Institute and Faculty of Actuaries

Machine Learning & Experience Analysis

Louis Rossouw

21 February 2023

Institute and Faculty of Actuaries

Machine Learning & Experience Analysis

Using machine learning to model claims experience and reporting delays for pricing and reserving

21 February 2023

Institute and Faculty of Actuaries

Using machine learning to model claims experience and reporting delays for pricing and reserving

By L Rossouw and R Richman

Presented at the Actuarial Society of South Africa's 2019 Convention
22–23 October 2019, Sandton Convention Centre

ABSTRACT

In this paper we review existing modelling approaches for analysing claims experience in the presence of reporting delays, reviewing the formulation of mortality incidence models such as GLMs. We then show how these approaches have traditionally been adjusted for late reporting of claims using either the IBNR approach or the more recent EBNER approach. We then go on to introduce a new model formulation that combines a model for late reported claims with a model for mortality incidence into a single model formulation. We then illustrate the use and performance of the traditional and the combined model formulations on data from a multinational reinsurer. We show how GLMs, lasso regression, gradient boosted trees and deep learning can be applied to the new formulation to produce results of superior accuracy compared to the traditional approaches.

KEYWORDS

Machine learning; IBNR; incurred but not reported; experience analysis; reinsurers; EBNER; analytics; gradient boosted trees; deep learning; mortality models; pricing and reserving

CONTACT DETAILS

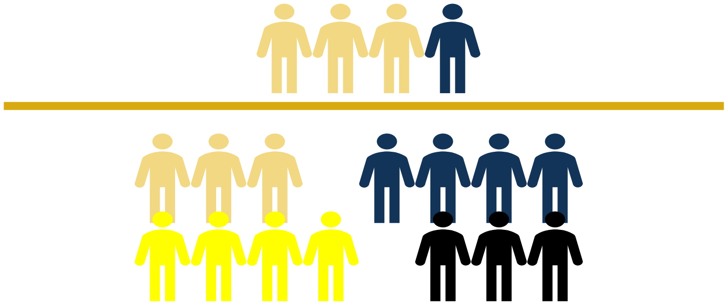
Mr Louis Rossouw, Gen Re, Cape Town; Email: LRossouw@genre.com
Mr Ronald Richman, QED, Johannesburg; Email: ronald.richman@qedact.com

Rossouw, L. and Richman, R. (2019) 'Using machine learning to model claims experience and reporting delays for pricing and reserving', in. *Actuarial Society of South Africa Convention 2019*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3465424.

Experience Analysis

- Experience analysis is key part of ongoing management of Life Insurance book
- Check and update mortality assumptions for pricing and reserving
- Understand the rate at which people claim
- In practise, often performed via AvE analysis

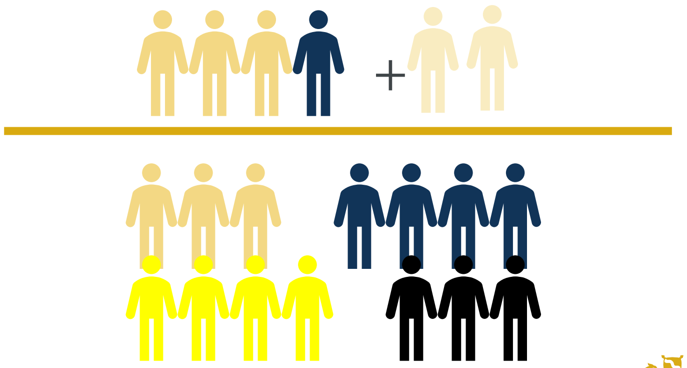
$$m_x = \frac{\theta_x}{E_x} =$$



Claims are reported late…

- Some allowance necessary for Incurred but not Reported (IBNR) claims
- More problematic for reinsurers than direct writers
- Less emphasis on IBNR methods for Life Insurers compared to GI

$$m_x = \frac{\theta_x + IBNR_x}{E_x} =$$



Institute and Faculty of Actuaries

How do we estimate IBNR?

