CERA Global Risk Conference 2021

from 14 to 17 June 2021

Presented on actuview

CERA Global Risk Conference 2021

from 14 to 17 June 2021

BELIEVING THE BOT – MODEL RISK IN THE ERA OF DEEP LEARNING

Ronald Richman

SA Taxi, Managing Head: Insurance Actuarial

INTRODUCTION

Presented by:

Ronald Richman, SA Taxi, rrichman@sataxi.co.za

Joint work with:

Nicolai von Rummell, QED Actuaries & Consultants nicolai.von.rummell@qedact.com

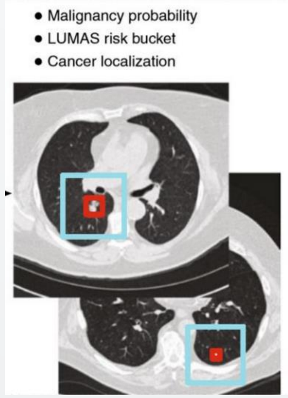
Mario V. Wüthrich, ETH Zürich mario.wuethrich@math.ethz.ch

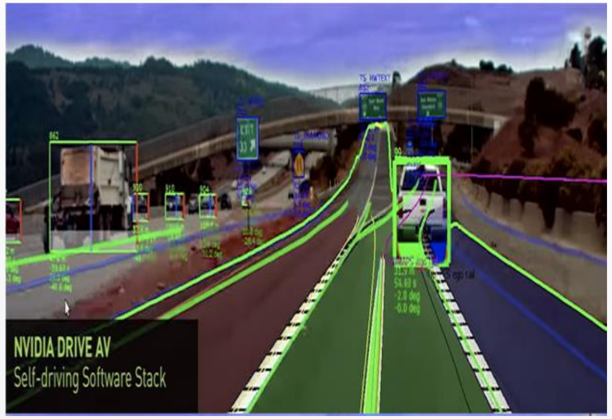
Paper:

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3444833

AI APPLICATIONS ON THE RISE

Man from www.thispersondoesnotexist.com/
Mona Lisa from Samsung AI team
Text from https://talktotransformer.com/
Self- driving from NVIDIA blog
Cancer detection from Nature Medicine





An exciting part of the world of finance is insurance

I think we all know that the insurance industry is exciting. I see it everywhere - the airlines, the cars, most all the businesses in the world. The insurance industry can really drive the economic innovation.

But one area of insurance that I really want to see develop more is financial advice. It might be a private sector service but insurance companies are not really there anymore. In general we are not allowed to talk to clients about financial solutions - we need to find a new solution. It would be fun to see what a private sector insurance can deliver.

PITFALLS WHEN USING AI MODELS

- Amazon recently scrapped its recruitment algorithm, that the company has been using since 2014
- Algorithm selected top candidates based on their resumés only
- Trained using past applications at Amazon introducing pro-male bias in training set
- Editing the algorithm could not ensure removal of bias
- Therefore, the algorithm was scrapped

CERA Global Risk Conference 2021
from 14 to 17 June 2021

- **From traditional modelling to deep learning**
- Introduction to model risk management
- Model risk management for DL Models

HOW TO DEFINE A MODEL

"a set of verifiable mathematical relationships or logical procedures which is used to represent observed, measurable real-world phenomena, to communicate alternative hypotheses about the causes of the phenomena, and to predict future behaviour of the phenomena or the purpose of decision making".

Base	Interpretation	Purpose
Model	Explanation	Decision making
Representation	Prediction	
Real World		

ACTUARIAL MODELS: EXPLANATION OR PREDICTION?

