

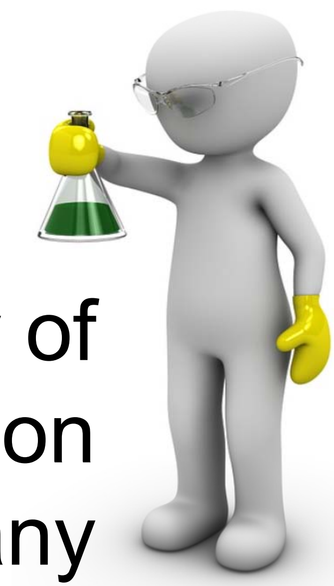
Missy Gordon, FSA, MAAA
Principal and Consulting Actuary
Milliman, Minneapolis

ILTCI

Introduction

How can we develop LTC assumptions that predict well for many years into the future?

Can predictive analytics help?



Explored case study of claim termination assumption for one company

Agenda

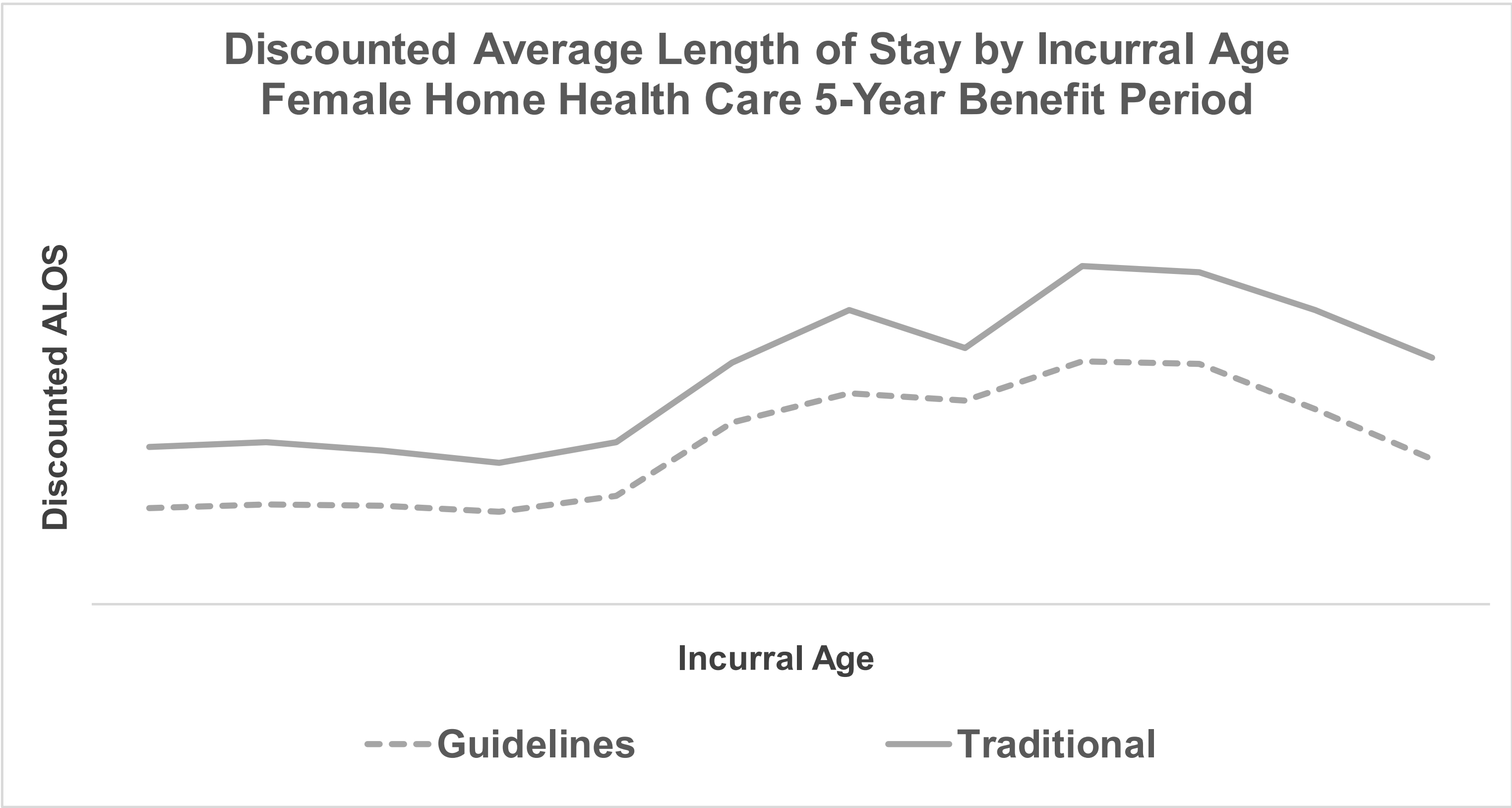
- Traditional method and its challenges
- Two predictive methods as solutions
  - – Build comfort by understanding similarities
  - – Expand solution with advanced methods
- What challenges remain?