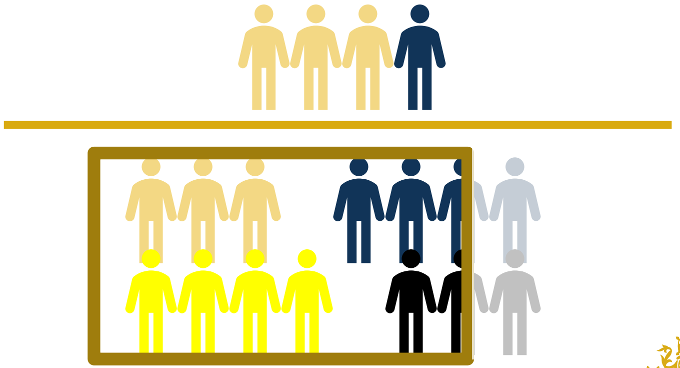
- Many methods available to estimate IBNR:
  - Exposure free methods - Chain-ladder
  - Exposure based methods - Bornhuetter Ferguson and Cape-Cod
- Key assumption – past pattern of claim development indicative of future claim development

Year	Delay 0	Delay 1	Delay 2	Delay 3
2016				
2017				
2018				
2019				

EBNER

- Exposed but Not Expected to be Reported
- Key principle: take away exposure instead of adding claims
- No need to decide where to add claims
- Appropriately down weights experience that is less developed
- Does not amplify volatility in cells where run-off is not complete

$$m_x = \frac{\theta_x}{E_x R_x} =$$



Lewis, P. and Rossouw, L. (2007) ‘IBNR versus EBNER – An Alternative Method of Allowing for Late Reported Claims’, in. *Actuarial Society of South Africa Convention 2007*. Available at: http://legacy.actuarialsociety.org.za/Portals/2/Documents/Convention-IBNRvsEBNR-PJL-LJR-2007.pdf (Accessed: 12 May 2021).

Modelling Mortality

Goal of mortality modelling – approximate experience with a function

$$m_x \approx \mu_x = f(\mathbf{x})$$

1. Reference table
2. Mortality law
3. GLM

$$\mu_x = \%\ of\ table$$

$$\mu_x = ae^{b(x-M)} + c$$

$$\mu_x = e^{\sum_i x_i \beta_i}$$

GLM Representation

(can include many different covariates)

$$\ln(\Theta_x) \approx \beta_0 + \sum_i x_i \beta_i + \ln(E_x)$$

Define our data

$$t_i' = \begin{cases} 1 \; if \; calendar \; year = i \\ 0 \; otherwise \end{cases}$$

$$d_i'' = \begin{cases} 1 \; if \; development \; year = i \\ 0 \; otherwise \end{cases}$$

Run-off triangles = GLMs!

Development Period

Loss Period

Year	Delay 0	Delay 1	Delay 2	Delay 3
2016	👦👩👦👩		👩	👦
2017	👦👩👦		👦	👩
2018	👦👩👦		👩	👦
2019	👦👩👦👩	👦👩	👩	👦

GLM Representation

$$\ln(\Theta^{t,d}) \approx \beta_0 + \sum_i t_i'\beta_i' + \sum_j d_j''\beta_j''$$

Loss Period

Development Period

Advantages

- Simple to code in R
- Can generalise
  - Smooth functions of t and/or d
  - Other variables
  - E.g. product, waiting period, interactions
  - Can decide how to include data e.g. triangles or parallelograms
- Any machine learning techniques

But, be careful when constructing data

Generalising

- Interactions

$$\ln\left(\theta_x^{t,d}\right) \approx \sum_i t_i' \beta_i' + \sum_j d_j'' \beta_j'' + \sum_j x_i d_j'' \beta_{ij}'''$$

- Our run-off can change based on underlying data
- Longer run-off for older people
- By duration etc.
- t and/or d can be in months, quarters or years

Institute and Faculty of Actuaries

But there is more…

Development Period

Exposure

Loss Period

Year	Delay 0	Delay 1	Delay 2	Delay 3
2016	👦👧👦👧		👧	👦
2017	👦👧👦		👦	👧
2018	👦👧👦		👧	👦
2019	👦👧👦👧	👦👧	👧	👦

GLM Representation

$$\ln(\Theta^{t,d}) \approx \beta_0 + \sum_i t_i' \beta_i' + \sum_j d_j'' \beta_j'' + \ln(E_{x_t})$$

Loss Period

Dev. Period

Mashup!

