BY RAQUEL HUMPHRIS, JAMES MORAN
AND WOUTER KARMAN

The actuary, a key player in achieving and demonstrating fair AI models

Artificial intelligence (AI) is becoming more and more important in the daily activities of insurers, who are increasingly placed under a social magnifying glass for the way they implement algorithms. With model developments like ChatGPT, interest in and development of AI models is peaking and exposure to the magnifying glass increases further. When AI-models are not properly designed, developed, monitored, and used, this can have serious consequences on peoples' lives, including financial exclusion and discrimination. For example, increasing premiums to unaffordable levels for people with bad credit due to lower education worsens their financial situation further. The EU AI act is the major EU legal initiative on data and AI that is intended to prevent AI models from infringing on human rights, health, and safety. The EU AI act will also have an impact on insurance companies. The question arises where this affects actuaries and where they fit in to assure that customers are treated fairly and ensure compliance of AI models. In this article, we discuss how the introduction of the EU AI Act affects insurance companies, with a particular focus on fairness, as well as what you can do as an actuary and how you can leverage your broad skill set and domain knowledge.

A EUROPEAN APPROACH TO AI

The EU AI Act could become the global standard to determine to what extent AI can affect your life, similar to how the General Data Protection Regulation (GDPR) became the standard for data privacy. Insurance companies need to be conscious here, as AI is used to perform tasks such as underwriting, fraud detection, claims processing, and credit approvals. Therefore, the EU AI Act will have major implications on their business processes and they need to be aware of how to comply with the standard.

We expect the biggest impact by the EU AI Act for the life and health insurance sector due to being explicitly noted as high-risk areas. A high-risk classification results in requirements on, for example, risk management, registration, monitoring, and fairness and discrimination. These requirements demand extensive knowledge on statistics, modelling and testing, combined with a deep understanding of the business and the involved products and processes. These requirements need to be fulfilled and verified in all three lines of defence to provide the organisation with the required comfort. In other words, with the introduction of the EU AI Act, actuaries are highly likely to obtain a new and crucial role to provide the necessary comfort to the organisation that it complies with legislation.

CURRENTLY, LIMITED REGULATIONS REGARDING ETHICS AND FAIRNESS IN MODELLING ARE IN PLACE FOR INSURERS

THE FAIRNESS COMPONENT OF THE EU AI ACT

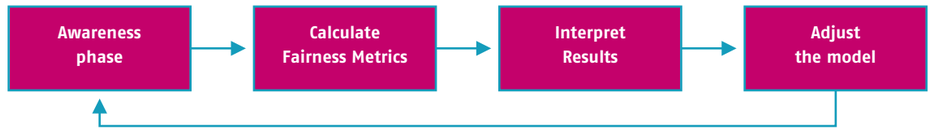
One key element to the EU AI Act are the requirements on discrimination prevention and fairness. As AI is transforming the insurance industry, businesses should implement algorithm fairness metrics and unfairness mitigation approaches to minimise the risk of AI algorithms picking up bias in historical data or from improper modelling. Also, to ensure the technology is being used fairly and ethically, businesses should explain where the data is coming from, how it is used, and how AI is deployed within the organisation. Currently, limited regulations regarding ethics and fairness in modelling are in place for insurers. While new, more complex models are being developed, the EU AI act is exactly offering this. Meaning, next to the business and moral process, a legal perspective on fairness is added to the activities provided by the actuary, especially for insurance products that involve sensitive features where age-discrimination is allowed, like car or life insurance. However, fairness evaluation depends on the purpose of the model, the place of the model in the organisation, the company ethics, client and society ethics, and the data and model quality, and therefore is highly complex and requires a deep understanding.

IN PARTICULAR, THEY NEED TO BECOME AWARE OF HOW TO MEASURE FAIRNESS

The new requirements mean that actuaries building AI models in collaboration with other experts on data, legal and operations, or reviewing/validating models as part of 2nd/3rd line teams, will need to upskill themselves. They will need to ensure they can understand the new type of models, their potential, risks and limitations, the outputs produced, and the testing performed. In particular, they need to become aware of how to measure fairness and make models and model outcomes explainable and transparent to a large variety of stakeholders. In order to take on this task, new skills and tools should be developed by actuaries in all three lines of defence. While the first line needs to focus on providing these insights as part of the design, the second and third line need to have the tools to demonstrate and validate transparency and fairness of the high-risk models. The second and third lines should prevent being surprised with new techniques by the first line and therefore need to take a proactive approach in developing their skill set and toolbox.

HOW TO EVALUATE MODEL FAIRNESS

In the process of evaluating and adjusting a model for fairness, one must understand the model, how it uses the data provided, and any possibly unexpected biases it may have learned. A typical cycle to evaluate fairness is illustrated below.



To assess whether the AI model has a bias, fairness metrics on, for example, gender and age, can be calculated. Although it is straightforward to calculate the fairness metrics, choosing the correct metric, scope and tolerance requires deep knowledge of the entire environment of the model, including ethical and strategic considerations of the organisation as a whole. A best practice observed in the financial industry is to create a framework and guidelines, based on the dimensions described above, that guides users and developers towards a recommended metric and scope. The results of the chosen metric and scope will indicate whether the model is (significantly) biased towards the identified groups, which could be stemming from one or more of the adopted variables. The model should be adjusted whenever it contains unacceptable bias, after which the evaluation starts over again. Bias mitigation techniques include adding or changing weights within the training data to remove bias, changing the model to one that considers fairness directly (e.g., Adversarial Learning), or adjusting the model predictions to reduce bias directly during optimization.

EXAMPLE APPLICATION

Insurers often have limited data about policyholders, especially new ones. They may know the age, gender, the insured object (in case of non-life insurance), self-reported information about health (life and health insurance) and the postal code. Additional data on the underlying risk is often unavailable. Data enrichment can provide access to additional data, allowing models to better estimate the underlying risks and limit the biases in available data.

Postal code is often available, but using the code itself can create proxies to sensitive attributes, such as ethnicity. However, the postal code can be used to gain additional insight into the residential environment of the policyholder. Instead of using the actual postal codes, one could cluster them to reduce the risk of discrimination. When clustering the postal codes on comparable underlying statistics, these clusters can be used to gain additional insight into the policyholder while using an appropriate data source. Attributes to cluster postal codes could for example be the socio-economic class, safety, environment, or the facilities in the residential area. Even after applying a mitigating technique like postal code clustering, you need to consider and evaluate the fairness of this approach and validate that you are not creating an unjustified and unintentional treatment for particular groups of people. For example, socio-economic class is a good risk factor for mortality, but can be correlated with ethnicity. In this case, one needs to ensure that the socio-economic class risk factor is captured accurately and that different ethnic groups are not treated unfairly.

CONCLUSION

The EU AI act is on its way and will influence the use of AI and statistical models. Models used in life and health insurance for risk assessment and pricing are explicitly considered 'high-risk', along with models for creditworthiness. There needs to be an increased awareness and capability on the evaluation of models by actuaries in all three lines of defence, as the regulation may affect many use cases. The actuary can become a key player to provide comfort and assurance to management and the organisation as a whole, as they are one of the few groups of experts who have the knowledge and expertise to assess AI models simultaneously on business impact, performance, ethics and fairness. To claim this role, actuaries need to upskill themselves and become familiar with AI modelling and fairness evaluation. ■

R. Humphris MSc (left) is a Risk Modelling Services Senior Associate.

J. Moran BSc (Hons) (middle) is a Risk Modelling Services Senior Associate.

Dr. W. Karman is a Risk Modelling Services Manager.

They all work at PwC the Netherlands.