Goal	Build casual understanding	Make accurate predictions
Consequences	• Favour techniques that are interpretable • Require unbiased models • Consider strength of inferences	• Use complex algorithms • Accept bias if it increases predictive accuracy • Quantify predictive accuracy

Common actuarial techniques usually applied for predictive purposes:

- *GLMs*
- *Chain-Ladder and Bornhuetter-Ferguson*
- *Lee-Carter*

FORMAL DEFINITION OF A MODEL

$$M(X; T; S(A, E); \Theta) = \hat{y}$$

X Matrix of known variables

S Specification of class of algorithms A and explicit model definition E

Unknown variable $\hat{y}$

T set of functions defining new variables X’

Θ Set of parameters

PARADIGM OF TRADITIONAL ACTUARIAL MODELLING

$$M\left(X; T; \sum \beta_i f_i(x_i); \Theta\right) = \hat{y}$$

X usually relatively simple and might be in the region of 50 variables or fewer

Quality tested using various techniques but usually not predictive performance

T often specified manually:
- For most variables select f_i asidentify or step functions
- Multiplying known variables to create interaction effects important in pricing

Model fit by minimizing loss function derived from error distribution of predicted variable $\hat{y}$ (i.e. stochastic data generating process specified))
Derives a set of model parameters β_i (no casual parameters because confounders not usually assessed)

Class of algorithms is usually linear (although good alternatives, such as GAMS, exist)
Model specification is choosing known variables for X'

MOVING TO MACHINE LEARNING MODELS

$$M(X;T;S(A,\tilde{E});\Theta) = \hat{y}$$

Characteristics of ML models

- Two most successful classes (gradient boosted trees, neural networks)
- Usually combined with extensive feature engineering involving complicated functions in T
- Structure of model determined by model class A
- No explicit model specification E
- Stochastic data generating process usually not considered – model fit using objective functions such as MSE

Differences from traditional actuarial modelling

- ML algorithms may produce highly non-linear models
- Interaction of multiple variables via feature engineering
- => Higher complexity and less transparency
- Many different models may achieve similar performance i.t.o. loss function but produce very different policy level results
- May produce high quality predictions at policy level but be biased at portfolio level
- Since models may involve random processes, results develop a dependency on seed

No unique best model due to under-determination of model (insufficiently specified loss function)

NEUTRAL NETWORKS AND DEEP LEARNING

Critique of Feature Engineering

1	Complexity	Difficult to choose most relevant features in complex problems (high-dimensional data or large-scale problems)
2	Effort	Requires substantial time spent on a problem by suitably qualified people
3	Expert Knowledge	Depends on a body of expert knowledge on how to tackle problems

