

ASSA Essay Competition 2026

Actuaries are all you need: Diffusing the mind of the actuary into the mind of AI

Introducing two minds with a shared perspective

The world of research has never been moving at a faster pace than it is right now. With the advent of Artificial Intelligence (AI), global research has entered a new stratosphere, largely through efficiency gains in generating new literature using Large Language Models (LLMs). However, there is a parallel contributor to this acceleration that is often underestimated, being research into advancing AI models themselves. The mind of AI is evolving by the day, as researchers push to break new frontiers in their quest to move from Artificial Intelligence to Artificial General Intelligence, and even to Artificial Super Intelligence.

There is risk that amongst the fast-paced evolution of AI models, the user's understanding of how these models work strays further away from the truth, and actuaries are by no means immune. The actuarial profession, albeit a very technical fraternity, is not as centred on understanding the technical architecture of AI as much as our data science and computer science counterparts. The future relevance of technical professions, such as our very own, may hinge on the ability to deeply understand how AI works and convey this to non-technical stakeholders. If actuaries lack this understanding at present day, this casts a looming shadow on the ability of our profession to stay relevant in the coming 10 years.

However, while the mind of AI continues to take shape, it is becoming increasingly clearer that it is converging on a conscience much closer to home: the mind of the actuary. There are abundant similarities between the cognitive processes underscoring both AI and the actuary, which creates a unique opportunity for actuaries in the era of AI. Actuaries are well-positioned to understand the complex mechanisms behind AI, be thought leaders in the future development of AI, and guide AI research into encompassing the very principles that paint the art of actuarial science. The mind of an actuary is by no means artificial, but it is certainly influential, and has a critical role in how AI develops over the next 10 years.

This essay focuses on two main themes:

- The model architecture behind modern AI is remarkably similar to the thought process and principles that underpin actuarial science.
- Actuaries must acquire a deep understanding of the mathematical architecture behind AI models if we are to thrive as a profession in the next 10 years.

Transforming AI from mystery to mathematics

In order for us to understand how actuarial thinking can be fed into the mind of AI, we must first explore the mind of AI itself. The brain behind modern generative AI is the transformer model, which was first unleashed through seminal research by Vaswani et

al. (2017) in their breakthrough paper “Attention Is All You Need”. The transformer model has taken the world by storm, and is applied most commonly in large language models (LLMs) by autoregressively predicting the next word in a sentence using previous words as context.

In reality, the true beauty of these models lies in the mechanics hidden beneath the surface. The transformer model is a confluence of elegant mathematics, ranging from vectors to probability distributions and even neural networks. You would expect these kinds of mathematical models to attract undivided actuarial interest, yet there seems to be a dissonance between actuaries and our understanding of the mathematical architecture behind the AI models we use on a daily basis.

However, in truth, actuaries are fully equipped to understand every last detail of how these models operate. While AI models may seem overly complex at first glance, they are being powered by simple matrix algebra beneath the surface. I get the feeling that actuaries perceive themselves as the inconsequential matrix algebra in the vast expanse of the mathematics behind AI, without realising that this is precisely where AI begins. In fact, the mind of AI and the mind of the actuary are more closely aligned than you may think.

Figure 1 below exhibits how the mathematical architecture of transformer models underpinning AI (as detailed in the “Attention Is All You Need” paper) can be parallelised to the cognitive architecture that actuaries use when solving problems. Each element of the transformer architecture can be mapped precisely to a step in the actuarial thought process, starting from data ingestion all the way through to final output. The following paragraphs explore key segments of this mapping in more detail, to exhibit how the mind of AI transforms into the mind of the actuary.

