Universidade de Pernambuco
Programa de Pós-Graduação em Engenharia da Computação (PPGEC)

Proposta de Dissertação ou Tese de Mestrado ou Doutorado

Área: Computação Inteligente

Título: Leveraging Multi-Objective Optimization and Explainable AI for Fair Decision-Making Systems

Orientador(a): Dimas Cassimiro do Nascimento Filho (dimas.cassimiro@ufape.edu.br)
Co-Orientador(a): Bruno Costa e Silva Nogueira (bruno@ic.ufal.br)

In modern decision-making pipelines, Artificial Intelligence (AI) and Optimization techniques have become essential tools across a wide variety of domains. For example, in credit scoring, AI models predict default risk and optimization methods allocate capital under regulatory constraints; in healthcare, decision support systems use optimized models to recommend treatments while considering fairness across demographic groups; and in resource allocation or scheduling, optimization algorithms decide how to distribute scarce resources optimally, guided by predictive models. However, as these systems increasingly affect people's lives, there is a growing concern that purely utility-driven optimization may lead to unfair or opaque decisions, especially when sensitive attributes like race, gender, or socioeconomic status are involved.

The main goal of this PhD research project is to develop a unified framework that leverages optimization and AI to support decision-making in a way that ensures both fairness and explainability. Specifically, the objectives are to design multi-objective optimization algorithms that jointly optimize predictive performance, different formal fairness metrics (e.g., group and individual), and interpretability or model simplicity; to embed explainable AI techniques such as feature importance, counterfactual explanations, or procedural explanations into the optimization loop; and to evaluate how this framework performs in realistic decision-making domains (e.g., credit scoring, hiring, healthcare, public spending analysis, fraud detection in financial investigations, school admissions, and resource allocation in public services) under different fairness-explainability trade-offs.

This research is expected to contribute both theoretically and practically. Theoretically, it will propose new optimization formulations that explicitly incorporate explanation quality or model transparency as part of the objective, alongside accuracy and fairness. It will also develop algorithms (e.g., based on Pareto optimization or multi-objective evolutionary/swarm methods) to find optimal trade-off frontiers. Practically, the framework will be validated on real-world datasets, showing how decision-makers can navigate trade-offs between utility, fairness, and interpretability. The work will also generate insights into how stakeholders perceive fairness when explanations are integrated into optimized decision systems.

Referências Bibliográficas:

[1] Amarasinghe, P. T. et al. (2024). Evolutionary multi-objective optimisation for fairness-aware self-adjusting memory classifiers. ACM Transactions on Autonomous and Adaptive Systems.

[1] Nagpal, R., Shahsavarifar, R., Goyal, V. & Gupta, A. (2024). Optimizing fairness and accuracy: a pareto-optimal approach for decision-making. AI and Ethics, 5, 1743–1756.

[2] Pham, D., Tran, B., Nguyen, S., Alahakoon, D. & Zhang, M. (2024). Fairness optimisation

with multi-objective swarms for explainable classifiers on data streams. Complex & Intelligent Systems, 10, 4741–4754.

[3] Yu, G., Ma, L., Wang, X., Du, W., Du, W. & Jin, Y. (2024). Towards fairness-aware multi-objective optimization. Complex & Intelligent Systems, 11(1), 50.

[4] Zhao, Y., Wang, Y. & Derr, T. (2023). Fairness and Explainability: Bridging the Gap Towards Fair Model Explanations. Proceedings of the AAAI Conference on Artificial Intelligence, 37(9), 8888–8895.

[5] Kamani, M. M., Haddadpour, F., Forsati, R. & Mahdavi, M. (2022). Efficient Fair Principal Component Analysis. Machine Learning, 111(10), 3671–3702.

[6] Heitz, C., Hannák, A. & Baumann, J. (2022). Enforcing Group Fairness in Algorithmic Decision Making: Utility Maximization Under Sufficiency. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), ACM, 2315–2326.

[7] Pan, W., Cui, S., Bian, J., Zhang, C. & Wang, F. (2021). Explaining Algorithmic Fairness Through Fairness-Aware Causal Path Decomposition. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, ACM, 1287–1297.