Disclaimer: The views and opinions expressed in this presentation are those of the presenter and do not reflect the official policy or position of Milliman, Inc.

Traditional method

- Starting expectation of claim terminations
- A:E adjustments with judgement
  - Amount of weight to give data
  - Variable selection and interactions
- Enhanced to adjust and re-normalize iteratively

Traditional method



Traditional method challenges

- Effort and cumbersome

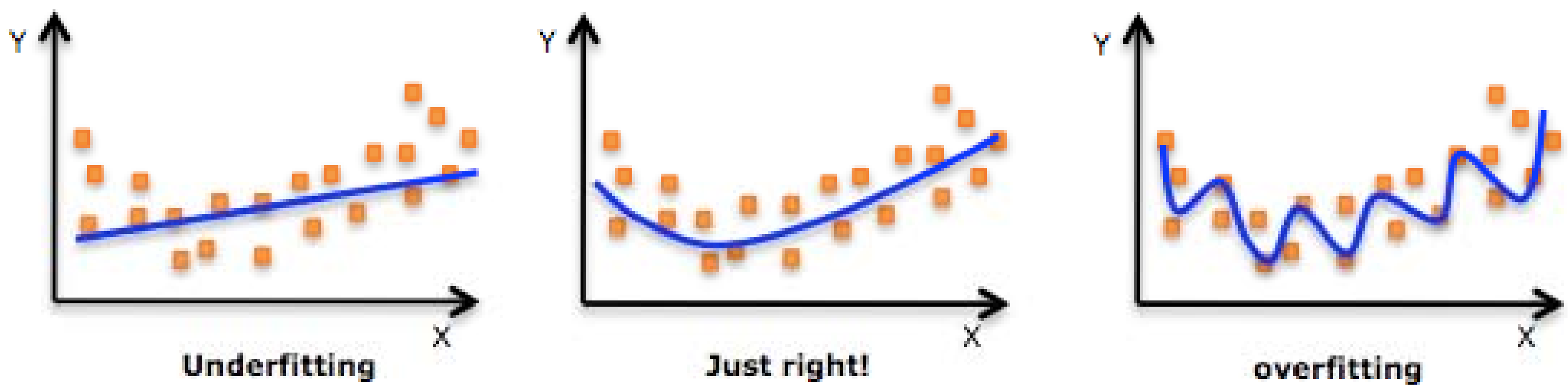
Judgement decisions:

- Variable selection
- Interactions/slices
- Weight given to data

- Does not tell us if works on unseen data

Traditional method biggest challenge

Traversing the bias-variance tradeoff



High bias
Low variance
Low data weight

High variance
Low bias
High data weight

Can predictive analytics help?

Two possibilities out of many methods:

1. Penalized GLM
   - Similarities to traditional
   - Addresses most challenges

2. GBM
   - Addresses remaining challenges
   - Powerful predictor, but less user control
   - Can be used to supplement penalized GLM