Can we combine the GLM representations for mortality (by gender, age, product etc) and IBNR?
=> a single experience analysis model
Model relates to the “partial” mortality rates in each period (similar to Mack’s Incremental Loss Ratio method)

GLM Representation of **Mortality**

+

GLM Representation of **IBNR**

$$\ln\left(\theta_x^{t,d}\right) \approx \beta_0 + \sum_i x_i \beta_i + \sum_i t_i' \beta_i' + \sum_j d_j'' \beta_j'' + \ln(E_x)$$

Institute and Faculty of Actuaries

Illustration of approach

Fit model on triangle of rates

$$m_{\mathbf{x}'}^{t,d} \approx \mu_{\mathbf{x}'}^{t,d} = f(\mathbf{x}')$$

$$m_{\mathbf{x}'}^{t,d} = \frac{\theta_{\mathbf{x}'}^{t,d}}{E_{\mathbf{x}'}^{t}}$$

Derive overall rate by adding up over cells

$$\mu_x = \sum_{t,d} \mu_{\mathbf{x}'}^{t,d}$$

Year	Delay 0	Delay 1	Delay 2	Delay 3
2016	$m_{\mathbf{x}'}^{2016,0}$	$m_{\mathbf{x}'}^{2016,1}$	$m_{\mathbf{x}'}^{2016,2}$	$m_{\mathbf{x}'}^{2016,3}$
2017	$m_{\mathbf{x}'}^{2017,0}$	$m_{\mathbf{x}'}^{2017,1}$	$m_{\mathbf{x}'}^{2017,2}$	
2018	$m_{\mathbf{x}'}^{2018,0}$	$m_{\mathbf{x}'}^{2018,1}$		
2019	$m_{\mathbf{x}'}^{2019,0}$			

Year	Delay 0	Delay 1	Delay 2	Delay 3
2016	$\mu_{\mathbf{x}'}^{2016,0}$	$\mu_{\mathbf{x}'}^{2016,1}$	$\mu_{\mathbf{x}'}^{2016,2}$	$\mu_{\mathbf{x}'}^{2016,3}$
2017	$\mu_{\mathbf{x}'}^{2017,0}$	$\mu_{\mathbf{x}'}^{2017,1}$	$\mu_{\mathbf{x}'}^{2017,2}$	$\mu_{\mathbf{x}'}^{2017,3}$
2018	$\mu_{\mathbf{x}'}^{2018,0}$	$\mu_{\mathbf{x}'}^{2018,1}$	$\mu_{\mathbf{x}'}^{2018,2}$	$\mu_{\mathbf{x}'}^{2018,3}$
2019	$\mu_{\mathbf{x}'}^{2019,0}$	$\mu_{\mathbf{x}'}^{2019,1}$	$\mu_{\mathbf{x}'}^{2019,2}$	$\mu_{\mathbf{x}'}^{2019,3}$
2020	$\mu_{\mathbf{x}'}^{2020,0}$	$\mu_{\mathbf{x}'}^{2020,1}$	$\mu_{\mathbf{x}'}^{2020,2}$	$\mu_{\mathbf{x}'}^{2020,3}$

What is the benefit?