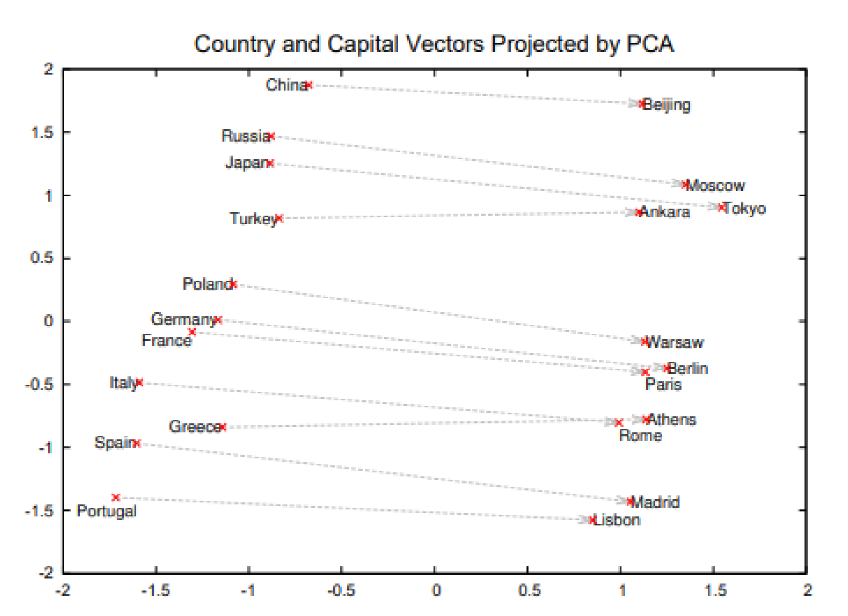
Relevance for actuarial modelling

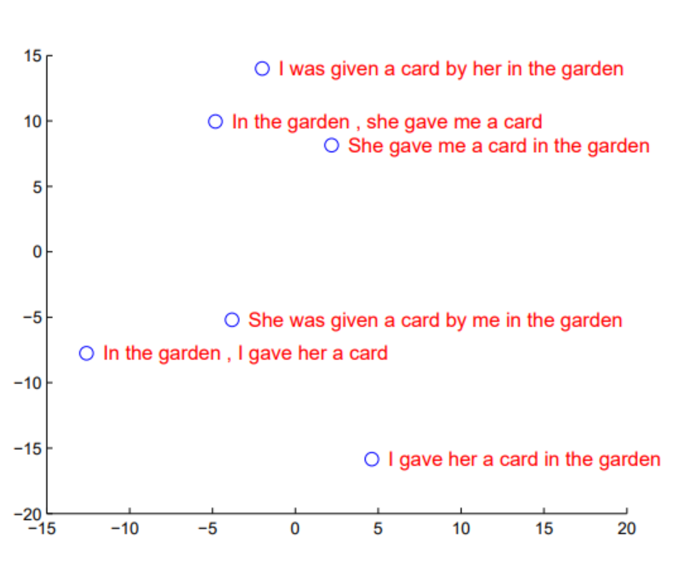
1. Are actuarial data complex? Relatively few data points collected for each policy
2. Many years of collective professional and institutional experience in modelling techniques

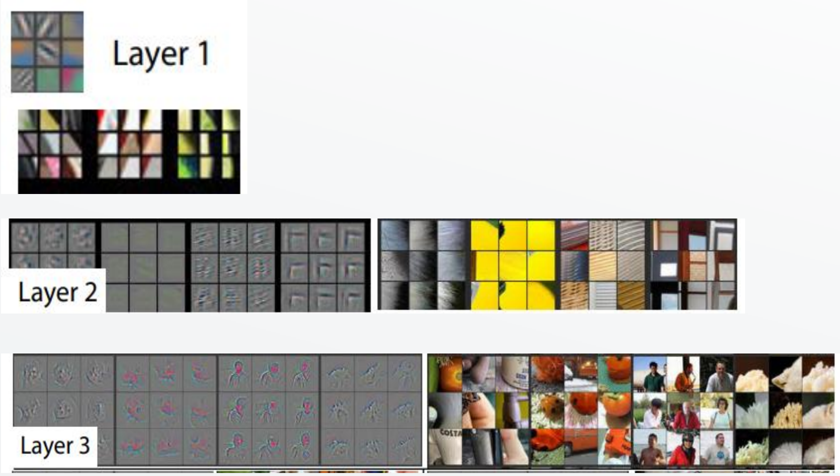
A. Techniques for large-scale actuarial problems not well developed
B. Complex data becoming more commonplace – telematics/wearables/individual claims
C. New deep learning techniques shown to outperform traditional techniques relatively easily

$$M\left(X;\tilde{T};S\left(A,\tilde{E}\right);\Theta\right)=\hat{y}$$

REPRESENTATION LEARNING - EXAMPLES







Cities & Capitals – Mikolov et al (2013)
Sentence embedding – Sutskever (2014)
Filters & Activations – Zeilier & Fergus (2013)