Penalized GLM

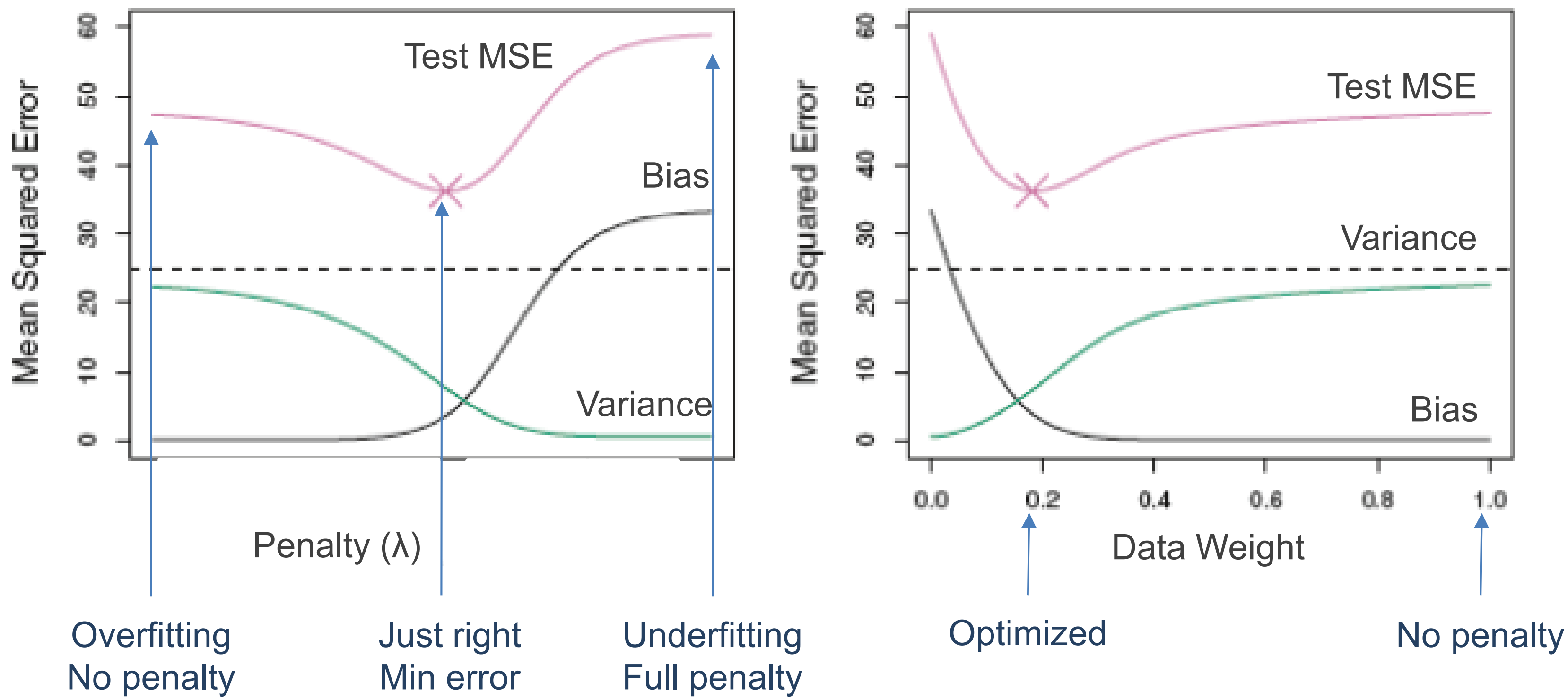
How does it work?

- Develops coefficients using GLM with E offset
  - Similar to A:E adjustments
- Penalizes coefficients to control overfitting
  - Determines amount of weight to give data
- Chooses penalty to minimize cross-validation error
  - Traverses bias-variance tradeoff
  - Produces better predictions on unseen data

