Survival Analysis of Heart Failure Patients

Souradeep Sen | August, 2023

Supervised By
Professor Krasimira Tsaneva
Dr Ayah Helal

Contents

Research Context

Heart failure is a syndrome that interferes with the heart's pumping ability, causing circulatory problems. Globally, around 64 million people were affected by heart failure as of 2017[1]. Precise risk prediction is vital for better patient outcomes. Traditional semi-parametric models are most popular amongst medical practitioners and clinicians to ascertain risk of mortality for subjects. We compare several traditional methodologies versus newer deep learning methods to gauge their performance.

Research Question

The questions posed in this study are:

1. Are deep learning methods better at predicting risk for HF patients when compared to traditional methods?
2. If so, what are the tradeoffs?

Aim and Objective

Consequently the aim of this study is to analyze a suitable dataset and predict survival of patients using traditional as well as deep learning models. Through rigorous experimentation, the validity of DL models is to be ascertained.

---

[1] S. L. James, D. Abate, K. H. Abate, S. M. Abay, C. Abbafati, N. Abbasi, and H. Abbastabar, Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017., The Lancet https://doi.org/10.1016/S0140-6736(18)32279-7 (2018).

Background

In continuous time survival models, the Cox Proportional Hazards model has been the 'gold standard'. Extensions for nonlinear hazards have been explored, while maintaining the proportional hazards assumption, where covariates exert a multiplicative impact on the hazard.

Deep learning, as well as traditional machine learning, is gaining traction in clinical settings, including survival analysis [2]-[3]. Recent approaches [4] propose using deep neural networks to handle complex electronic health record (EHR) data for risk prediction, showing promise in congestive heart failure cases, while [5] explores real-time risk prediction. Discrete time survival predictions [6] enable probabilistic mortality risk estimation.

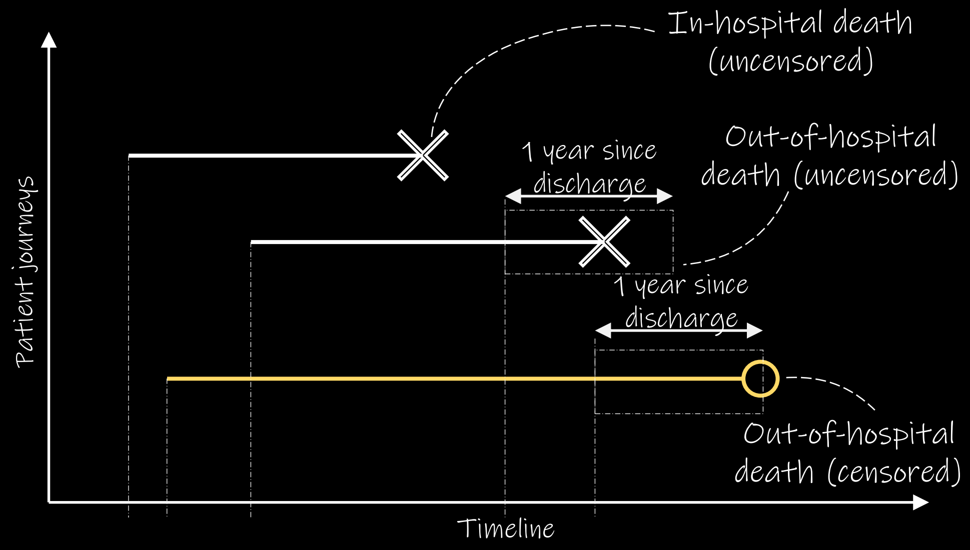
[2] M. Gjoreski, A. Gradisek, B. Budna, M. Gams, and G. Poglajen, Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds, IEEE Access 8, 20313 (2020).
[3] L. Wang, L. Sha, J. R. Lakin, J. Bynum, D. W. Bates, P. Hong, and L. Zhou, Development and Validation of a Deep Learning Algorithm for Mortality Prediction in Selecting Patients With Dementia for Earlier Palliative Care Interventions, JAMA Network Open 2, e196972 (2019).
[4] Z. Che, Y. Cheng, Z. Sun, and Y. Liu, Exploiting convolutional neural network for risk prediction with medical feature embedding, arXiv preprint arXiv:1701.07474 (2017).
[5] L. Brand, A. Patel, I. Singh, and C. Brand, Real time mortality risk prediction: A convolutional neural network approach., in HEALTHINF (2018) pp. 463–470.
[6] C. Lee, J. Yoon, and M. van der Schaar, Dynamicdeephit: A deep learning approach for dynamic survival analysis with competing risks based on longitudinal data, IEEE Transactions on Biomedical Engineering https://doi.org/10.1109/TBME.2019.2909027 (2020).