ACTUARIAL EXAMPLES OF DL

	Pricing	Reserving	Telematics	Mortality Forecasting	Quantitative Risk Management
Feed-forward Nets	• Ferrario, Noll and Wüthrich (2018) • Noll, Salzmann and Wüthrich (2018) • Rossouw and Richman (2019) • Wüthrich and Buser (2018)	• Castellani, Fiore, Marino et al. (2018) • Doyle and Groendyke (2018) • Gabrielli and Wüthrich (2018) • Hejazi and Jackson (2016, 2017) • Kuo (2019) • Wüthrich (2018) • Zarkadoulas (2017)	• Gao and Wüthrich (2017) • Gao, Meng and Wüthrich (2018) • Gao, Wüthrich and Yang (2018)		• Castellani, Fiore, Marino et al. (2018) • Hejazi and Jackson (2016, 2017)
Convolutional Neural Nets			• Gao and Wüthrich (2019)		
Recurrent Neural Nets		• Kuo (2018a, 2018b)		• Nigri, Levantesi, Marino et al. (2019)	
Embedding Layers	• Richman (2018) • Schelldorfer and Wüthrich (2019) • Wüthrich and Merz (2019)	• Gabrielli, Richman and Wüthrich (2018) • Gabrielli (2019) • Poon (2019)		• Richman and Wüthrich (2018)	
Autoencoders			• Richman (2018)	• Hainaut (2018) • Richman (2018)	

CERA Global Risk Conference 2021

SUMMARY

Traditional Actuarial

$$M\left(X;T;\sum \beta_i f_i(xi)\,;\Theta\right) = \hat{y}$$

- Linear model specification, for f_i identity (GLM), f_i spline function (GAM)
- β_i regression parameters

Machine Learning

$$M(X;T;S(A,\tilde{E});\Theta) = \hat{y}$$

- Implicit Specification of the model $\tilde{E}$ by a class of algorithms A

Deep Learning

$$M(X;\tilde{T};S(A,\tilde{E});\Theta) = \hat{y}$$

- Representation Learning: Implicit Specification of functions $\tilde{T}$ to derive features X'
- Explicit use of loss function $\mathrm{L}(y,\hat{y})$ to measure predictive accuracy

ASPECTS OF A TYPICAL MODELLING PROCESS

1	Design	Choosing all the defined parts of a Model M
2	Calibration	Determination of parameters θ, depending of Algorithms A and Loss function
3	Validation	Assessing model accuracy, robustness and suitability for decision making
4	Implementation	Technical configuration and set-up of model on suitable platform
5	Operation	Running the model and producing of model predictions $\hat{y}$

These steps are a simplified representation of a modelling process

CERA Global Risk Conference 2021
from 14 to 17 June 2021

- From traditional modelling to deep learning
- **Introduction to model risk management**
- Model risk management for DL Models

SOME EXAMPLES OF MODEL RISK[1]

LTCM Hedge Fund

- In 1998 LTCM was bailed out by the Federal Reserve with $3.6bn to prevent a chain reaction in the financial system
- Before the LTCM adopted aggressive strategy to maintain high returns for their clients
- This strategy included merger arbitrage with small margins and more and more highly leveraged positions
- In 1998 Russia effectively defaulted on its bonds and markets fell by 35% (Europe) and 20% (US)
- LTCM lost $4.4bn of $4.7bn of its capital in less than a year

JP Morgan London Whale

- Since 2010 the CIO of JPM bought synthetic CDS derivates to protect the bank against economic downturn (SCP Portfolio)
- While intended as a protection, the SCP portfolio became a speculative investment and included short-positions that served as a source of profits betting on an upturn of the markets
- Investments in SCP increased to £157bn in early 2012
- Existing risk management tools highlighted the risk but instead of reducing the risk, JPM changed the assessment method including a spreadsheet error
- In 2012 European sovereign debt crisis led to losses of £6bn

1: IFoA Model Risk Working Party (2015): Examples taken from the paper

WHAT IS MODEL RISK?

"the potential loss an institution may incur, as a consequence of decisions that could be principally based on the output of internal models, due to errors in the development, implementation or use of such models"

Cause → Event → Impact

Errors in the development, implementation or use of models → Decisions that could be principally based on the output of internal models → Potential loss an institution may incur