Penalized GLM a.k.a., Generalized linear model with regularization

Penalized GLM

- Penalty minimizes cross-validation error
- Determines data weight



Penalized GLM

K-fold cross-validation

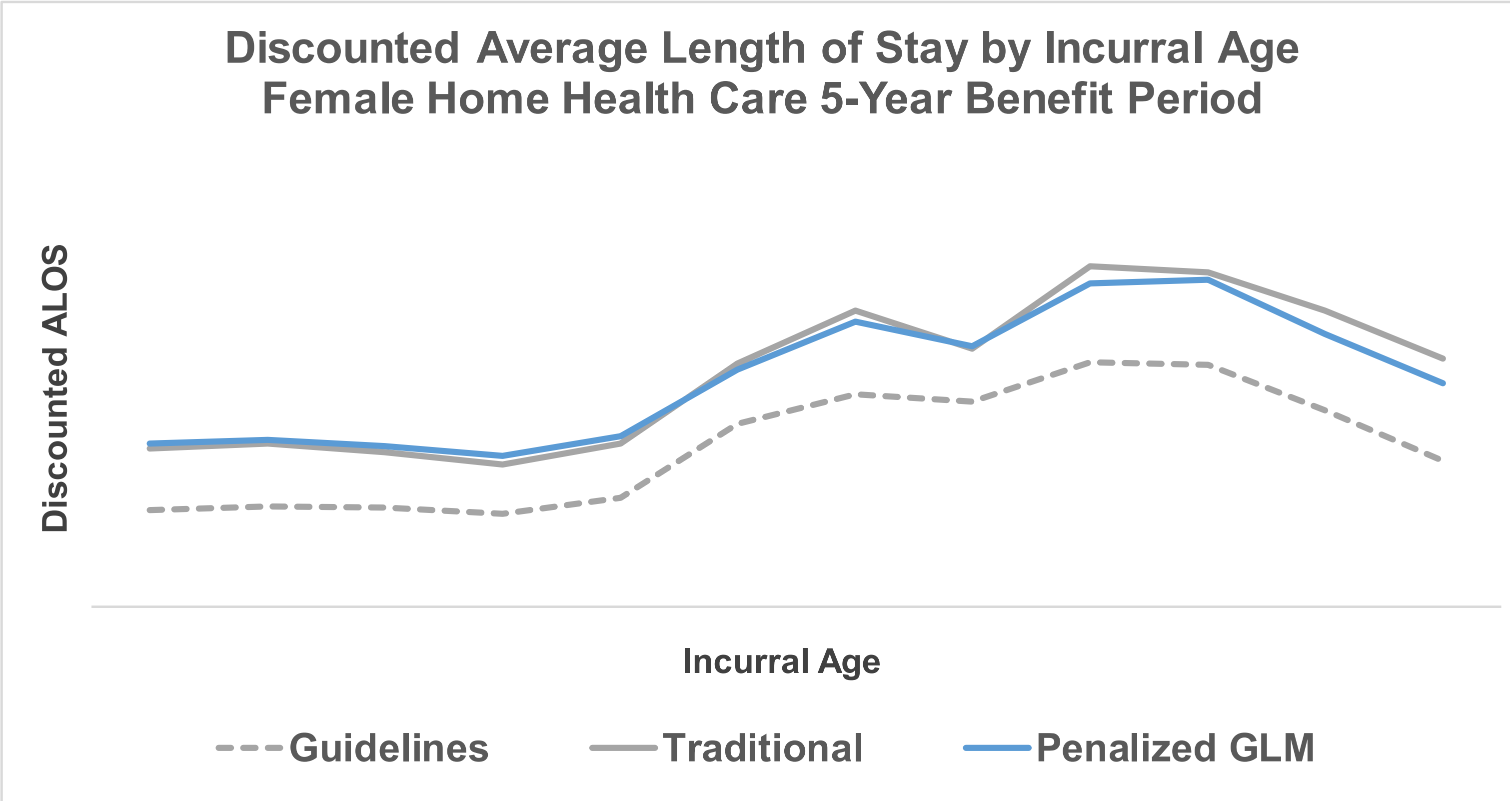
– Calibrates model using training data

– Tests how well model predicts unseen data

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Testing	Training	Training	Training	Training
Training	Testing	Training	Training	Training
Training	Training	Testing	Training	Training
Training	Training	Training	Testing	Training
Training	Training	Training	Training	Testing

Prediction Error Statistics	Average
	Testing
	Testing
	Testing
	Testing
	Testing

Penalized GLM



Future Profit Margin
15.5% Traditional
16.2% Penalized GLM
0.7% Impact

- Apples-to-apples, except method
- Penalized GLM gave less weight to data

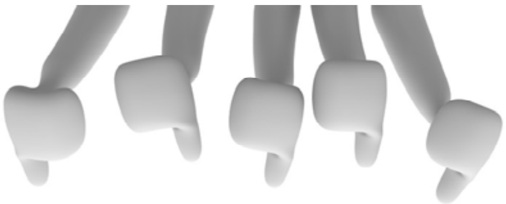
Penalized GLM

Why do we like it?

- Automates bias-variance tradeoff
  - Choice of data weight
  - Tests prediction on unseen data
- Efficient to update/modify