- Approach is more complicated than traditional AvE model
- But significant benefits:
  - Simplifies the approach: 2 models → 1 model
  - Impact of modelling choices understood
  - Prediction uncertainty
  - Interactions of multiple parameters
  - Better weighting
  - No need to decide where to add claims
  - Estimates mortality
  - Estimates for IBNR reserves

Generalising

- Allowing or expected mortality and development

$$\ln\left(\theta_{\mathbf{x}}^{t,d}\right) \approx \beta_0 + \sum_i x_i\beta_i + \sum_i t_i'\beta_i' + \sum_j d_j''\beta_j'' + \ln(\mu_x^{prior} E_{\mathbf{x}} w_{\boldsymbol{d}})$$

- μ_x^{prior} is an existing mortality table
- $w_{\boldsymbol{d}}$ is an existing expected reporting pattern

Generalising

- Interactions

$$\ln\left(\theta_{\mathbf{x}}^{t,d}\right) \approx \sum_{i} x_i \beta_i + \sum_{i} t_i' \beta_i' + \sum_{j} d_j'' \beta_j'' + \sum_{j} x_i d_j'' \beta_{ij}''' + \ln\left(\mu_x^{prior} E_{\mathbf{x}} w_{\boldsymbol{d}}\right)$$

- Our run-off can change based on underlying data
- Longer run-off for older people
- By duration etc.
- Care is needed in construction of data
  – E.g. cells with no claims
- Much more data in exposure

Machine Learning

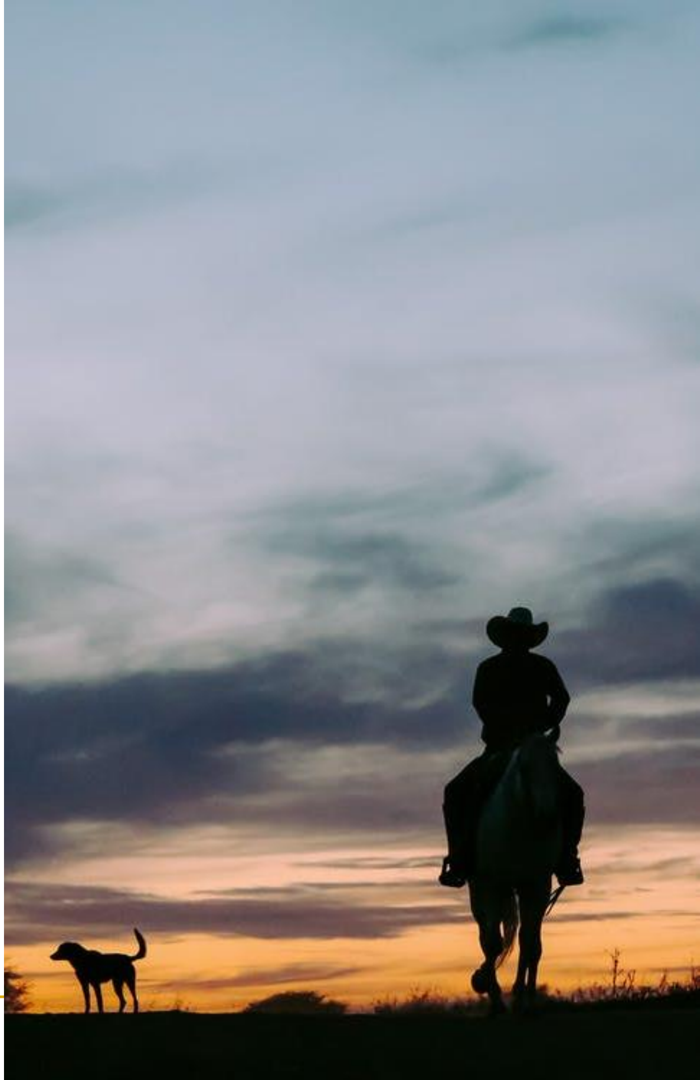
To apply the proposed modelling approach, we considered following categories of models:

- Traditional GLMs
- Lasso Regression
- Gradient Boosted Trees
- Deep Learning

Lasso Regression

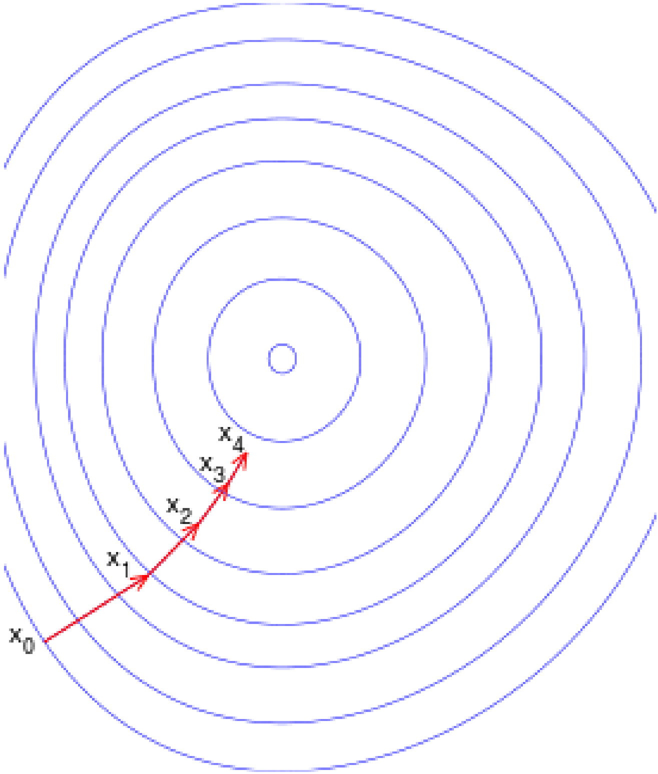
- GLM with penalty on size of coefficients
- Least absolute shrinkage and selection operator (LASSO)
- Advantages of technique:
  - Automated Variable Selection
  - Regularisation
- All controlled with single parameter, set via cross-validation

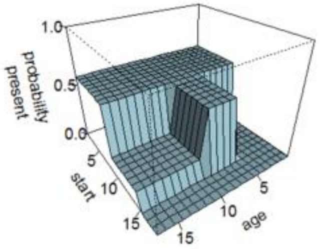
$$\sum_i \left(y_i - \beta_0 - \sum_j \boldsymbol{x}_{ij} \beta_j \right)^2 + \lambda \sum_j \left| \beta_j \right|$$

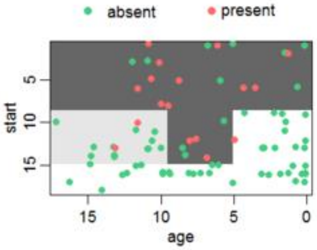


Gradient Boosted Trees

- Very successful approach for tabular data
- General principle of boosting – fit successive models to minimize error => fit highly complex functions
- Boosted tree models fit successive decision trees
- “Descending” to optimal solution

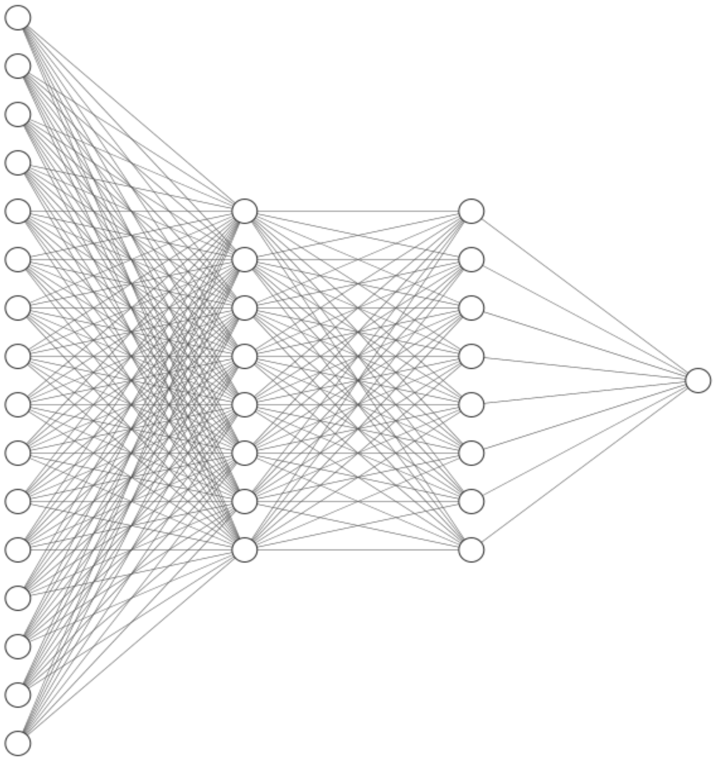






Deep Learning

- Deep Learning automatically constructs hierarchies of complex features to represent abstract concepts
- Features in lower layers composed of simpler features constructed at higher layers => complex concepts can be represented automatically
- Typical example of deep learning is feed-forward neural networks, which are multi-layered machine learning models, where each layer learns a new representation of the features.
- The principle: Provide raw data to the network and let it figure out what and how to learn.



