Fairness-Aware Causal Active Learning for Genomics AI (FACA-GEN)

Amber Qayum Hawabaz *1

*Equal contribution [1]*Independent Researcher*. Correspondence to: Amber Qayum Hawabaz amberhawabaz54@gmail.com.

1. Introduction

AI-driven genomics has transformed precision medicine, yet fairness remains a critical challenge. Existing **Active Learning (AL)** techniques in Genomics AI prioritise informativeness, leading to biased sample selection and under-representation of certain genetic groups. Traditional fairness-aware AI methods fail to address this issue, either imposing rigid constraints that harm accuracy or neglecting hidden biases in genetic data. We propose **Fairness-Aware Causal Active Learning for Genomics AI (FACA-GEN)**, a novel AL framework designed to:

- Balance **informativeness and fairness** dynamically using a **multi-arm bandit approach (MAB)**.
- Integrate **Causal Representation Learning (CRL)** to prevent shortcut learning from confounding factors.
- Optimise fairness trade-offs through **reinforcement learning (RL)** to improve generalisation without reducing accuracy.

Experiments on the **IGSR genomics dataset** show that **FACA-GEN reduces demographic bias by 45% (Demographic Parity improvement) and improves Equalised Odds by 70%** compared to standard AL methods. These results highlight the potential of fairness-aware AL to mitigate bias while maintaining predictive performance in Genomics AI.

2. Methodology

Traditional Active Learning (AL) methods prioritise information accuracy but overlook fairness, leading to biases in Genomics AI. FACA-GEN addresses this by integrating **fairness-aware sample selection** and **causal representation learning**, ensuring equitable data selection while preserving accuracy.

2.1 Key Methodology

FACA-GEN employs a **Multi-Armed Bandit (MAB)** approach for fairness-aware sample selection. At each AL step, the framework selects training samples based on three criteria:

- **Informativeness**: Samples with high uncertainty, measured using entropy-based AL.
- **Fairness**: Samples that reduce demographic bias, quantified using **Demographic Parity Difference**.
- **Causal Consistency**: Ensures AI learns true genetic markers rather than confounding factors such as ancestry.

To optimise these trade-offs, FACA-GEN employs **Reinforcement Learning (RL)** to dynamically adjust fairness constraints during training. This prevents fairness-accuracy trade-offs, ensuring models generalize well across diverse populations.

2.2 Evaluation Metrics

FACA-GEN is evaluated on the **IGSR genomics dataset**, using fairness metrics such as:

- **Demographic Parity (DP)**: Measures fairness in sample selection.
- **Equalized Odds (EO)**: Evaluates bias in prediction errors.
- **Optimal Transport Distance (OTD)**: Assesses distributional fairness.

Results show **a 45% reduction in DP bias and 70% improvement in EO**, demonstrating FACA-GEN's effectiveness in mitigating bias without sacrificing accuracy.

2.3 Related work

Active Learning (AL) is widely used in Genomics AI to optimize data selection, but traditional methods prioritize informativeness over fairness, leading to biases in underrepresented populations [1]. Prior fairness-aware approaches in AI [2, 3] attempt to mitigate bias but often impose rigid constraints that degrade model accuracy. Moreover, existing fairness-aware AL methods fail to address causal dependencies in genomic data, which can lead to shortcut learning from confounding factors like ancestry rather than true genetic markers [4]. Real-world consequences of fairness failures in AI-driven decision-making have been observed in domains such as healthcare [5] and recruitment [6], where biased AI models reinforce systemic disparities. Addressing these limitations, FACA-GEN integrates fairness-aware AL with Causal Representation Learning (CRL) to prevent reliance on spurious correlations and uses Reinforcement Learning (RL) to dynamically optimize fairness constraints. Unlike previous works that treat fairness as a static constraint, FACA-GEN actively balances informativeness, fairness, and causal consistency, ensuring robust and equitable model training in genomics.

2.4 Results and Tables

The performance and fairness improvements of **FACA-GEN** over traditional Active Learning (AL) methods are summarized in

Table 1.

Table 1: Fairness evaluation of FACA-GEN compared to standard AL methods. (While OTD slightly increases, this is a known trade-off in fairness-aware selection, as it rebalances underrepresented samples.)

Metric	Before AL	After FACA-GEN	Improvement
Demographic Parity (DP)	0.2429 ± 0.013	**0.1333 ± 0.009**	45% ↓
Equalized Odds (EO)	1.0000 ± 0.000	**0.3011 ± 0.022**	70% ↓
Optimal Transport Distance (OTD)	0.0072 ± 0.001	**0.0097 ± 0.001**	Slight ↑

References

[1] Burr Settles. From theories to queries: Active learning in practice. In *Active learning and experimental design workshop in conjunction with AISTATS 2010*, pages 1–18. JMLR Workshop and Conference Proceedings, 2011.

[2] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6):1–35, 2021.

[3] Alekh Agarwal, Alina Beygelzimer, Miroslav Dudík, John Langford, and Hanna Wallach. A reductions approach to fair classification. In *International conference on machine learning*, pages 60–69. PMLR, 2018.

[4] Bernhard Schölkopf, Francesco Locatello, Stefan Bauer, Nan Rosemary Ke, Nal Kalchbrenner, Anirudh Goyal, and Yoshua Bengio. Toward causal representation learning. *Proceedings of the IEEE*, 109(5):612–634, 2021.

[5] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.

[6] Jeffrey Dastin. Amazon scraps secret ai recruiting tool that showed bias against women. In *Ethics of data and analytics*, pages 296–299. Auerbach Publications, 2022.