- Similarities to traditional
- Reduces human error

Penalized GLM

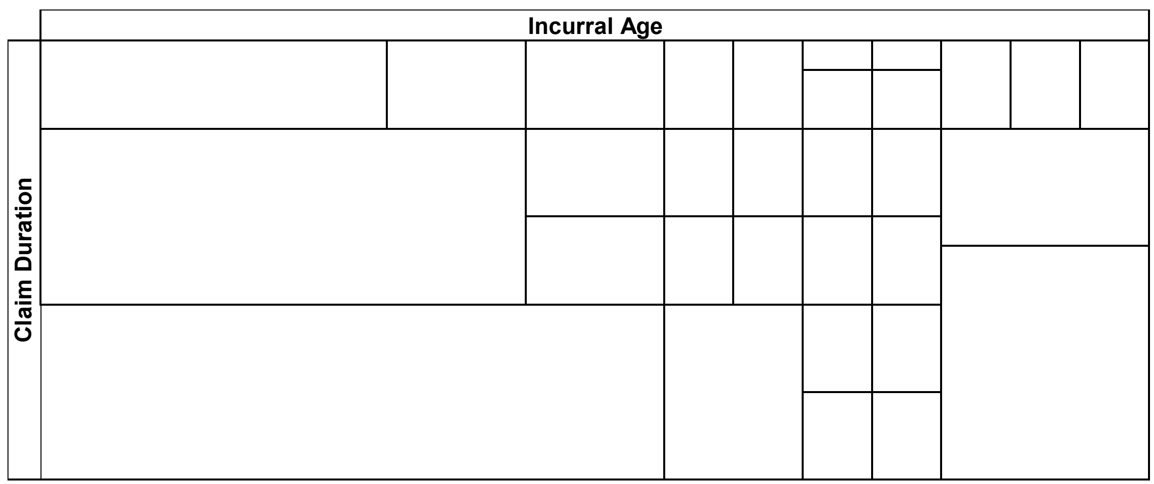


What challenge remains?

Navigating complex interactions

- What are the key interactions?
- How do we slice the data?
- Do the slices vary with interactions?
- Are the adjustments similar enough to keep rolled-up?
- Is there enough data in a slice?

If only there was something to help…



GBM

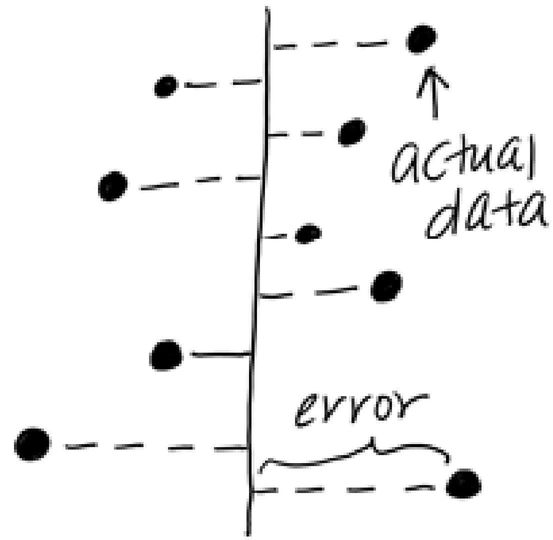
How does it work?

- Develops A:E adjustments
- Builds layers of decision trees to minimize error
  - Slices data to create variable buckets
  - Finds complex interactions
- Control for overfitting using cross-validation

Gradient Boosting Machine (GBM)