Data Storage and Processing

The study uses the large publicly available database MIMIC-IV [7], which consists of critical care data from hospital and ICU admissions for almost 300,000 patients admitted to intensive care units at the Beth Israel Deaconess Medical Center (BIDMC). For a more thorough treatment of the data, see MIMIC-IV website.

As per the maintainers' recommendation (see: github), the data was set up on a local postgres server to allow fast querying through Python.

This study explores two approaches: time-invariant and time-varying. Preprocessing is similar for both. Patients with Heart Failure ICD-10 codes are selected, including admission/discharge times, and static as well as time-varying covariates.



[7] A. E. Johnson, L. Bulgarelli, L. Shen, A. Gayles, A. Shammout, S. Horng, T. J. Pollard, S. Hao, B. Moody, B. Gow, and et al., Mimic-iv, a freely accessible electronic health record dataset, Scientific Data 10, 10.1038/s41597-022-01899-x (2023).

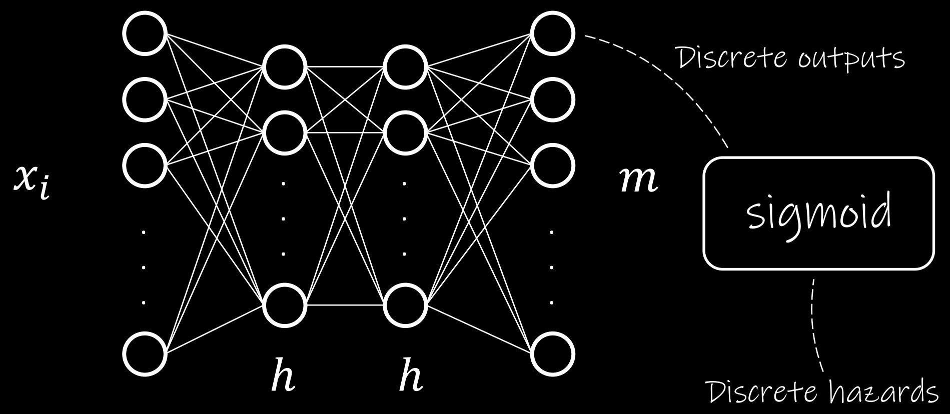
Although the data is freely-available, ethics approval was sought for the project. See: Application ID: 2669275

Date of death (dod) for discharged patients is collected from state records if they died within 1 year of discharge. Otherwise, dod column is left blank (indicating censoring).

