Evaluating Fairness in ICU Mortality Prediction with MIMIC-IV

Yun Jing Chang, Lesley

(Dated: February 5, 2025)

I. INTRODUCTION

Biases in Artificial Intelligence (AI) systems can lead to disparities in healthcare outcomes, disproportionately affecting vulnerable patient groups based on sensitive attributes such as race, gender, age, and socioeconomic status. This project is motivated by the urgent need to ensure AI advancements do not exacerbate existing health disparities. In intensive care units (ICUs), where clinical decisions can be life-saving, biased algorithms risk unequal treatment recommendations, delayed interventions, and poorer outcomes for specific demographic groups. This challenge is compounded by the fact that ICUs and emergency rooms (ERs) often face overwhelming patient loads and staffing shortages, particularly in resource-limited hospitals. In such environments, AI can play a crucial role in supporting clinical decisions, optimizing resource allocation, and reducing the cognitive burden on healthcare professionals.

AI integration into healthcare has transformed clinical decision-making, improving diagnostics, patient monitoring, and outcome predictions. AI models excel at analyzing complex medical data, identifying patterns, and generating prognostic insights. AI-driven systems provide timely risk assessments, flag high-risk patients, and assist in prioritizing critical interventions, enhancing patient care even in resource-constrained settings. However, concerns about fairness and bias have emerged, particularly where healthcare professionals heavily rely on AI systems due to staffing shortages. In such settings, biased AI can perpetuate inequities, leading to adverse outcomes for vulnerable populations. This research aims to contribute to developing decision-support tools that promote equitable healthcare delivery.

II. DATASET DESCRIPTION

This study utilizes the Medical Information Mart for Intensive Care (MIMIC-IV), a publicly available database containing de-identified health-related data from over 50,000 ICU admissions at the Beth Israel Deaconess Medical Center between 2008 and 2019. MIMIC-IV includes detailed information on demographics, vital signs, laboratory measurements, treatment interventions, and clinical outcomes. Its extensive coverage and diversity allow for developing predictive models and examining their performance across demographic groups, enabling analysis of potential biases and fairness. The dataset's granularity and diversity provide an opportunity to explore how AI models perform across demographic groups and assess disparities in predictive outcomes.

RESEARCH QUESTIONS

This project aims to address the following primary research question: How can AI models be developed and evaluated to ensure fairness in ICU mortality prediction, particularly concerning sensitive attributes within the MIMIC-IV dataset?

Supporting sub-questions include: 1. What sensitive attributes in the MIMIC-IV dataset contribute to biased outcomes in ICU mortality prediction? 2. How do Logistic Regression, XGBoost, and Multilayer Perceptron (MLP) models differ in predictive performance and fairness across demographic groups? 3. Which fairness metrics are most effective in evaluating biases in ICU mortality prediction models? 4. What bias mitigation strategies can improve model fairness without compromising predictive accuracy?

OBJECTIVES

This project aims to develop predictive models using Logistic Regression, XGBoost, and Multilayer Perceptron (MLP) to forecast ICU mortality. It focuses on identifying potential biases within these models, specifically examining sensitive attributes such as race, gender, age, insurance type, ethnicity, language preference, marital status, and socioeconomic status. To evaluate model fairness, the project will utilize fairness metrics, including demographic parity, equal opportunity, equalized odds, disparate impact, and calibration across subgroups. Additionally, performance metrics like AUC(min), AUC(macro-avg), and AUC(minority) will assess disparities, with equalized odds difference measuring consistency in true and false positive rates.

The project also proposes bias mitigation strategies to improve model fairness while maintaining predictive accuracy. These strategies include reweighting (adjusting training sample weights), threshold modification (adjusting decision thresholds), and the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalances. Ultimately, this research seeks to advance the development of equitable AI systems in healthcare, ensuring predictive models support unbiased decision-making, optimize resource allocation, and reduce the cognitive burden on healthcare professionals.

DATA AVAILABILITY

The data used in this study is available from the MIMIC-IV database (version 3.1), which can be accessed at https://physionet.org/content/mimiciv/3.1/. Access to the dataset requires credentialed approval through the PhysioNet platform after completing the required data use agreements and training on human subjects research.

---

[1] Nikos Afxentis. Predicting mortality and algorithmic fairness of icu patients. Master's thesis, 2024.

[2] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, and H. E. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation [Online]*, 101(23):e215–e220, 2000.

[3] A. E. W. Johnson, L. Bulgarelli, L. Shen, and et al. Mimic-iv, a freely accessible electronic health record dataset. *Scientific Data*, 10(1):1, 2023.

[4] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Brian Gow, Benjamin Moody, Steven Horng, Leo Anthony Celi, and Roger Mark. Mimic-iv (version 3.1). PhysioNet, 2024.

[5] Chuizheng Meng, Loc Trinh, Nan Xu, James Enouen, and Yan Liu. Interpretability and fairness evaluation of deep learning models on mimic-iv dataset. *Scientific Reports*, 12(1):7166, 2022.

[6] Anand Murugan. Implementing fairness in real-world healthcare machine learning through datasheet for database. Master's thesis, University of Waterloo, 2024.

[5] [6] [1] [4] [3] [2]