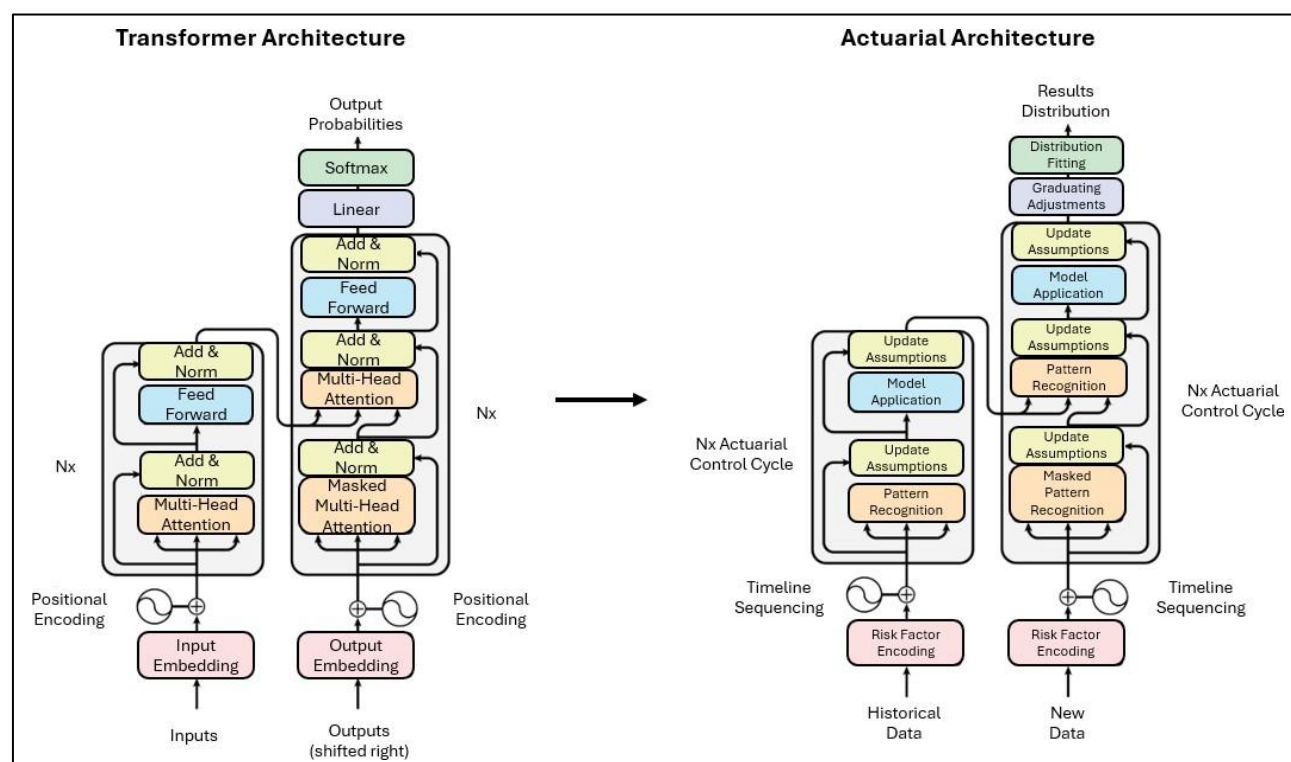


Figure 1: Comparison between transformer and actuarial architecture.

Embedding actuarial understanding in the AI context

A well-trained transformer starts with input tokens, most commonly sequences of words in a sentence. Each input token is mapped to an embedding vector, which is essentially a unique fingerprint that encapsulates base-level meaning of that word. The transformer then applies positional encoding to each embedding to help the model understand the original order of the words in the input sentence.

This is no different to any well-trained actuary, whose input tokens could be represented by policyholders in a historical dataset, for which an embedding can be formed through the actuary's holistic knowledge of their demographics, health status and other risk factors. As actuaries, we are taught from outset in the A211 Financial Mathematics exam to begin any solution by plotting the sequence of ordered events on a timeline, our homegrown form of positional encoding.

Hence, we observe that context is a cornerstone of both the transformer and actuarial architectures. Transformers require context to get an initial sense for what might come next in a sentence, while actuaries require context to anchor their forecasts on what might manifest in the financial world of the future. Actuaries are well placed to convey the importance of context in AI models to non-technical stakeholders, particularly given that we acknowledge that the ordering of context matters. Intelligent prompting of AI models is a clear application of our competitive advantage in this respect, by applying the most coherent sequential order to input context, so as to optimise the quality of response from the AI models we use.

Actuarial attention is the key to solving any query

The centrepiece of the transformer model, and the main reason for its dominance in driving AI, is a mathematical concept called attention. While transformers see the entire history of context that is inputted, these models learn exactly which words from the past are most important for predicting the words to come, almost paying attention to the words that truly matter. This is the genius of the transformer, and is powered through query vectors that know what information to look for, and key vectors that know what information to offer in return. Transformers repeat these matches over numerous attention layers in the model, ensuring that the semantic meaning of each word in the input can be gradually updated as the model learns more from the words that surround.

The concept of queries, keys and attention layers are in essence a reconfiguration of our very own actuarial control cycle. Our queries are established when we initially define the problem, unearthing the critical pieces of information that characterise the task at hand. Our keys are engineered once we develop the solution, piecing together each step towards our final destination. The true artistry of the actuary lies in the confluence of these queries and keys, knowing exactly what information from our problem should be given most attention when constructing our solution. Just like a transformer's numerous attention layers, we complete the cycle by iterating through this process multiple times when monitoring experience, ensuring that our ecosystem remains robust and relevant in an evolving environment around us. It is this element of attention that sets actuaries apart: our ability to filter through the myriad of chaotic information at our disposal, and focus on the hidden patterns that unveil the

solutions within.