Institute and Faculty of Actuaries

Application of combined model

- Model tested on data from 4 portfolios:
  - 2 countries - UK & South Africa
  - 2 products - Mortality & CI
- Several co-variates available for modelling
- Models from classes discussed above trained on dataset
- Assumption of Poisson loss function (standard for modelling mortality with GLM)
- Also measured AvE
- Fit 2 traditional models – estimate IBNR and then estimate mortality
- Fit 4 combined models – using 4x models (GLM/LASSO/GBT/DL)

Field	Description
Company	A-D (Portfolio)
Benefits	Death (with and without accelerator), Critical Illness (with and without accelerator)
Product type/code	Whole life Term (level and decreasing)
Gender	Male or Female
Smoker status	Smoker or Non-smoker
Country	United Kingdom or South Africa
Joint life indicator	Joint Life First Death or Single
Rate	Standard Extra mortality loading Per mille loading
Policy_Year	Curate years since policy start
Calendar year	Each calendar year for exposure and claims events
Underwriting year	Policy commencement year

Train, Validate, Test

- Data split into 3 sets:
  - Training (to estimate models)
  - Validation (initial assessment of performance)
  - Testing (final quantification of predictive ability)
- Train data on claims reported up to 2009
- Validate data on claims reported in 2010
- Test data on claims reported in 2011 and 2012

Year	Delay 0	Delay 1	Delay 2	Delay 3
<=2006	Train	Train	Train	Train
2007	Train	Train	Train	Validate
2008	Train	Train	Validate	Test
2009	Train	Validate	Test	Test
2010	Validate	Test	Test	
2011	Test	Test		
2012	Test			