MODEL RISK CLASSIFICATION

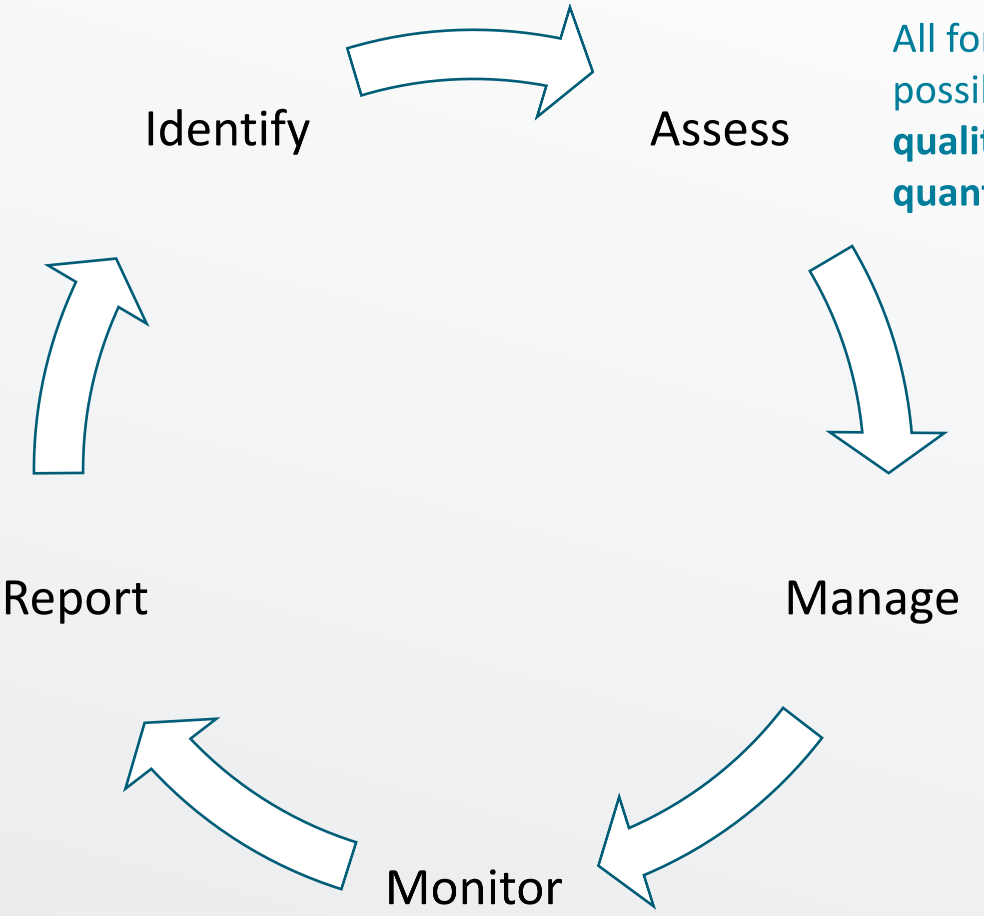
Model Risk

$$\boldsymbol{M(X;T;S(A,E);\Theta)=\hat{y}}$$

- **Structural Risk**
  - Specification Risk S(A,E)
  - Parameter Risk Θ
  - Numerical and Simulation Errors
- **Operational Risk**
  - Data Risk
  - Decision Risk
  - Other process Risk

RISK MANAGEMENT PROCESS

Identification of existing models, model purpose (**Model Inventory**)



All forms of model risk assessment is possible. From **qualitative** (RAG), to **semi-qualitative** (Frequency/Severity) to **quantitative**

Comparison of risk assessment to **risk appetite**, define **management actions** if risk is unacceptable

Tracking of management action and implementation of measures

CONTROLS FOR TRADITIONAL ACTUARIAL MODELS

Structural Risk

Specification Risk:
- Clear definition of the problem or business question the model is going to solve,
- Documentation and review of model specification
- Peer review of the model specification
- Governance (committee, minutes) to review model specification and decisions in model development

Parameter Risk:
- Calculate confidence intervals, sensitivities and stress tests
- Reasonability checks on the input and output

Operational Risk

Data Risk:
- Assessment of the input data X to determine sufficient data quality i.e. clean and fit for purpose,
- defined metrics and reporting formats to track data quality of the inputs
- Documentation and review of the manipulation or initial transformation T of the input data X