Diffusing actuarial clarity amongst the AI noise

The link between the mind of AI and the mind of the actuary is now well-established, however the question still remains: how can actuaries capitalise on this similarity to thrive in an AI-driven world? As an avenue to predict how the role of actuaries will transform in line with emerging AI models, we can analyse the reaction of actuaries to existing AI concepts in times past.

Figure 2 below exhibits the first instance of key machine learning/AI terms appearing in an academic paper in the global AI research context, as well as in a paper published in the South African Actuarial Journal (SAAJ). While it is important to note that there is a censoring effect given that SAAJ only dates back to 2001, it is intriguing nonetheless to compare the lag between global AI research and South African Actuarial research in exploring these key AI terms in academic literature. The key finding is that while South African actuarial research was a slow adopter of early machine learning/AI models (notably GLMs, random forests and neural networks), this gap has narrowed significantly for newer models (including transformers through Richman et al. (2024) and large language models through Balona (2023)).

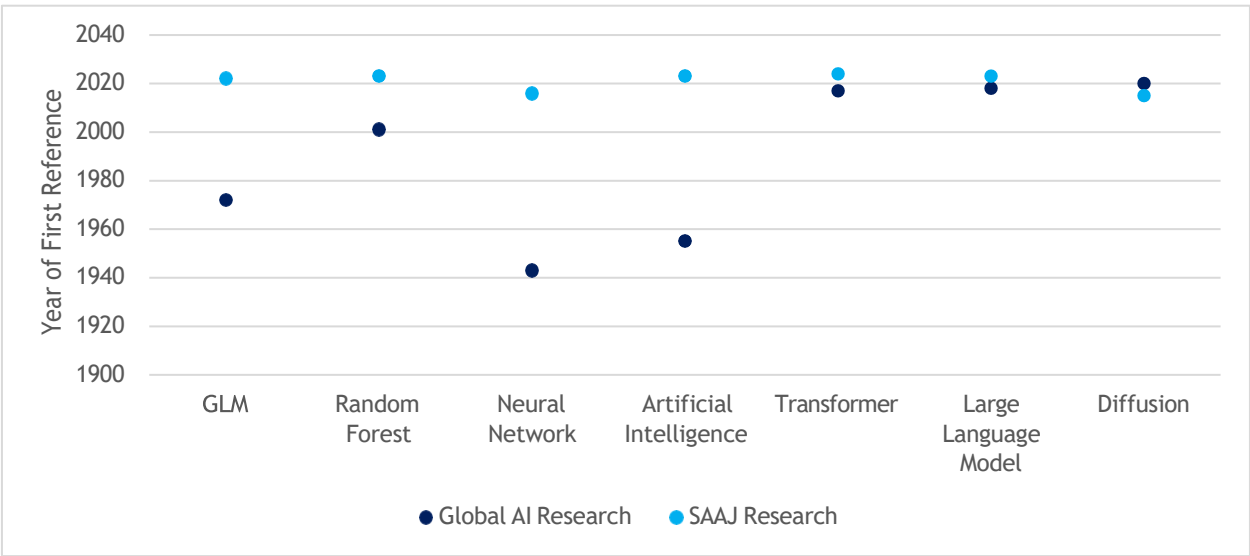


Figure 2: Year of first reference to key machine learning/AI terms in global AI research and South African Actuarial Journal.

However, critically, we observe an inverted lag for diffusion models. Diffusion models have been central to actuarial science for many years with clear applications in stock price forecasting, as exhibited by Hong & Zou (2015) in their paper on diffusion-based semimartingale models (the first evidence of diffusion models in South African actuarial research). However, diffusion models have only made a relatively recent foray in global AI research, most pertinently by Ho et al. (2020) in their research titled “Denoising Diffusion Probabilistic Models”, which forms the core backbone of AI image and video generation as we know it.

Diffusion models themselves represent another clear parallel between the mind of AI and the mind of an actuary. In the generative AI context, diffusion models start with

random noise, an image comparable to a snowy television screen. The diffusion model then incrementally denoises this image by reversing the process of Brownian Motion, in order to generate a crystal clear image based on the input prompt. In the actuarial context, we too often start with random noise, a financial world characterised by chaos and data that appears convoluted on the surface. However, step by step, we incrementally decode this noise, to uncover the crystal clear image of the scenario that lies beneath, and develop an actuarially sound strategy for the future. The diffusion model analogy aligns perfectly with ASSA's fresh slogan of "Clarity Amidst Uncertainty", where us actuaries act as the diffusion models ourselves, guiding our stakeholders from the initial image of uncertainty, towards the targeted output of clarity.