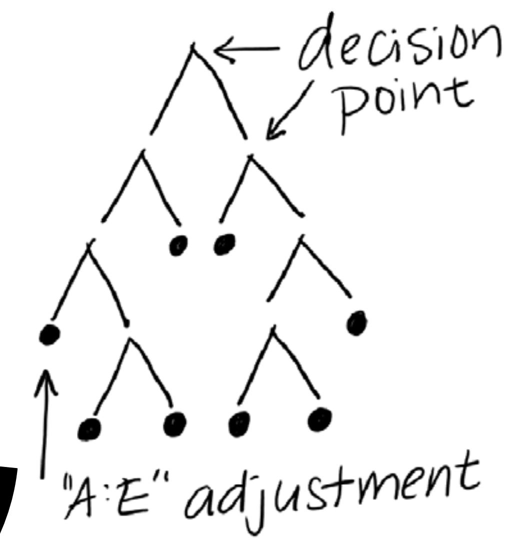
GBM

1. Start with base E



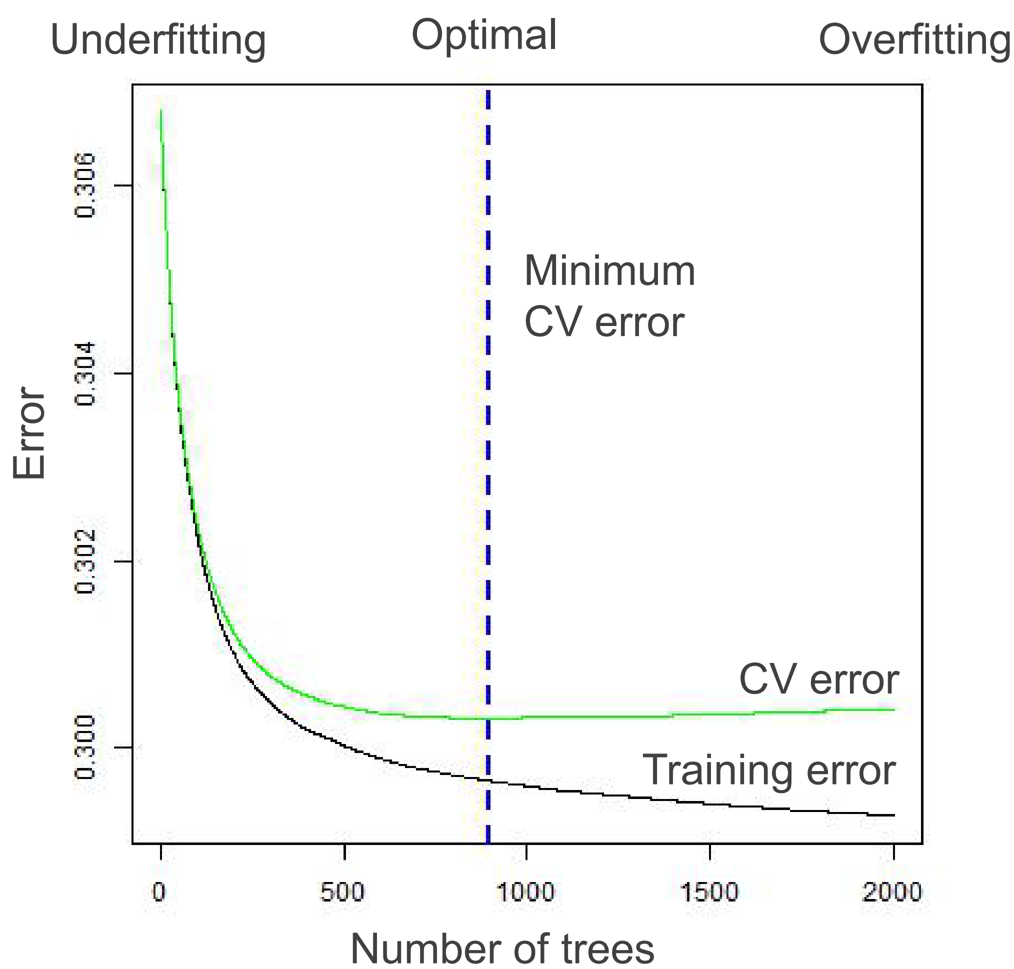
2. Subsample and calculate error

3. Tree develops A:E adjustments to minimize error



4. Calculate new "E" by layering A:E adjustments and repeat…. ~1,000 times

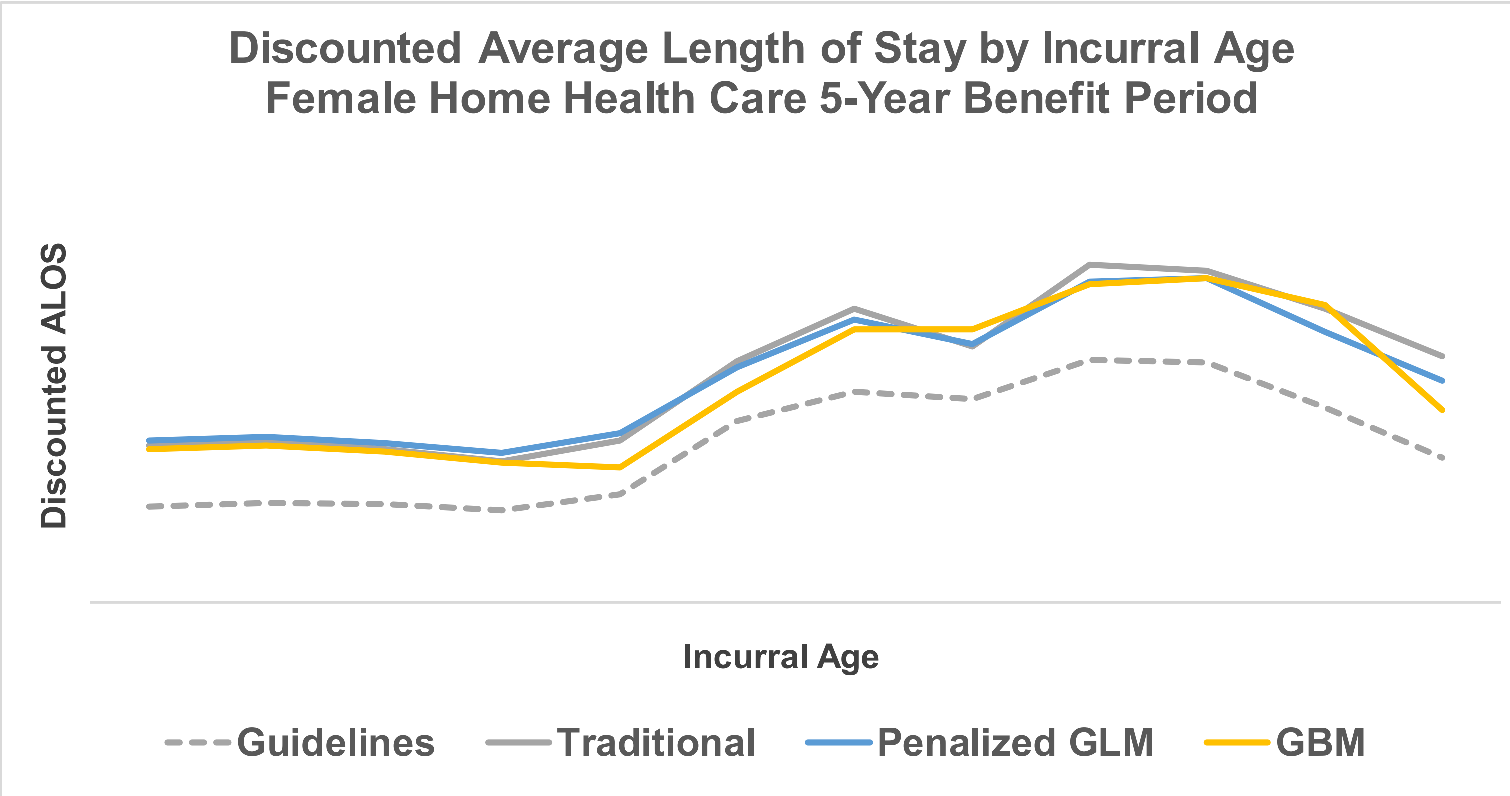
Control overfitting using cross-validation