Decision Risk:
- Testing of output results against defined criteria
- Training/Communication for decision makers
- Identify areas of human intervention
- Documentation of board discussion on results, underlying assumptions and appropriateness
- Defining clear performance metrics to assess the predictive accuracy of the model

CERA Global Risk Conference 2021
from 14 to 17 June 2021

- From traditional modelling to deep learning
- Introduction to model risk management
- **Model risk management for DL Models**

RECAP: FROM TRADITIONAL ACTUARIAL MODELLING TO DL

Traditional Actuarial

$$M\left(X;T;\sum\beta_i f_i(xi);\Theta\right)=\hat{y}$$

- Linear model specification, for f_i identity (GLM), f_i spline function (GAM)
- β_i regression parameters

Machine Learning

$$M(X;T;S(A,\tilde{E});\Theta)=\hat{y}$$

- Implicit Specification of the model $\tilde{E}$ by a class of algorithms A

Deep Learning

$$M(X;\tilde{T};S(A,\tilde{E});\Theta)=\hat{y}$$

- Representation Learning: Implicit Specification of functions $\tilde{T}$ to derive features X’
- Explicit use of loss function $\mathrm{L}(y,\hat{y})$ to measure predictive accuracy

CONSEQUENCES OF IMPLICIT MODEL SPECIFICATION FOR ML MODELS[1]

Simulatability	Being able to comprehend the model as a whole	• Develop expert knowledge on model architecture and heuristics to assess models • Issue pertains to some traditional models as well
Decomposability	All components of the model can be inspected and make sense	• Inspect learned representations in last layer of the model • Manual intervention and inclusion of prior knowledge is more difficult
Algorithmic Transparency[2]	Transparency of the learning algorithm and corresponding techniques	• Many techniques to mitigate risks of instability and consistency of DLM still to be developed • Ensembling many models and use of toy models to test outcomes of DLM

Increased exposure to model risk for ML/DL models and additional controls required

1: Lipton (2016) defines framework for model interpretability to assess two basic questions: Transparency or "How does the model work?" and Post-hoc Interpretability or "What else can the model tell me?"
2: Some of the techniques used for DLM are transparent but the consequences of applying those techniques are not always fully understood.

POST-HOC INTERPRETABILITY OR WHAT ELSE CN THE MODEL TELL ME?

Post-hoc interpretability means applying expert knowledge to the model outcome through using various techniques like visualization, examination of examples etc. to assess quality of results for decision making

Extraction information relevant for decision making

Many techniques exist for DL and are being developed (SHAP, LIME, local explanations etc.)

Both traditional and DL models are equally amendable to post-hoc interpretation…
…but some assumptions within these methods are not suitable for actuarial modelling
Consider the MACQ method: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3809674

CONSEQUENCES OF REPRESENTATION LEARNING: BIAS

Representation learning may lead to models exhibiting unwanted biases

Models might learn a proxy for factors that are illegal. Example – learning gender on the basis of motor vehicle utilization.

Data may contain biases that models may learn to reproduce accurately. Example – Amazon recruitment model.

Countering bias in Deep Learning models

- Companies should formally define list of unethical/illegal/unwanted biases for the organisation
- Explicitly test representations/output against those biases or close proxies (if possible)
- Build two models –including and excluding factors representing unwanted bias and test the differences
- Explicitly counter biases (if possible) using loss function that enforce equality or build another model to remove the effects of bias
- Another way is to apply the Discrimination Free Insurance Pricing method

UNINTENDED CONSEQUENCES OF REPRESENTATION LEARNING

Deep learning yields models that are highly adapted to the training data. If the training data does not match the intended use case well, then a risk of **unintended learning** may occur leading to **poor generalization** of models

General Examples

- Identify the breed of dog from an image
- Would hope that model learns characteristics of dog breed
- But snow can effectively identify a husky!
- Other recent example – identifying skin cancer from image of slides using dermatologist's markings

