Joint Webinar on Machine Learning Q&A Session

Held on Monday, 14 July 2025

Widespread adoption

Thomas Sithole (Guest) 07:39 AM

Do you think we will ever see these methodologies being widely adopted

Answer: I am amazed by the response to this webinar, with a total of 426 registrants and 253 attendees joining us online. I hope with those attending the session, we can be agents of the changes we want to see in the profession.

Learning resources

Anonymous Attendee 07:42 AM

In past, I've learnt some of the topics discussed in the session but can't recollect, maybe because I'm not using it. Any tips here?

Also, what could be baby steps to start learning again?

Answer: With this seminar, we will be sharing the recording as well as the slides for attendees to use as a reference. In the slides we have included a number of links to learning resources we have created and curated, which we hope will help with the learning process.

Colin Priest (inadvertently removed)

I recommend Actuaries Institute's Data Science Applications microcredential course for those who wish to learn neural network programming skills, natural language processing, and practical LLM prompting

Answer: Agree, the course has good coverage of important relevant skills in the application of ML.

Sarsiz Chauhan (Guest) 07:45 AM

Having technical knowledge, and someone wanting to work in insurance industry, what all areas would you recommend to learn/understand of the insurance industry and where one can learn more on it.

I come from reinsurance background, have very limited exposure of insurance side.

Answer: I think this question is broader than the scope of the talk. I'd consider the actuarial education courses, CPD events, industry events, industry magazines available in your location to learn more about insurance.

Accuracy of ML methods

Anonymous Attendee 07:36 AM

Based on your implementation of ML in reserving, were their improvements in accuracy of your reserving eg better AvEs?

Walther Neuhaus (Guest) 07:51 AM

Benchmarking ML results to chain ladder results (with or without actuarial judgement) doesn't make for a high bar.

Gowdhami sakthivel (Guest) 07:41 AM

What have been the results and reliability of applying ML techniques to long-tail lines of business (LOBs) like liability insurance so far?

*Answer to questions above: My argument in favour of ML would be – if you have sufficient data of good quality, then using **all** the data, logically you should be able to achieve same or better results than a projection which only uses **some** of the data.*

An attendee shared this story, and I am confident there are many more out there:

Colin Priest (Guest) 07:42 AM

I can anecdotally say that I've worked with some European insurers who achieved more reliable reserving results in times of change (e.g. covid) by using these individual claims reserving models.

Data requirements

Djiby Fall (Guest) 07:47 AM

What types of data are recommended for the ML models, I mean ML models might have cofactors outside the traditional models say GLMs. Or does it not matter we can work with usual insurance claims data.

Joaquín Benavides (Guest) 07:47 AM

Have you, or any of the colleagues of the party, ever tried these solutions on underwriting year triangles? How did it go?

Answer to questions above: In the talk, I referenced examples where ML methods were applied to triangles, same as the traditional methods. However, personally I see the value of ML methods is highest when applied to insurance claims data of the highest detail of data you have available.

Actuarial judgement

Ronald T Kozlowski (Guest) 07:45 AM

Do you think that these models require fewer "actuarial judgments" or do they require more with model fitting? Can you give examples? For example, selecting different averages of LDFs?

DAMODARAN G (Guest) 07:47 AM

ruled based system probably which requires on judgmental approach with amendments better serve for mankind instead Blackbox more of numerical data prediction algorithm. Any possibility for incorporation of the same for counter check or warning dynamically.

Answer to questions above: I see actuarial judgement playing a different role in an MLR context – closer to how a pricing actuary may apply judgement. Rather than selecting individual factors, you may consider reasonableness checks, features to include, accuracy of diagnostics, model drift, including to what extent historical data is representative of anticipated trends etc.

One vs multiple models

Anonymous Attendee 07:50 AM

how do you know whether it's better to fit one ML model with data from two product types, with the product type as an independent variable vs fitting two independent ML models?

Answer: You may consider your understanding of your data and the algorithm. Do you have common features in the two product types and do you expect them to drive your data similarly? If there are commonalities, a single ML can find and leverage them. If not then makes sense for them to be independent.

If fitting two very different datasets and responses, you need to consider how well the chosen ML algorithm can fit bimodal response features. Some standardisation of the output feature may be required.

R vs Python

Anonymous Attendee 07:48 AM

Hi Jacky, Great Insights, which programming language is preferred better for Actuarial machine learning analysis in Reserving, is it Python or R?

Answer: Python has become the more popular choice for data analytics in the industry, so is my preferred if your team is getting started today. In a workplace environment, due to its wider popularity in the tech industry, it may be easier to get engagement and support from your IT team.

Python is definite recommendation if you are intending to work with neural networks. Whilst R versions are available, the most common neural network frameworks: Pytorch by Meta, Tensorflow by Google and JAX also by Google are Python-first.

R remains a solid choice if using more established ML and statistical techniques. Our blog and book has a significant amount of R content that you can leverage.

Sequence with Exposure

Colin Priest (inadvertently removed)

I can see how you can model IBNER/ultimate cost using the approaches you outlined e.g. sequence modelling.

How do you propose modelling IBNR? Is that a sequence based on each exposure? That's how I've done it.

Answer: To date, the Working Party has mostly focused on the IBNER with sequence modelling. But I had previously given this some thought and I am of the same view as you. You would start with the exposure and the exposure records with no claims would just stop there.

Exposure weights

Walther Neuhaus (Guest) 07:51 AM

Have you considered more explicit recognition of exposures to get apriori measures of likely ultimate cost?

Anonymous Attendee 07:51 AM

Have you attempted to use exposure/policy count as a predictor variable in your GLM model, as an offset, and if so, did you see an improvement in prediction?

Answer to questions above: Our examples include examples using weights in the data. It makes perfect sense to incorporate exposures where available.

Joining the group

Anonymous Attendee 07:55 AM

What kind of qualifications are you looking for from your volunteers?

Answer: We have enthusiasts from both academic and industry backgrounds. Please reach out to Sarah MacDonnell our chair for a chat to see if you'd be the right fit!