Key hyperparameters

- Sub-sampling
- Shrinkage
- Interaction depth
- Minimum observations
- Number of trees

GBM



Future Profit Margin
15.5% Traditional
15.2% GBM
-0.3% Impact

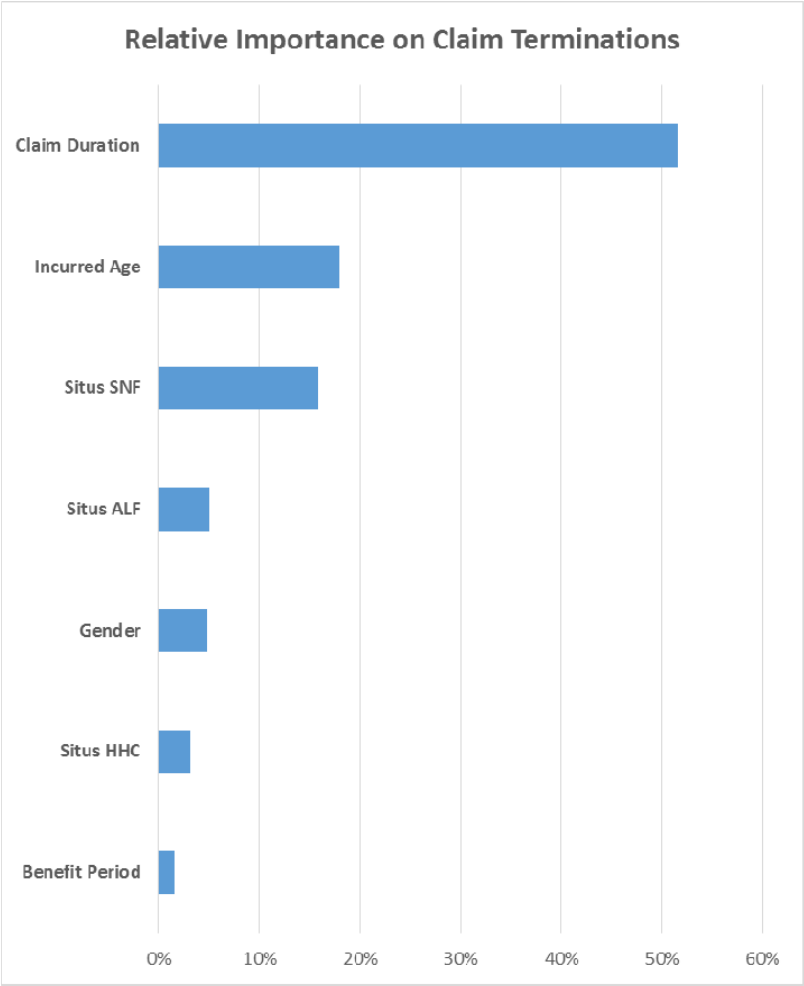
GBM introduces new interactions and buckets