Insurance Modelling

- Unclear to what extent this is a risk for actuarial modelling
- If circumstances are changing rapidly (e.g. introduction of advanced driver assistance systems) then older data less useful => heightened risk with deep learning
- Inspect learned representations of models to ensure these accord with expert knowledge

PREDICTIVE ACCURACY OF DEEP LEARNING MODELS

	Shifting modelling paradigm:			Explicitly calculating Loss Function
	From			by
Why?	Description / causal explanation		How?	Splitting Train/Test Data
	To			to
	Prediction			Mitigate Decision Risk

Techniques to increase predictive accuracy also mitigate decision risk, apply for traditional modelling as well

LOSS FUNCTION AND ADDITIONAL CRITERIA FOR USE OF MODEL OUTPUT

Simple loss functions will typically not identify unique best model but class of sufficient good models

Predictive accuracy is only one criteria for using model outputs, other criteria (e.g. commercial considerations for pricing) are applied as well

Loss Function → Model → Decision

Investigate if additional criteria can be used to enhance loss function

STABILITY AND CONSISTENCY OF NEUTRAL NETWORKS

Training time

- Neural networks trained in stochastic manner (batch selection, parameter initialization, regularization)
- Induces dependence on random seed i.e. results may not be reproducible and hard to evaluate performance
- Multiplicity of models – may have significant implications for individual policyholders
- Potential solution – ensembling of models

Recalibration

- Consistency of neural networks over time not yet researched
- Ideally, should avoid major changes in outputs unless corresponding shift in training data
- Use stable techniques to assess if shift over time is reasonable

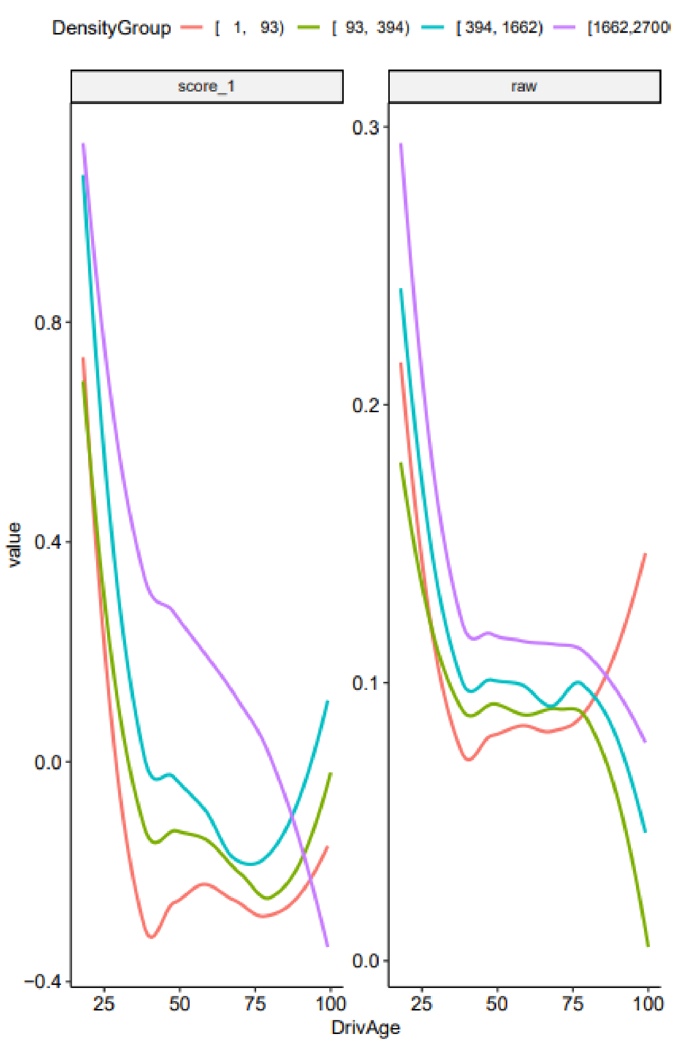
EXAMPLE 1 – FRENCH MOTOR TPL

Description

- Dataset consists of policy characteristics and claims on French Motor Third Party Liability portfolio
- Deep learning shown in several studies to outperform other approaches
- Analyzed using deep neural net with:
  - 5 layers of neurons
  - Embedding layers
- Code provided in paper