Institute and Faculty of Actuaries

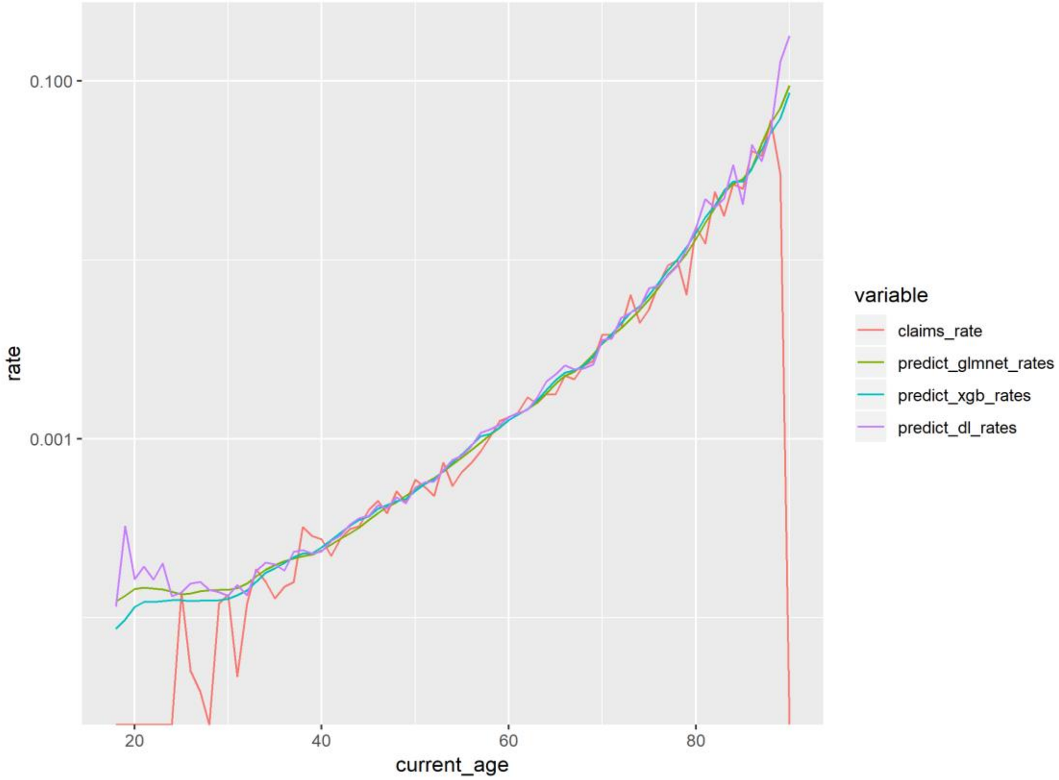
Performance – all claims

Table shows **test set** performance of models.

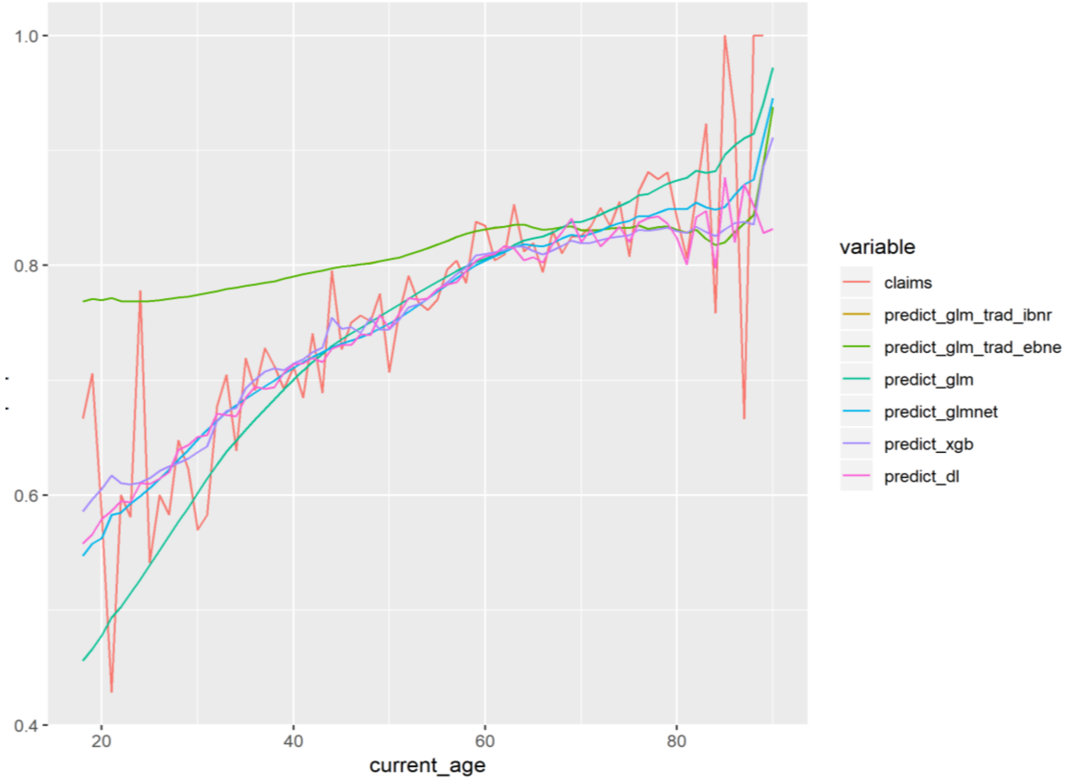
- ML models outperform GLMs
- Combined models outperform traditional models
- Apparent trade-off between goodness of fit (Poisson Deviance) and AvE (which measures bias)

Model	Poisson Deviance
IBNR + GLM	22 944
EBNER + GLM	22 947
GLM	22 883
Lasso	22 826
Gradient Boosted Tree	22 822
Deep Learning	22 799