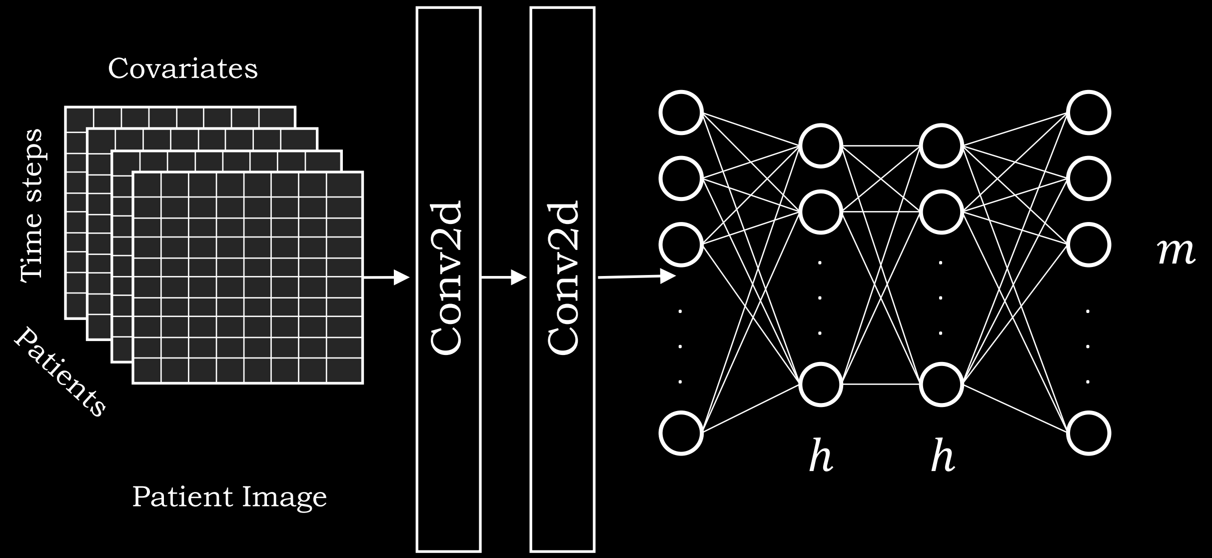
For the time-varying approach, patients need records across datasets and at least 10 time steps, reducing patient count compared to the time-invariant version.

Experimental Design - I

Time-Invariant



Time-Variant



Survival from cumulative hazards [8]

$$S(\tau_j) = \prod_{k=1}^{j} \left(1 - h(\tau_k)\right)$$

Composite Loss Function [8]-[9]

$$\mathcal{L} = \alpha\mathcal{L}_1 + (1-\alpha)\mathcal{L}_2$$

[8] H. Kvamme and Ø. Borgan, Continuous and discrete time survival prediction with neural networks (2019), arXiv:1910.06724

[9] V. C. Raykar, H. Steck, B. Krishnapuram, C. Dehingoberije, and L. Philippe, On ranking in survival analysis: Bounds on the concordance index, Advances in Neural Information Processing Systems (2007).

Experimental Design - II

Traditional Fitters

Cox Proportional Hazards[10]
Weibull Accelerated Failure Time[11]
Random Survival Forest[12]

Deep Learning Fitters

Deep Survival Machines [12]
PyCox with Logistic Hazards*[8]