This diffusion model anecdote foreshadows a clear opportunity for actuaries to influence the future of AI. It is undoubted that AI research will continue to progress quickly in the next 10 years, with a multitude of "unborn" AI models still yet to be added to the horizontal axis of Figure 2 above. There is a clear risk that if actuaries do not have a deep mathematical understanding of existing AI models today, then it will be difficult for us to keep pace with emerging mathematical techniques underpinning future AI models. This may result in an accelerated reduction in our relevance to dictate conversation around the use of AI. However, parallel to this risk lies an even larger opportunity. Actuaries undoubtedly have the technical capability to be one of the only professions to develop our very own mathematical AI model infrastructure, completely novel to the rest of the AI fraternity. We can shift the perspective from actuaries adapting existing AI models to actuarial problems, to leading the way by getting AI researchers to weave actuarial principles into the backbone of emerging AI model infrastructure. Within the next 10 years, the actuarial education syllabus must extend beyond traditional statistical foundations, towards including new AI mathematical concepts as they emerge. This will allow the actuarial profession to climb the proverbial AI food chain, and pivot from being end users of AI, to being key creators of AI itself.

Beware discounting the value of the actuary in the era of AI

While this essay is themed around the role of actuaries in forging new paths for future AI development, I would like to conclude by paying homage to traditional actuarial principles, particularly casting light on our professional responsibility in an AI-driven world. The future of AI in South Africa is akin to an actuarial discounted cashflow model, with two concurrent forces at play: the rate of growth at which AI tech giants elevate the capability of their models, and the rate of discount at which the general population can make best use of these models. If the rate of growth exceeds the rate of discount, we risk witnessing the widening of the existing equity gap in South Africa, where entities with AI access advantages pull away from the general population, potentially including the actuarial profession.

However, if the converse effect manifests, where the rate of discount exceeds the rate of growth, we unleash the opportunity for all South Africans to leverage AI on a democratised playing field, and eliminate any remaining national inequity gaps for good. Given our technical nous to understand the internal training of AI models, coupled with our professional commitment to ensuring that the very same training of

these models is immune to any inherent bias, actuaries belong as a critical variable in this discounted cashflow equation. The actuarial profession has a distinct role to play in controlling these rates of growth and discount, to ensure that South Africa, and our very own profession, thrives in an AI-driven world for not just the next 10 years, but rather in perpetuity. (Word Count: 2,312)

Note: The opinions expressed in this essay are those of the individual author and do not necessarily reflect the views of the Actuarial Society of South Africa. The author notes that AI was used for research purposes when compiling evidence for this essay.

References

Balona, C (2023). ActuaryGPT: Applications of large language models to insurance and actuarial work. *Proceedings of the Actuarial Society of South Africa's 2023 Convention*, Sandton Convention Centre, 11–12 October 2023.

Breiman, L (2001). Random forests. *Machine Learning* 45(1), 5–32.

Devlin, J, Chang, M-W, Lee, K & Toutanova, K (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*.

Ho, J, Jain, A & Abbeel, P (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* 33, 6840–6851.

Hong, L & Zou, J (2015). Jump tests for semimartingales. *South African Actuarial Journal* 15(1), 93–108.

McCarthy, J, Minsky, ML, Rochester, N & Shannon, CE (1955). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Reprinted in *AI Magazine* 27(4), 12–14 (2006).

McCulloch, WS & Pitts, W (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5(4), 115–133.

Nelder, JA & Wedderburn, RWM (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A (General)* 135(3), 370–384.

Richman, R & Wüthrich, MV (2022). LASSO regularisation within the LocalGLMnet architecture. *Proceedings of the Actuarial Society of South Africa's 2022 Convention*, Cape Town, 25–28 October 2022.

Richman, R, Scognamiglio, S & Wüthrich, MV (2024). The credibility transformer. *Proceedings of the Actuarial Society of South Africa's 2024 Convention*, Cape Town International Convention Centre, 20–21 November 2024.

Siriram, C & Harris, R (2023). Applications of advanced analytics in healthcare. *South African Actuarial Journal* 23(1), 1–22.

Smith, ML, Beyers, FJC & de Villiers, JP (2016). A method of parameterising a feed forward multi-layered perceptron artificial neural network, with reference to South

African financial markets. *South African Actuarial Journal* 16(1), 35–67.

Vaswani, A, Shazeer, N, Parmar, N, Uszkoreit, J, Jones, L, Gomez, AN, Kaiser, Ł & Polosukhin, I (2017). Attention is all you need. *Advances in Neural Information Processing Systems* 30, 5998–6008.