Model Risk Considerations

- Learned representations (after PCA) shown in figure and compared to observed frequencies
- In line with intuitions:
  - Shape is similar to data implied rates
  - Younger drivers have higher values than older drivers
  - High density has higher frequency than low density
- At oldest ages, model not in line with data and needs further investigation:
  - Modify model at oldest ages
  - Use GLM at older ages
- Not much ability to check biases within this dataset
- Could consider whether we expect learned representations to generalize well



CERA Global Risk Conference 2021

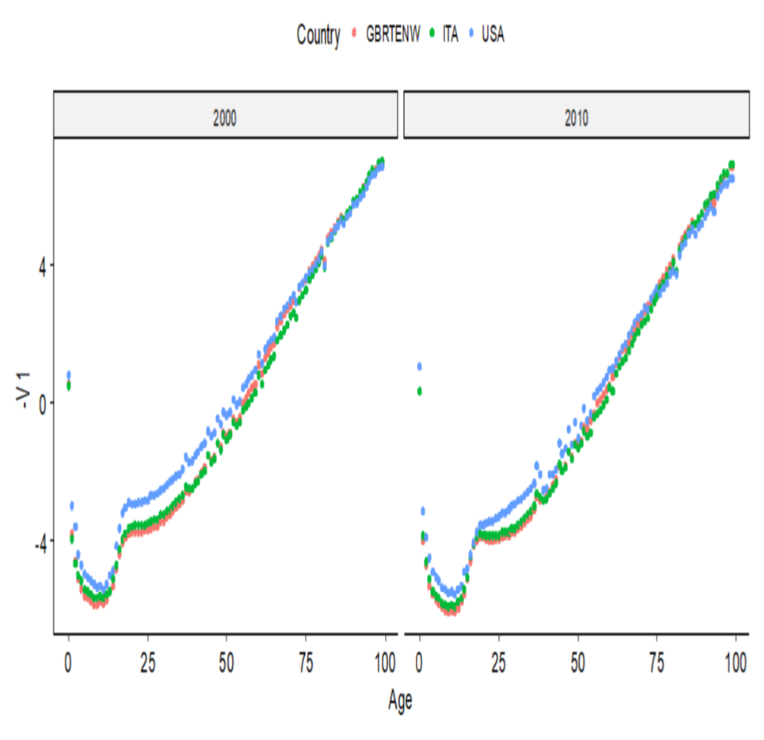
EXAMPLE 2 – MORTALITY FORECASTING

Description

- Dataset consists of mortality rates observed in 38 different countries in the period 1950-2016, for males and females
- Analyzed using deep neural networks in Richman and Wuthrich (2019):
- 5 layer deep neural network
- Embedding layers
- Code provided in the referenced paper

Model Risk Considerations

- Learned representations (after PCA) again in line with intuitions:
- Overall representation is in shape of lifetable
- Representations lower in 2010 compared to 2000
- Ranking of mortality by country maintained
- Greater risk compared to motor example, since need to forecasts out-of-sample and out-of-time => special controls required to ensure this part of the model is functioning correctly
- Unlikely to suffer from unwanted bias, but should be aware of “regime shift” to lower improvements in certain countries



CERA Global Risk Conference 2021

CONCLUSION

1. Defined formal framework for actuarial models (traditional, ML, DL)
2. Assessment of differences between models and consequences for model risk management
3. Implicit model specification introduces opaqueness to actuarial models
4. Representation learning exacerbates issues of bias in actuarial modelling
5. Controls have been identified to mitigate some of the additional exposure
6. Those controls require that actuaries acquire good understanding of DL models and practices

Success of DL models (predictive accuracy) and application in actuarial science indicate that further investigation into techniques to mitigate model risk is advisable