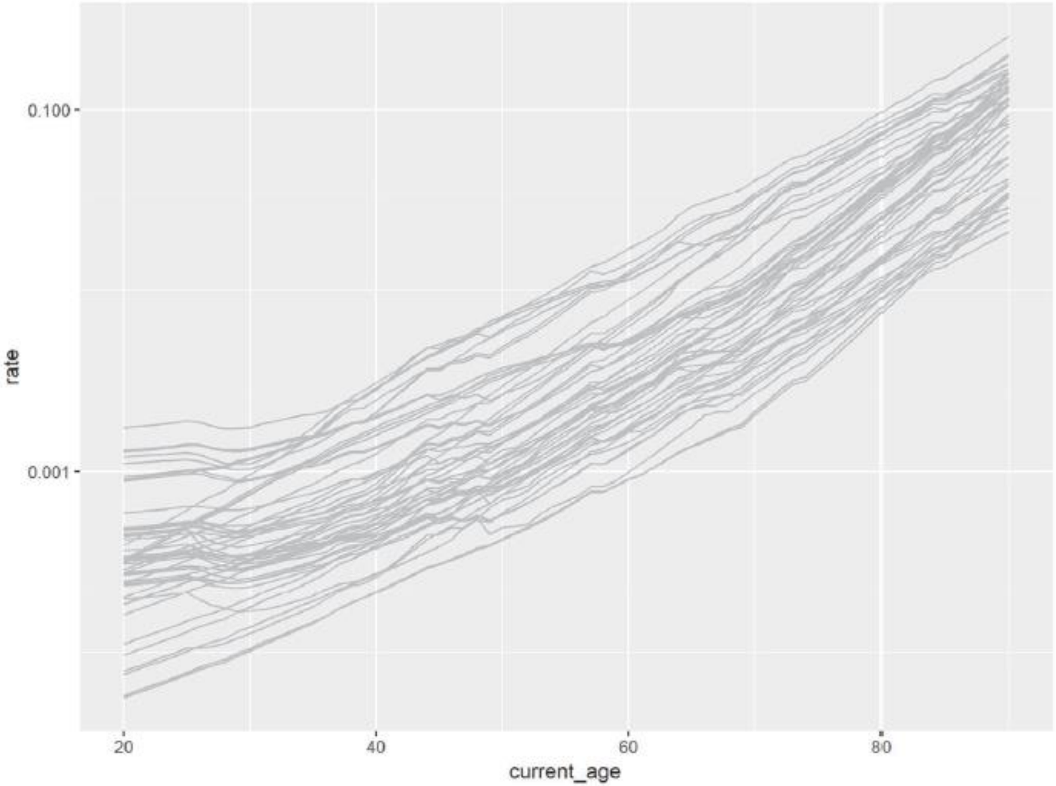
Mortality Rates by Age



Proportion of Claims Reported in First Year by Age



Individual Conditional Expectation - Gradient Boosting



Institute and Faculty of Actuaries

Key Take-Aways

- Combined modelling technique improves accuracy!
- Links between delay and other variables
- ML improves accuracy
  - GLMs can be good
  - Regular GLMs need lots of fine tuning “per hand”
- Lasso Regression gets there quickly (but less interpretable)
- Care is needed when extrapolating (any model)
- Must decide on suitable loss metrics…
- … and ensure that ML techniques are corrected for bias
- ML has a place in mortality modelling

Questions

Comments

The views expressed in this [publication/presentation] are those of invited contributors and not necessarily those of the IFoA. The IFoA do not endorse any of the views stated, nor any claims or representations made in this [publication/presentation] and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this [publication/presentation].

The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this [publication/presentation] be reproduced without the written permission of the IFoA [*or authors, in the case of non-IFoA research*].

Institute and Faculty of Actuaries

Institute and Faculty of Actuaries

Thank You!

Louis Rossouw
Gen Re | Segment Head of Research & Analytics
LRossouw@GenRe.com

Institute and Faculty of Actuaries

Machine Learning & Experience Analysis

Louis Rossouw

21 February 2023