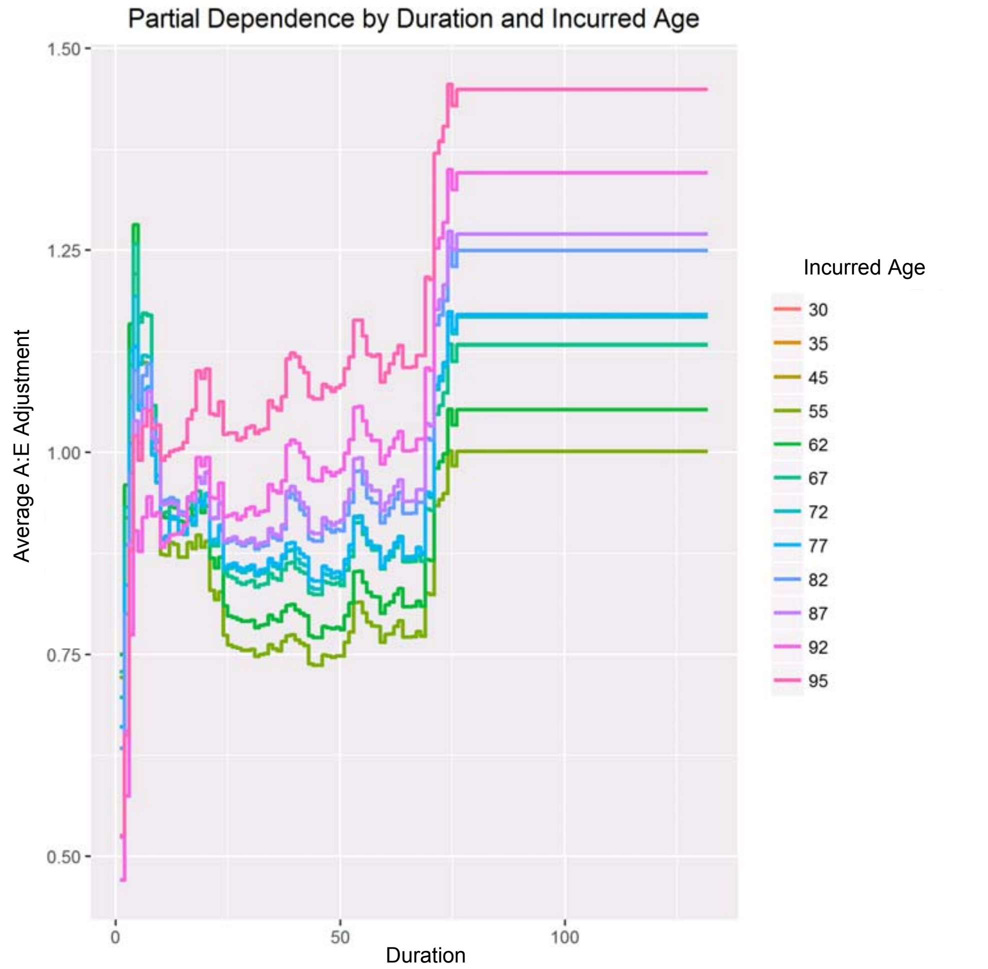
GBM

Why do we like it?

- Little input from researcher
- Uncover complex interactions
- Powerful predictor
- Aid to build interactions for penalized GLM
- Understanding variables' relative importance

GBM





Useful in understanding complex data

What challenges remain?

- Limited data
  - – Supplement with industry data
- Beyond experience data
  - – Hold level or grade off adjustment
- Trend
  - – Understand driver
  - – Hold level or grade off

Summary of case study

Penalized GLM and GBM

Similarities to traditional, but…

- Automates bias-variance tradeoff
- Efficient to update/modify
- Reduces human error
- Uncovers complex interactions

Remember to still consider…

- Supplementing with industry data
- Using judgement with extrapolation and trend

Questions?

References and resources

- 2017 ILTCI predictive modeling workshop homework and resources http://www.iltciconf.org/predictivemodelingmaterials.htm
- Armstrong, B. & Parkes, S. (2015). Calibrating Risk Score Model with Partial Credibility. Forecasting and Futurism, 11, 25-29.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. New York, NY: Springer-Verlag New York.
- Friedman, J., Hastie, T., Tibshirani, R., (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33(1). Retrieved from http://www.jstatsoft.org/v33/i01/
- Miller, H. (2015). A Discussion on Credibility and Penalised Regression, with Implications for Actuarial Work.
- R Development Core Team (2011), R: A Language and Environment for Statistical Computing. Vienna, Austria : the R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at http://www.R-project.org/.
- Hastie, T., Tibshirani, R., & Friedman, T. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition. New York, NY: Springer-Verlag New York.