Fair Dynamic Survival Prediction on Longitudinal Electronic Health Record

Xin Huang, Xiangyang Meng
Department of Information Systems, University of Maryland, Baltimore County

Introduction

- The longitudinal record in an EHR enables doctors to trend labs over multiple encounters for a more holistic and longitudinal overview of their patient's health. **Joint analysis of longitudinal data and survival outcomes** is necessary to obtain unbiased inference when the two outcomes are correlated.
- Cox proportional hazards model, a statistical model, can prioritize the process of providing care to the ones in need of treatments. However, the cox model doesn't consider **the time varying covariates** when fitting for individual and practical treatment questions. Dynamic-DeepHit flexibly incorporates the available longitudinal data comprising various repeated measurements[2], but **fair treatment on different group of patients (gender, race…) remains unexplored**.

Goal

- We aim to develop a **fairness-aware** dynamic survival model for **longitudinal** EHR, to reduce bias, provide a fair prioritization process, and promote fair treatment on certain individuals or demographic groups equally.

Accomplishment

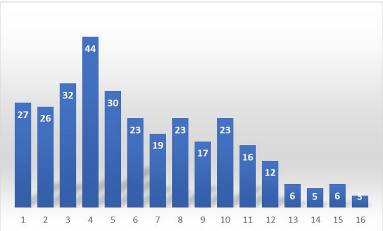
- Implemented **individual and group fairness metrics** to evaluate the dynamic survival analysis with time dependent covariates.
- Incorporated **fairness constraint in loss function** to optimize fairness-aware dynamic survival prediction.

Related Work

[1]Kamrun Naher Keya, Rashidul Islam, Shimei Pan, Ian Stockwell, and James Foulds. Equitable Allocation of Healthcare Resources with Fair Survival Models. SIAM Int. Conference on Data Mining (SDM), 2021.
[2]C. Lee, etc, "Dynamic-DeepHit: A Deep Learning Approach for Dynamic Survival Analysis With Competing Risks Based on Longitudinal Data," IEEE Trans. on Biomedical Engineering. 2020

Dataset

pbcseq: Mayo Clinic Primary Biliary Cirrhosis, sequential data, two events,
312 randomized patients|15 features| 1~16 sequential measurements



Patient measurements frequencies Distribution.

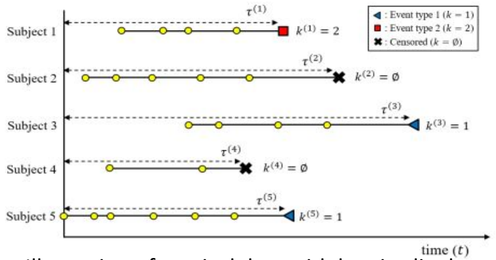
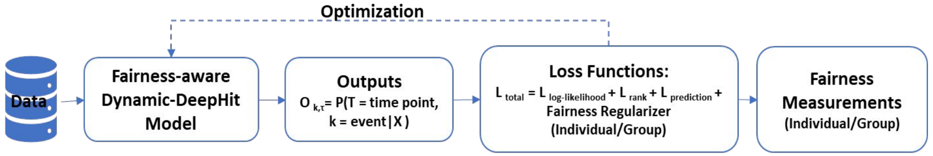


Illustration of survival data with longitudinal measurements where subjects are aligned.[2]

Proposed Methods



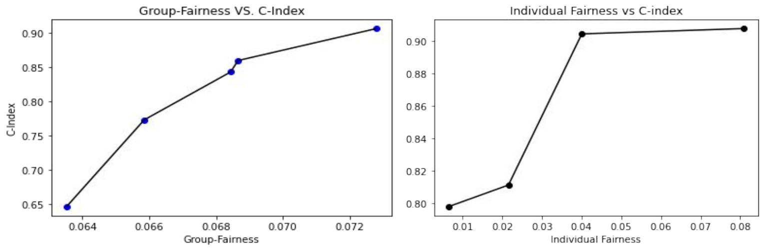
Fairness metric for time varying covariates:
•Individual fairness based on Cumulative Incidence Function: F(t) = P(T<=t)).
•Group fairness based on CIF.

Individual fairness:
$$F_I = \sum_{i=1}^{N} \sum_{j=i+1}^{N} Max(0, |o_{k,\tau}(x_i) - o_{k,\tau}(x_j)| - D_{Euclidean}(x_{i0}, x_{j0}))$$

Group fairness:
$$F_G = Max_{a \in A} |E[o_{k,\tau}(a)] - E[o_{k,\tau}(x)]|$$

Fairness-aware loss in dynamic prediction model:
Total Loss = Log-likelihood Loss + Ranking Loss + Prediction Loss
+Individual/Group Fairness Regularizer

Experiment Results



T = 156 (weeks)	Delta T = 12	Delta T = 36	Delta T = 60
Original			
C-Index	0.948882	0.929326	0.91657
BRIER-SCORE	0.082197	0.104873	0.117934
Group Fairness	0.064016	0.064859	0.06711
Individual Fairness	0.018866	0.0173635	0.048403
With Group Fair Metric			
C-Index	0.939297	0.923575	0.920046
BRIER-SCORE	0.080842	0.102076	0.114453
Group Fairness	0.061551	0.058288	0.066522
With Individual Fair Metric			
C-Index	0.942492	0.923575	0.920046
BRIER-SCORE	0.081845	0.104302	0.116512
Individual Fairness	0.002008	0.0025195	0.019375

Future Work

- Implement ranking-based fairness metric for dynamic survival analysis with competitive risks
- Apply the fairness metrics to more dynamic survival models and conduct more experiments on diverse datasets.
- Improve in leveraging a combination of loss functions to achieve a more efficient accuracy-fairness trade-off.