Proposed Fitters

Time-Invariant Survival
Time-Variant Survival

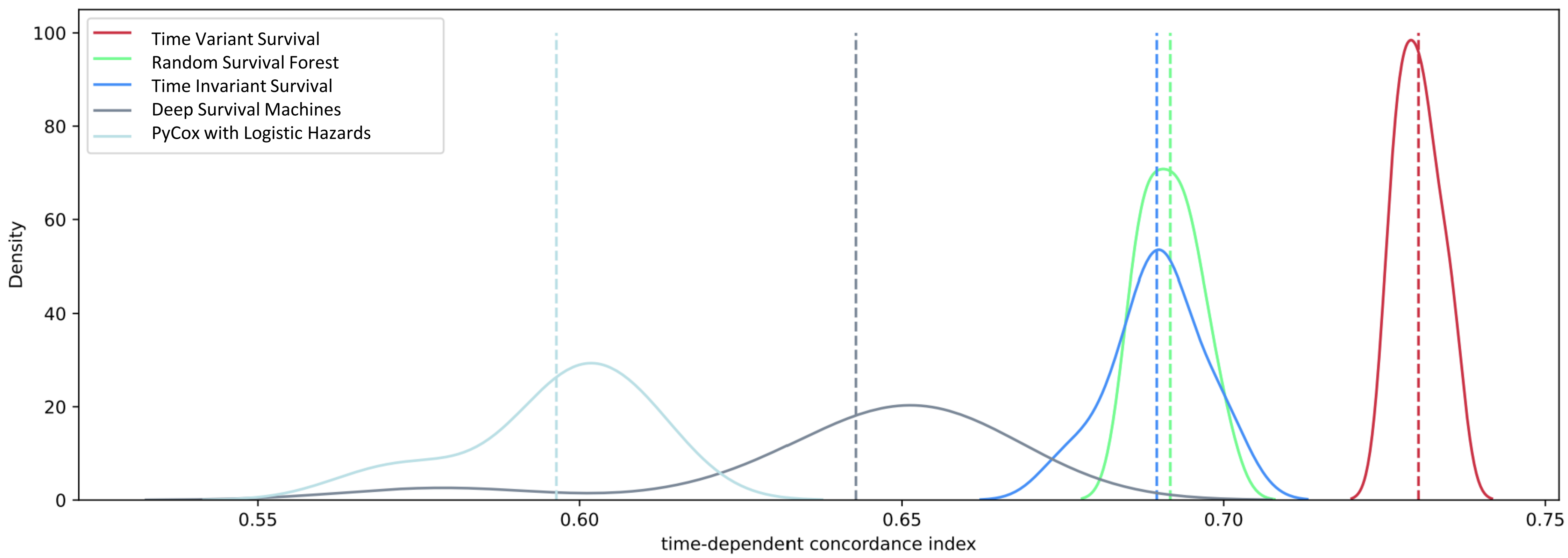
* Applied architecture is torchtuples.practical.MLPVanilla

[10] D. R. Cox, Regression models and life-tables., Journal of the Royal Statistical Society (1972).
[11] W. R. Swindell, Accelerated failure time models provide a useful statistical framework for aging research, Experimental Gerontology https://doi.org/10.1016/j.exger.2008.10.005 (2009).
[12] H. Ishwaran, U. B. Kogalur, E. H. Blackstone, and M. S. Lauer, Random survival forests, The Annals of Applied Statistics https://doi.org/10.1214/08-AOAS169 (2008).
[13] C. Nagpal, X. R. Li, and A. Dubrawski, Deep survival machines: Fully parametric survival regression and representation learning for censored data with competing risks (2021), arXiv:2003.01176 [cs, stat]

