Breakout Session 8: Track A

Ethically Optimize Machine Learning Models with Real-World Data to Improve Algorithmic Fairness

Dr. Xiaoqian Jiang

Professor and Chair, University of Texas Health Science Center at Houston

Finding combinatorial drug repositioning therapy for Alzheimer's disease and related dementias

– Ethically optimize machine learning models with real-world data to improve algorithmic fairness

Xiaoqian Jiang (PI)

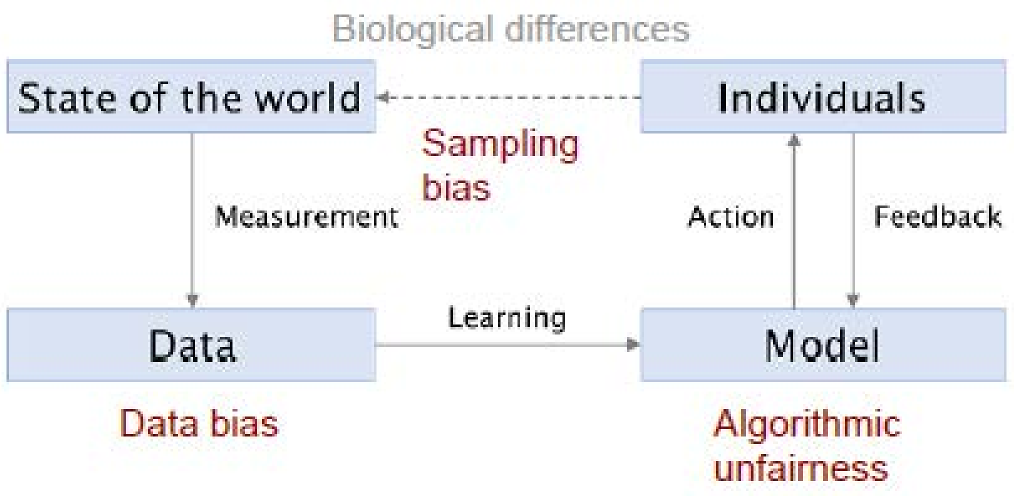
- UTHealth Houston

Introduction

- **Fairness in Healthcare: A Paramount Issue**
  - Systemic disparities in healthcare, including unequal access and allocation of resources, remain a paramount issue. The progress of Deep Learning in healthcare brings enhancements but also raises fairness issues.
- **Enhancements Brought by Deep Learning**
  - Improves personalized medicine,[1] predicts risk,[2] and aids in the development of treatments.[3]
  - Depends on real-world data that may be biased, mirroring historical injustices.[4]
- **Strategies for Achieving Fairness:** Identifying sources of bias is crucial in shaping strategies for achieving fairness:
  - **During Data Collection:** Pre-processing algorithms address biases in the input data.[5]
  - **In the Model Training Phase:** In-processing algorithms strive for a balance between accuracy and fairness.[6]
  - **For Decision-Making:** Post-processing algorithms adjust outputs to meet fairness standards.[7]

Summary

- Launched in September 2022 to enhance fairness in healthcare AI/ML.
- Aims to mitigate data biases affecting minority and underserved communities.
- Prevents perpetuation of health disparities through biased AI/ML predictions.
- Develops bias mitigation and ethical ML training methodologies.
- Strives for reduced health disparities and better patient outcomes.



We propose to develop novel methodologies to address potential algorithmic unfairness and ameliorate health disparity in machine learning models.

This project will address data bias by transforming and disentangling data (aim 1) and addressing model bias algorithmically by fair optimization for various parity metrics and equal opportunity (aim 2).

Achievements

Our research has led to AMIA Student Best Paper Runner-Up, AMIA CRI Best Student Paper award, and publications in the Journal of Biomedical Informatics (JBI)

Our work led to receiving the AIM-AHEAD award for addressing bias and AI fairness issues in heart transplantation.

