AKUR8

Explainable AI for Actuarial Applications

May 2025







Parhaam Behnoudnia

Actuarial Data Scientist

Edwin Graham

Principal Actuarial Data Scientist

First, let's comprehensively define Al



Al Overview

Yes, a Gradient Boosting Machine (GBM) is a type of artificial intellige (AI): <a># Comparison

Yes, a large language model (LL

Al Overview

Artificial intelligence (AI) is a set of technologies that allow machines to perform tasks that typically require human intelligence. Al uses math, computer science, and cognitive science to simulate human behavior. Al systems can:

- See, understand, and translate written and spoken language
- Analyze data

Al Overview

- Make recommendations
- Solve problems

Act to achieve a goal *o*

Actually, let's not

Al Overview

Yes, a Generative Language Model (GLM) is a type of machine learning

model that can be used to create AI: 🥀 🔔

The most basic form of artificial intelligence is **reactive machines**, which react to an input with a simplistic output programmed into the machine. In this form of AI, the program does not actually learn a new concept or have the ability to make predictions based on a dataset. May 15, 2023



Let's Define AI for the Purpose of this Presentation





Al is a Magic Box



If I am specific, this talk will be irrelevant in a month





No Matter the AI, it is an Input-Output Machine

This is the simplest visualisation



How are AI tools different than normal tools?



Manual Work is also an Input-Output Machine

At its simplest



It's more complicated than this...



After each output, manual work must be done to create a new output





After each output, manual work must be done to create a new output





After each output, manual work must be done to create a new output





After each output, manual work must be done to create a new output





Why Do We Iterate?



Why do we iterate?

Three major categories





People are involved in each step





Does AI Solve Iteration?



No Matter the AI, it is an Input-Output Machine

The inputs can get extremely complicated, but it is still an algorithm





Actuaries Will View using Al as a Loss of Control



Actuaries Want Full Control of the Process

For good reason





When Used Blindly, AI Does Remove Control from Actuaries Don't do this.





How Can Actuaries Use AI Properly?

We need to review the results, right?





Actuaries Must Apply Judgment When Using Al



Reputational Risk of Blindly Applying Al



Thomas



They did a great job spot tuckpointing. It was great to go with a company that is focused on brickwork instead of a general contractor. There were a few missed spots and after I called, they came out ...

Response from the owner

ChatGPT said: ChatGPT Thank you for your feedback! We're glad you're pleased with our tuckpointing work and appreciate your understanding regarding the missed spots. We're happy to have promptly addr...





Risk of Blindly Applying Al

This is the best statistical fit for this variable from a magic box output





The Statistically Best Model is Almost Never the Best Actuarial Model



No Matter the AI, it is an Input-Output Machine

AI is always data and assumption driven



Has your reserve estimate ever been **purely statistical** and involved **no form of actuarial judgment**?

If so, why are there actuaries?



No Matter the AI, it is an Input-Output Machine

AI is always data and assumption driven



How do we apply actuarial judgment when using AI?



How NOT to Apply Actuarial Judgment When Using Al



This Magic Box Doesn't Work

Don't iterate with a magic box





How to Succeed at Using Al



This Magic Box Works!

One example of a potential application of AI in actuarial work





This Magic Box Works!

One example of a potential application of AI in actuarial work





How to Apply Control in an Al-Driven Process



Al Replaces Iteration in a Data-Driven Manner

Actuaries are used to applying statistical and judgmental considerations during building



Al Replaces Iteration in a Data-Driven Manner

Actuaries are used to applying statistical and judgmental considerations during building




Actuaries are used to applying statistical and judgmental considerations during building



Actuaries are used to applying statistical and judgmental considerations during building



Actuaries are used to applying statistical and judgmental considerations during building



39

Actuaries are used to applying statistical and judgmental considerations during building



The Output Must be Adjustable By the Actuary

Without the ability to apply actuarial judgment, implementations of AI will be unsuccessful



Why Does the Magic Box Work?



Good Data



No Matter the AI, it Cannot Overcome Bad Input Data

There is no workaround for bad data



Minimising Data Issues is Essential for the Use of AI (Or use AI to detect bad data)



How to Fail at Using Al



This Magic Box is a Waste of Time

How do we fix this?





Make All the Correct Assumptions



Make <u>Reasonable</u> Assumptions



This Magic Box Works!

This user will accept AI





An Actuarial Magic Box

This user will accept AI



A Reserving Actuarial Magic Box

This user will accept AI





Recommendations for the Use of AI in Actuarial Practice



Remember Why We Iterate

Three major categories

















Apply Actuarial Judgment to Inputs and Outputs

You can't apply judgment within the magic box.





Creating a Large Range of Outputs



The Ideal Outputs

One example of a potential application of AI in actuarial work





The Ideal Outputs - N Dimensional Outputs

One example of a potential application of AI in actuarial work





You Need 5 Measures to Address Reserve Adequacy

What do you ask ReserveMax9000?

Hello ReserveMax9000, what are the top 5 measures I should take to address reserve adequacy?

Hello ReserveMax9000, what are the top 15 measures I should take to address reserve adequacy?