Results - I

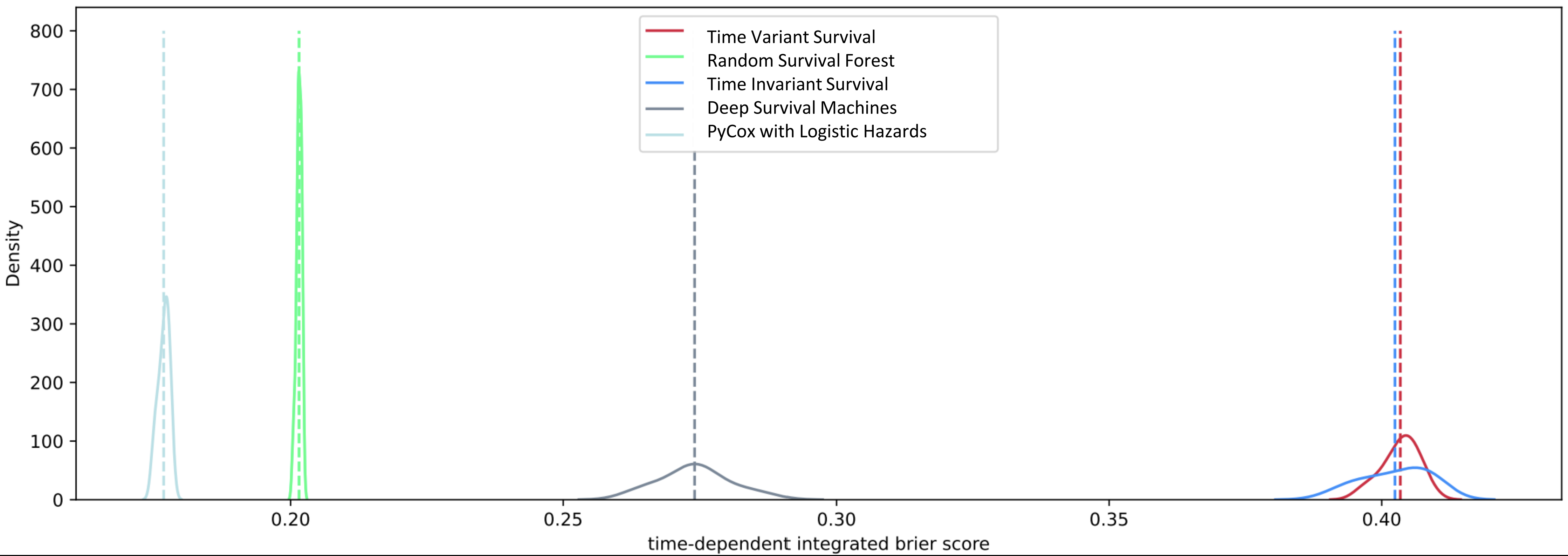
Model	C-index (0.05 , 0.95)	IBS (0.05 , 0.95)
Cox Proportional Hazards	0.6953	***0.1731***
Weibull Accelerated Failure	0.6953	***0.1741***
Random Survival Forest	0.6911 (0.6855, 0.6973)	0.2013 (0.1998, 0.2024)
PyCox Logistic Hazard	0.5984 (0.5723, 0.6104)	0.1768 (0.1751, 0.1778)
Deep Survival Machines	0.6498 (0.6016, 0.6598)	0.2739 (0.2651, 0.2834)
Time-Invariant Survival	0.6903 (0.6789, 0.6994)	0.4030 (0.3937, 0.4090)
Time-Variant Survival	***0.7301 (0.7263 , 0.7352)***	0.4039 (0.3981, 0.4072)

Results - II



Distribution of c-index shows high discriminatory power for Time-Invariant and Time-Variant models

