Fair and Transparent Digital Differentiation in Car Insurance.

Sara Tulimat
HU University of Applied Sciences Utrecht, the Netherlands

Abstract. Artificial Intelligence (AI) is transforming car insurance through digital differentiation (DD), using machine learning to assess risk and set premiums. While improving efficiency, AI may unintentionally reinforce biases, such as postal codes acting as proxies for protected attributes. This study explores how the 4D methodology (Detect Discrimination in Digital Differentiation), combined with Fairlearn and AIF360 helps insurers detect and mitigate bias. By integrating bias detection, fairness optimization, and evaluation techniques insurers can ensure fair yet profitable premium differentiation, reducing discrimination while maintaining economic viability.

INTRODUCTION

Artificial intelligence (AI) is playing an increasingly significant role in the financial sector. Insurers use digital differentiation (DD) to assess individual financial risks through self-learning algorithms. This influences premium pricing and acceptance decisions. While this enhances efficiency and profitability, it also introduces risks such as unintended discrimination. Algorithms can reinforce existing social inequalities, for example, through the use of postal codes.

This study explores how the 4D methodology (Detect Discrimination in Digital Differentiation) can be integrated into insurance processes to ensure fair and profitable premium differentiation. The focus is on implementing machine learning techniques, including optimization algorithms and dataset adjustments, to strike a balance between fairness and profitability.

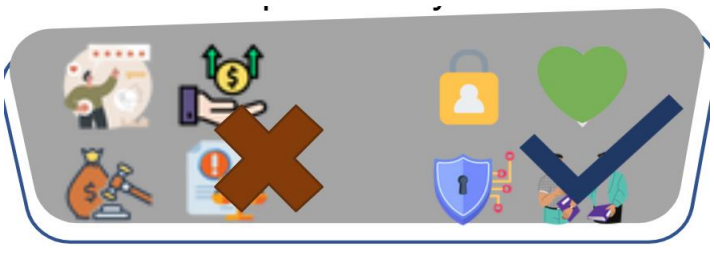
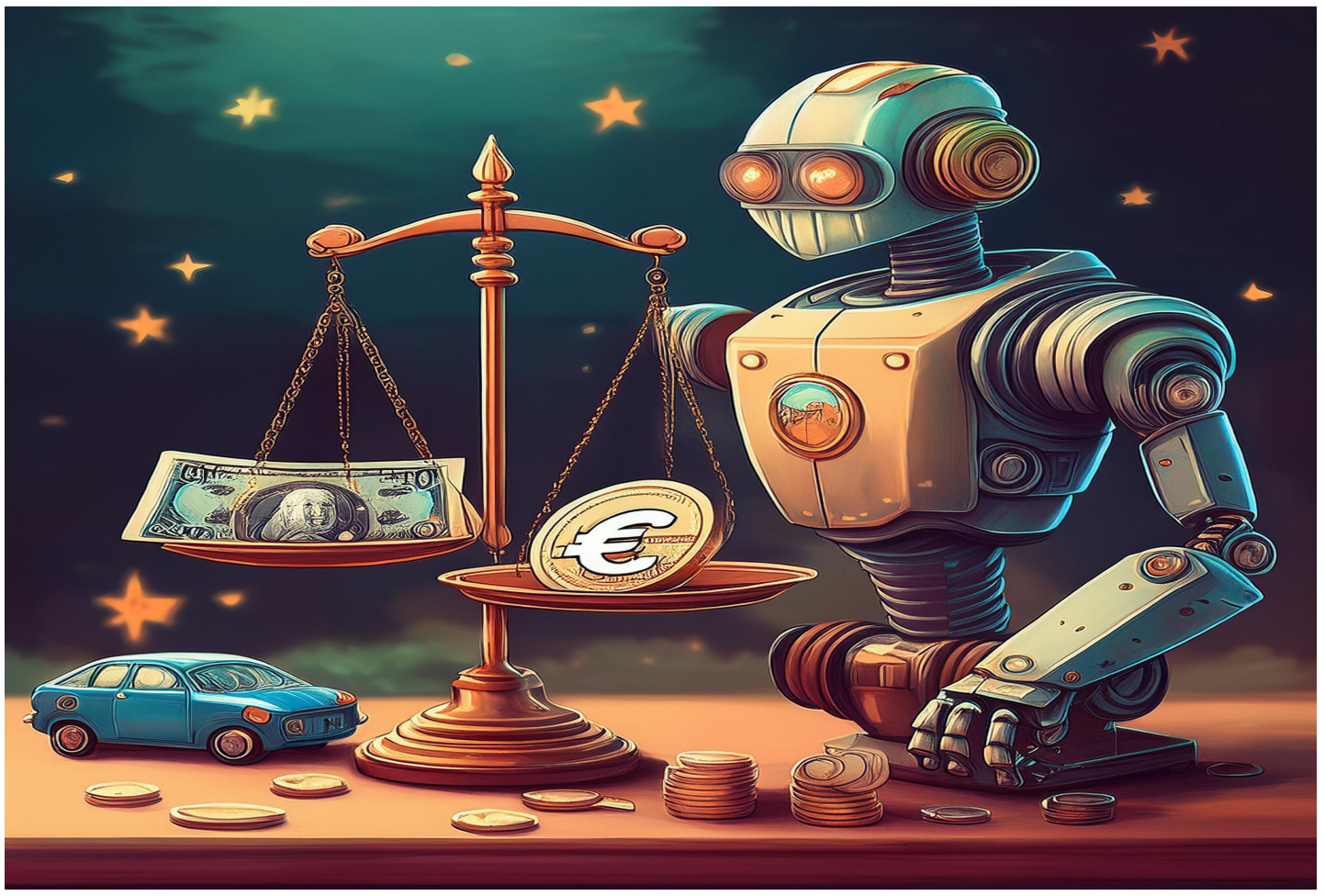


Fig. 1. Instead of facing high costs, reputational damage, and potential lawsuits, the 4D method ensures fairness, inclusivity, and ethical practices.