Ding S, Tang R, Zha D, Zou N, Zhang K, Jiang X, Hu X. Fairly Predicting Graft Failure in Liver Transplant for Organ Assigning. AMIA Annu Symp Proc 2022;2022:415–424. PMID:37128420 (Best Student Paper Runner-Up)

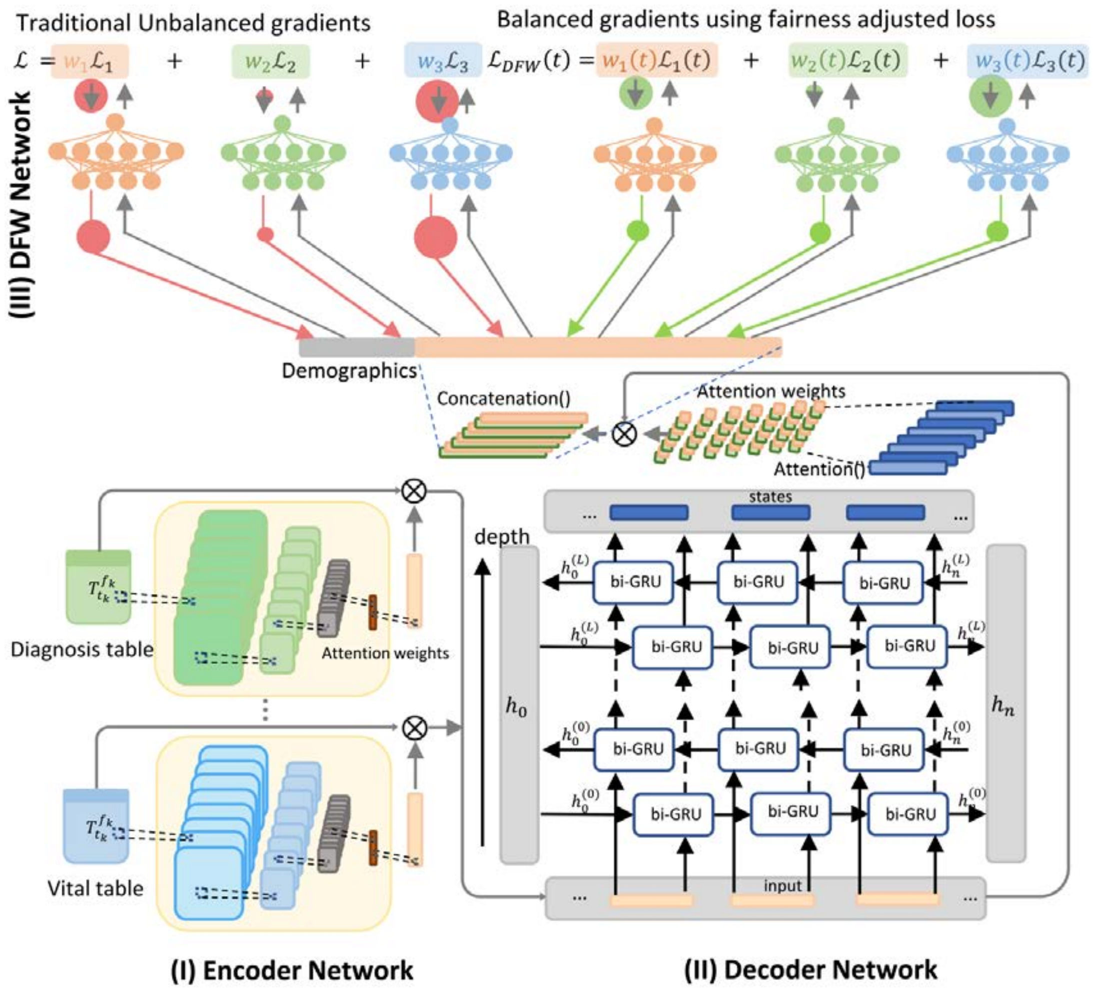
Li C, Jiang X, Zhang K. A transformer-based deep learning approach for fairly predicting post-liver transplant risk factors. J Biomed inform 2023 Nov 20;149:104545. PMID:37992791

Li C, Lai D, Jiang X, Zhang K. FERI: A Multitask-based Fairness Achieving Algorithm with Applications to Fair Organ Transplantation. AMIA Informatics Summit, Boston, MA, 2024. (Best Student Paper award)

Li C, Ding S, Zou N, Hu X, Jiang X, Zhang K. Multi-task learning with dynamic re-weighting to achieve fairness in healthcare predictive modeling. J Biomed Inform 2023 Jul;143:104399. PMID:37211197