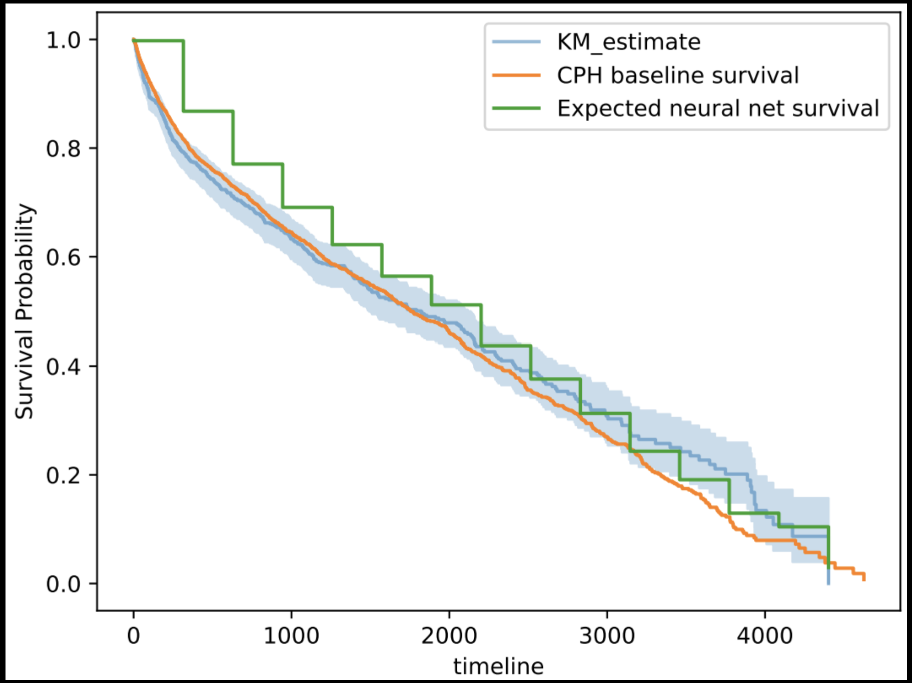
Results - III

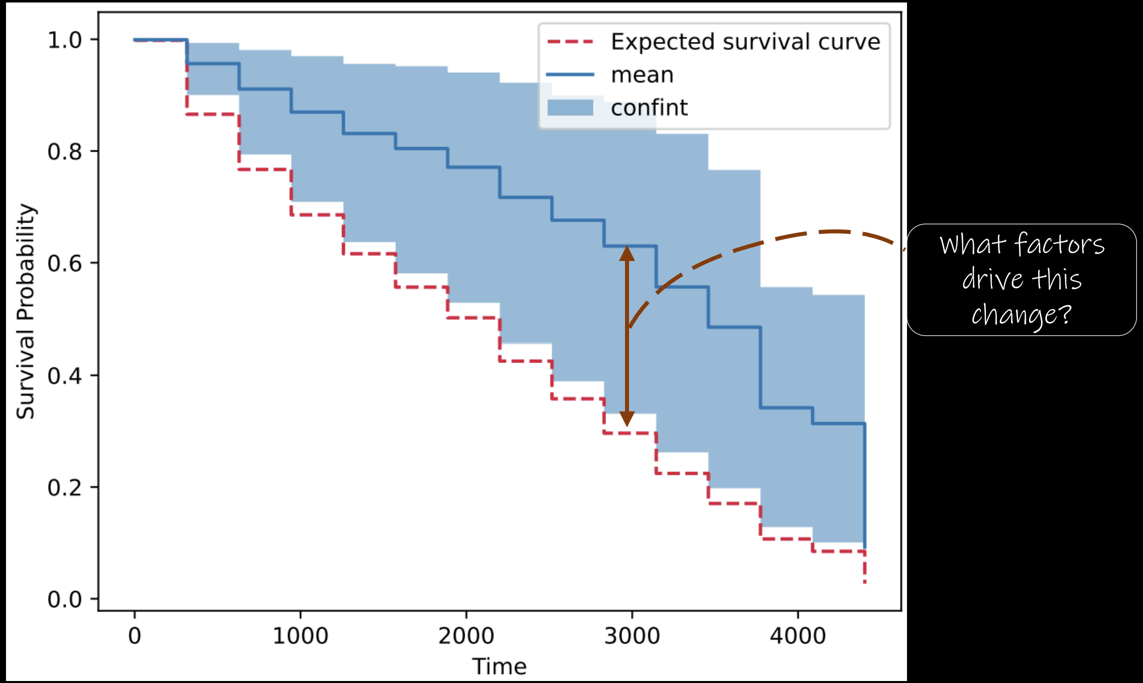


However, the distribution of IBS shows poor calibration power for both models

Results - IV

This implementation is equipped with MC dropout [14] (for generating confidence intervals on individual survival curves) and SHAP [15] (for explaining individual survival curves).