LITERATURE REVIEW

AI is transforming car insurance through digital differentiation (DD), using machine learning to assess risk and personalize premiums. While improving efficiency and profitability, AI raises fairness concerns, as models may unintentionally reinforce biases in historical data, with factors like postal codes.



To address this, fairness-focused AI frameworks such as Fairlearn and AIF360 provide tools to detect and mitigate bias. Fairlearn offers techniques like demographic parity and equalized odds, enabling insurers to measure and adjust fairness constraints in pricing models AIF360, developed by IBM, includes bias mitigation algorithms and explainability tools to ensure transparent decision-making .

While existing research highlights the challenges of balancing fairness and profitability, practical applications remain limited. Integrating Fairlearn and AI360 into insurers' workflows can ensure fair yet data-driven premium differentiation, reducing bias while maintaining economic viability.

RESEARCH METHODOLOGY

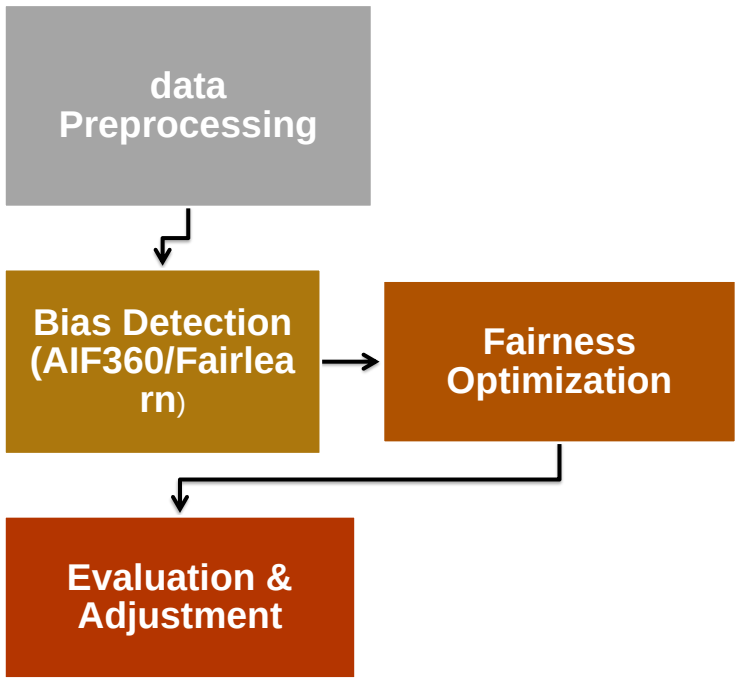


Fig. 2. The research method to balance fairness and profit involves using fairness metrics, bias detection, and optimization techniques.

This research uses a mixed-method approach, combining literature review, data analysis, and model evaluation. It examines discrimination in auto insurance, preprocesses data to mitigate bias, and applies AI fairness toolkits (AIF360/Fairlearn). Strategies balancing profitability and fairness are explored, and model evaluation assesses bias mitigation using accuracy, fairness, and profitability metrics.

CONCLUSIONS

Ensuring fairness in AI-driven insurance requires integrating bias detection and enables insurers to achieve fair and data-driven premium differentiation. mitigation tools into pricing models. Fairlearn and AIF360 offer practical solutions to measure and correct bias while maintaining economic viability. A structured workflow spanning data collection, preprocessing, bias detection, optimization, and evaluation

References.
CBS. (2021, oktober 15). ICT-gebruik bij bedrijven - ICT, kennis en economie 2021. ICT-gebruik Bij Bedrijven - ICT, Kennis en Economie 2021 | CBS. https://longreads.cbs.nl/ict-kennis-en-economie-2021/ict-gebruik-bij-bedrijven/
Zuiderveen Borges, F. J. (n.d.). *Digitale discriminatie en differentiatie: Het recht is er nog niet klaar voor.* Retrieved from https://pure.tue.nl/ws/portalfiles/portal/347507864/23-0389.pdf
Fairlearn. (n.d.). *Fairness in machine learning.* Retrieved from https://fairlearn.org/
Springer. (2022). *Digital discrimination and fairness in insurance.* Retrieved from https://link.springer.com/article/10.1007/s43681-022-00138-8/figures/1
Verbond van Verzekeraars. (2023). *Balanceren tussen solidariteit en differentiatie.* Retrieved from https://www.verzekeraars.nl/media/11441/balanceren-tussen-solidariteit-en-differentiatie-26-april.pdf

Acknowledgements. The HCAIM (the Human-Centred AI Master's Programme) Project is Co-Financed by the Connecting Europe Facility of the European Union Under Grant №CEF-TC-2020-1 Digital Skills 2020-EU-IA-0068. This poster was created as part of the Blended Intensive Programme organized under the Erasmus + Programme of the European Union