Novel Model Architecture

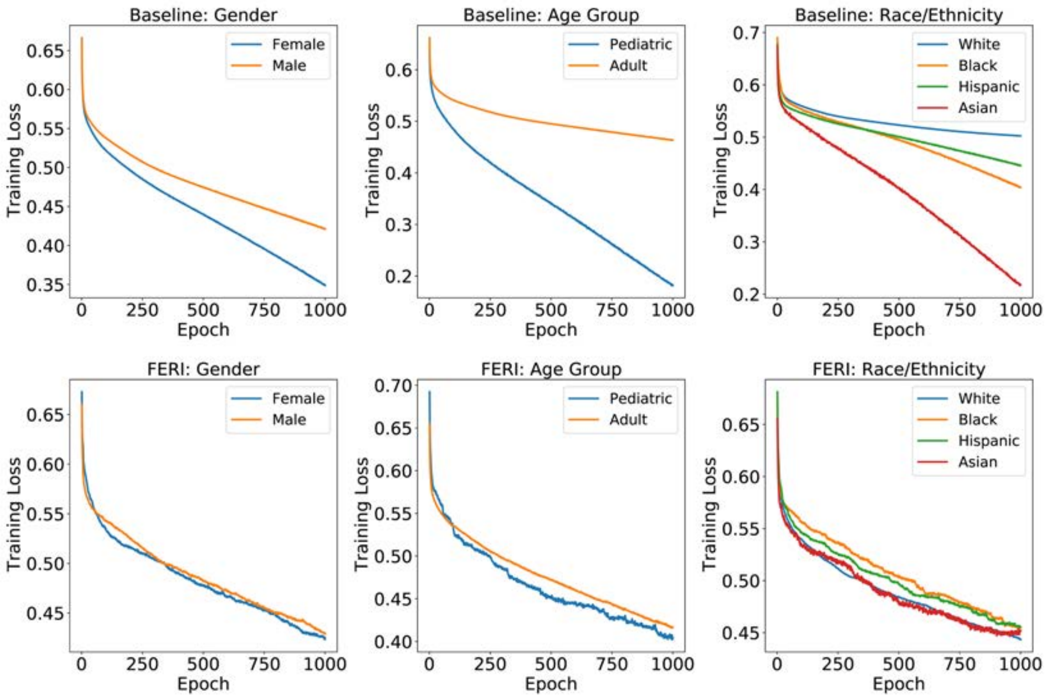
- Designed for fairness in algorithmic predictions.
- Improves traditional unbalanced gradients with a balanced gradients method.
- Incorporates a fairness-adjusted loss function.
- Features an encoder with attention mechanisms for diverse data inputs.
- Utilizes bi-GRUs in the decoder for accurate predictions.
- Ensures equitable model through fairness-adjusted loss function.



Our Innovation

- Introduces Dynamic Fairness re-Weighting (DFW) for demographic fairness.
- DFW adjusts task priority based on subgroup performance.
- Utilizes fairness metrics for gradient re-weighting during back-propagation.
- Dynamically maintains fairness throughout training.
  - Ensure the gradient for different tasks on the same scale during backpropagation. Therefore, the neural network has equal learning speed on each task;
  - Adjust the weight of each task dynamically based on the prediction performance of each task.