BAD



Create a "Too Large" Range of Inputs



A "Too Large" Range of Inputs

Help to ensure that you aren't missing appropriate outputs





A "Tool Large" Range of Inputs

All edges should be actuarially inappropriate

 $\blacksquare \textcircled{}$

1	Output	Output	Output	Output	Output	Output
	Output	Output	Output	Output	Output	Output
	Output	Output	Output	Output	Output	Output
	Output	Output	Output	Output	Output	Output
	Output	Output	Output	Output	Output	Output
	Output	Output	Output	Output	Output	Output
	Output	Output	Output	Output	Output	Output



Ask Yourself: Do I <u>Need</u> to Apply Actuarial Judgment To My Al Outputs?



Ask Yourself: <u>Can</u> I Apply Actuarial Judgment To My Al Outputs?



Black Box AI is Appropriate for Some Actuarial Analysis



Black Box Outputs

Can You Apply Judgment?

Hey ReserveMax9000, what are 15 actions I should take to avoid insolvency?

When the output is the creation of a range of results that can be independently evaluated and tested, black box AI may be appropriate.



Black Box AI is Inappropriate for Some Actuarial Analysis



Black Box Outputs

Can You Apply Judgment?

Hey ReserveMax9000, what's my range of reasonable reserves?

Hey ReserveMax9000, are you going to elaborate on this?

Can you Apply Judgment: No

No.



Black Box AI Should Not Be Used for a Pricing Estimate

If you can't explain or apply judgment to the estimate, you should not use it





Remember, Statistically Optimal is not Actuarially Optimal

This is the best statistical fit for this variable from a magic box output




Black Boxes Al as a Validation Tool





Used only as validation

Hey ReserveMax9000, what's your estimated reserves for this company?

Great, that's what I had. I'm confident in my estimate

Great, that's what I had. I'm confident in my estimate now so there's no need to do my normal validation



Black Boxes are Trusted Too Much

When AI is Incorrect, it biases "reviewed" results

Can incorrect artificial intelligence (AI) results impact radiologists, and if so, what can we do about it? A multi-reader pilot study of lung cancer detection with chest radiography

Michael H. Bernstein, #1,2,3 Michael K. Atalay, #1,3 Elizabeth H. Dibble,1

<u>Aaron W. P. Maxwell,^{1,3} Adib R. Karam,¹ Saurabh Agarwal,¹ Robert C. Ward,¹</u>

Terrance T. Healey,¹ and Grayson L. Baird

Conclusion

Incorrect AI causes radiologists to make incorrect follow-up decisions when they were correct without AI.

This is the actuarial equivalent of determining a company is insolvent and then changing the conclusion after AI consultation.



Reputational Risk is Enormous

The public will be unforgiving if AI is misused

Why were you confident in your reserve estimate and overlook this qualitative information?

ReserveBot9000 validated my assumptions!

Straight to Jail



Remember the Reputational Risk of Blindly Applying Al

The public will not forgive blind trust in AI



Thomas

They did a great job spot tuckpointing. It was great to go with a company that is focused on brickwork instead of a general contractor. There were a few missed spots and after I called, they came out ...

Response from the owner

ChatGPT said: ChatGPT Thank you for your feedback! We're glad you're pleased with our tuckpointing work and appreciate your understanding regarding the missed spots. We're happy to have promptly addr...



Explainable AI vs. Transparent AI



Producing Risk Scores: Transparency vs. Explainability

For example, a pricing model

	Transparency	Explainability
Level of Intuition	Easily understand exactly what the model is doing for each individual	Easily understand generally what the model is doing in aggregate
Example Models	GLM	GBM
Key Difference	Given a change to an item's characteristics, I know how the output score will change for this individual item without running the model.	Given a change to an item's characteristics, I must run the model to know how the output score will change for this individual item.

Local Transparency vs. Global Explainability for ReserveMax9000

For example, a reserving model

	Local Transparency: ReserveMaxTransparent	Global Explainability: ReserveMaxExplainable
Level of Intuition	Al outputs exactly how reserves, development, income determine the solvency estimate for each company.	Al outputs in general how reserves, development, and income when determine solvency in aggregate.
Example Reasoning	Exact calculation of solvency estimate is easily explained	Relative importance to solvency estimate is provided
Key Difference	Comparison between two similar is simple and easy to understand	Comparison between two similar may not be easily understood



Is Explainability Sufficient?



Adjusting Explainable Models like GBM

This does not work well.



Just multiply everyone by a constant and it'll work out, right?

人 AKUR8

ICE Graph: Yellow is PDP, Black Lines are Individual Risks

Example from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10235827/

Model behavior will likely not be smooth for individual risks, and cannot be made smooth.





What if we apply sufficient constraints to our explainable model?



Applying Sufficient Constraints May Not be Possible



Actuaries will Use Transparent Al for Most Use Cases







Actuaries Will use Black Box and Explainable AI to Create Transparent Output



Black Box Outputs

Can You Apply Judgment?

Hey ReserveMax9000, what are 15 actions I should take to avoid insolvency?

When the output is the creation of a range of results that can be independently evaluated and tested, black box AI may be appropriate.



If Actuarial Judgment is not Needed, Give it Completely to AI

If actuarial judgment is not needed, why are you doing this task?



An actuary must acknowledge the total lack of human intervention