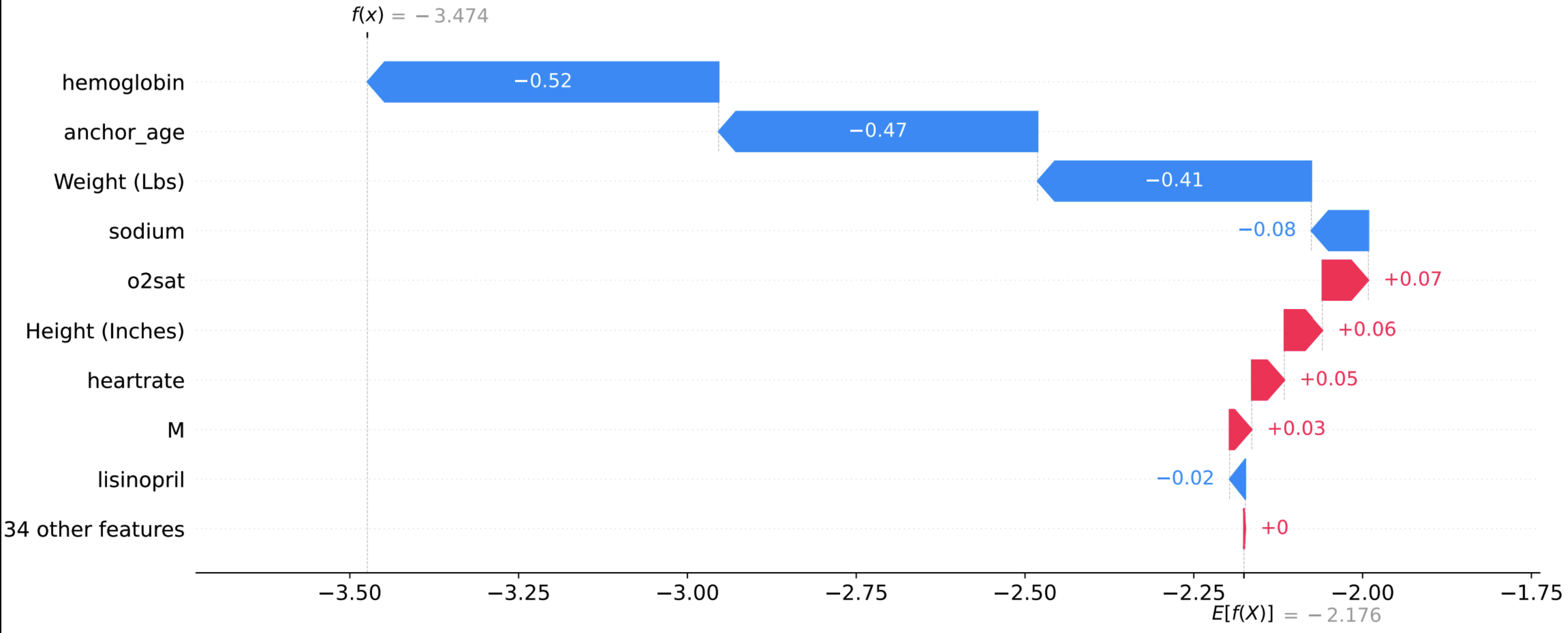
SHAP introduces the notion of an 'expected' survival curve, which may be comparable to an empirical fit or a baseline survival.

How does the survival curve for an individual subject deviate from the 'expected' curve?

[14] Y. Gal and Z. Ghahramani, Dropout as a bayesian approximation: Representing model uncertainty in deep learning, in Proceedings of The 33rd International Conference on Machine Learning, Proceedings of Machine Learning Research, Vol. 48

[15] S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions (2017).

Results - V



SHAP [15] waterfall charts attempt to allocate credit to model covariates which lets users understand which of these had a say in the model's output and how much that effect was. E[f(X)] and f(x) are shown on the predictor scale. Once passed through a sigmoid function, they become valid discrete hazards for the concerned subject over the chosen discrete time period.

Additional Findings

[1] S. L. James, D. Abate, K. H. Abate, S. M. Abay, C. Abbafati, N. Abbasi, and H. Abbastabar, Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017., The Lancet https://doi.org/10.1016/S0140-6736(18)32279-7 (2018).

Further Work

[1] S. L. James, D. Abate, K. H. Abate, S. M. Abay, C. Abbafati, N. Abbasi, and H. Abbastabar, Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017., The Lancet https://doi.org/10.1016/S0140-6736(18)32279-7 (2018).

Thank you for watching!