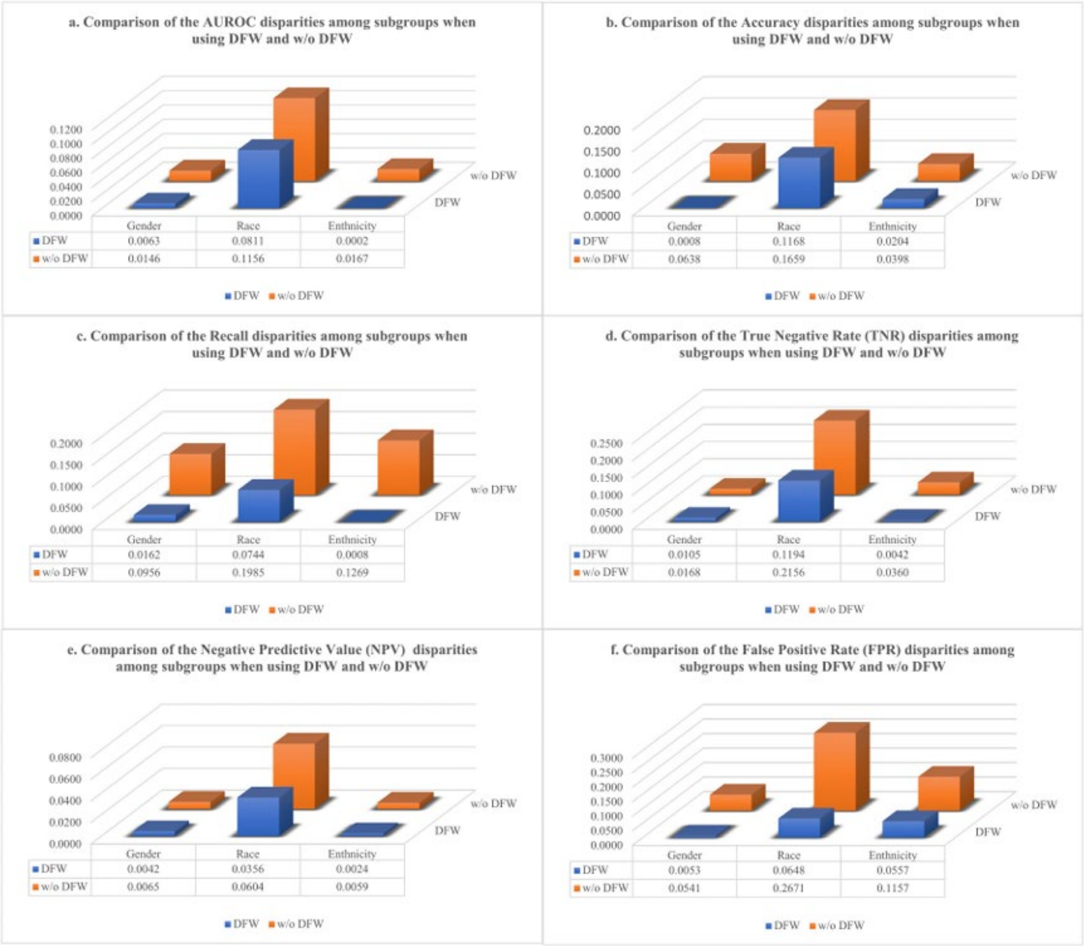
Loss Progression Analysis: Ours vs. Baseline



- In the baseline model, we can see a clear disparity between the male and female.
- In our algorithm's performance, both the male and female subgroups reached a loss value of 0.5 at about the 300-epoch. This consistency emphasizes the model's capability to equitably enhance performance across subgroups

Our Result

- Our method outperforms others in fairness metrics across demographics.
- Achieves better parity in AUROC, accuracy, recall, TNR, NPV, and FPR across groups than traditional approach.



Lessons Learned

- Employing Multitask Learning and adjusting the weights of loss functions dynamically can enhance the fairness of models via in-process learning.
- Conflicting fairness goals, such as Statistical Parity versus Equality of Odds, can arise.
- Due to some fairness metrics being non-differentiable, exploring various mitigation approaches, including continuous relaxation and smooth approximation, is essential.

Best Practices

Continuous fairness metrics monitoring and adjustment.

Promotes interdisciplinary collaboration for bias understanding and mitigation.

Advocates for transparency in model development and inform decision-making.

Future Plans

Enhance	Enhance our tool’s capabilities, expand its functionality.
Collaborate	Collaborate with healthcare professionals to address real world challenges.
Develop	Develop better representation learning to deal with large imbalanceness in subgroups.
Improve	Improve hyperparameter optimization in fine-tuning for optimal performance.

Reference

- *1. Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, et al. Deep Learning in Medical Imaging. Neurospine. 2019 Dec;16(4):657–68.*
- *2. Pan D, Huang Y, Zeng A, Jia L, Song X. Early Diagnosis of Alzheimer's Disease Based on Deep Learning and GWAS. In: Human Brain and Artificial Intelligence. Springer Singapore; 2019. p. 52–68.*
- *3. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A Survey on Bias and Fairness in Machine Learning. ACM Comput Surv. 2021 Jul 13;54(6):1–35.*
- *4. Du M, Yang F, Zou N, Hu X. Fairness in Deep Learning: A Computational Perspective. IEEE Intell Syst. 2021 Jul;36(4):25–34.*
- *5. Pessach D, Shmueli E. A Review on Fairness in Machine Learning. ACM Comput Surv. 2022 Feb 3;55(3):1–44.*
- *6. Nabi R, Malinsky D, Shpitser I. Learning Optimal Fair Policies. Proc Mach Learn Res. 2019 Jun;97:4674–82.*
- *7. Chai J, Wang X. Fairness with Adaptive Weights. Chaudhuri K, Jegelka S, Song L, Szepesvari C, Niu G, Sabato S, editors. 17--23 Jul 2022;162:2853–66.*