MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Neural Survival Clustering

Vincent Jeanselme, Brian Tom and Jessica Barrett

Conference on Health, Inference, and Learning
7th - 8th April 2022

Motivation

1. Phenotyping
2. Leveraging Neural Network
3. Interpretability

Literature

1. Survival Analysis
2. Cox PH
3. Deep Surv
4. Deep Cox Mixture

Survival Analysis

Time-to-event modelling with **censored** patients

Population

x Input
t Time of event
d Observed outcome

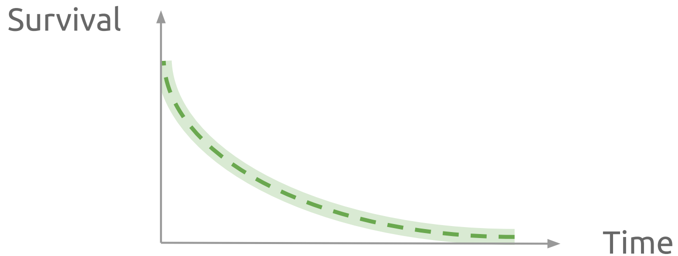
Survival Analysis

Time-to-event modelling with **censored** patients

Population

x Input
t Time of event
d Observed outcome

Aim to model the survival



$$S(t \mid x) = \mathbb{P}(T > t \mid x)$$
$$= e^{-\underline{\Lambda(t \mid x)}}$$
$$= e^{-\int_0^t \underline{\lambda(u \mid x)du}}$$

Cox Model - Proportional Hazard

$$\underbrace{\lambda(s \mid x)}_{\text{Hazard}} \;=\; \overbrace{\lambda_0(s)}^{\text{Baseline Hazard}}\; \underbrace{\exp\left(\beta^T x\right)}_{\text{Covariates Drift}}$$

Regression Models and Life-Tables by D Cox, 1972

DeepSurv

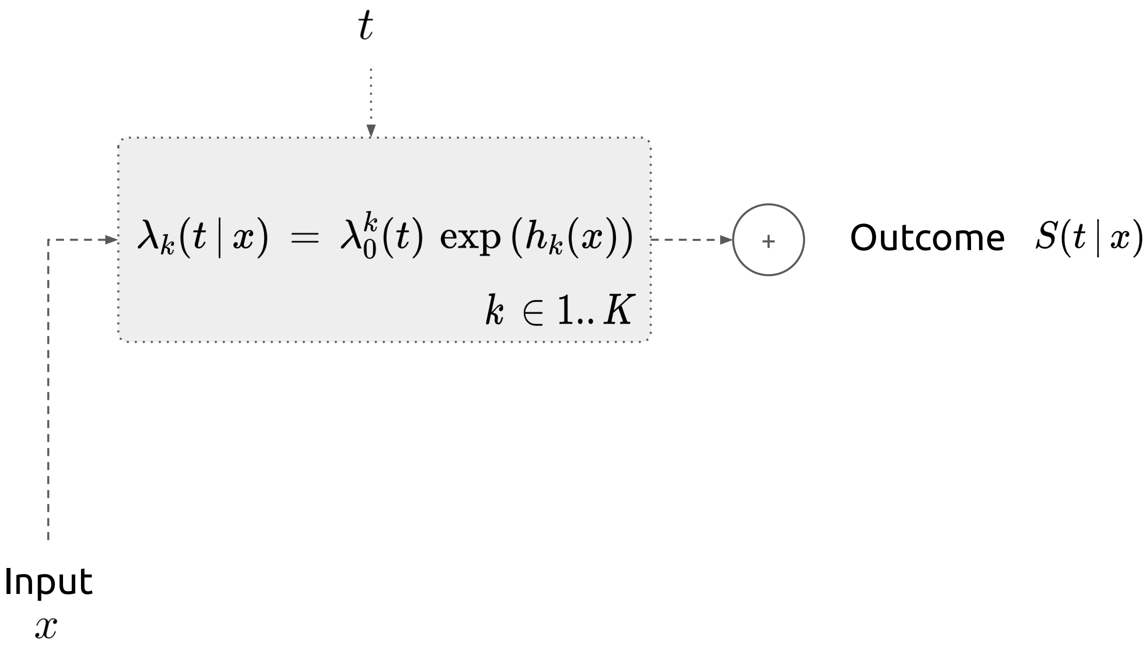
$$\lambda(s \mid x) = \lambda_0(s) \exp\left(\cancel{\beta^T x}\right)$$

$$\lambda(s \mid x) = \lambda_0(s) \exp\left(h(x)\right)$$

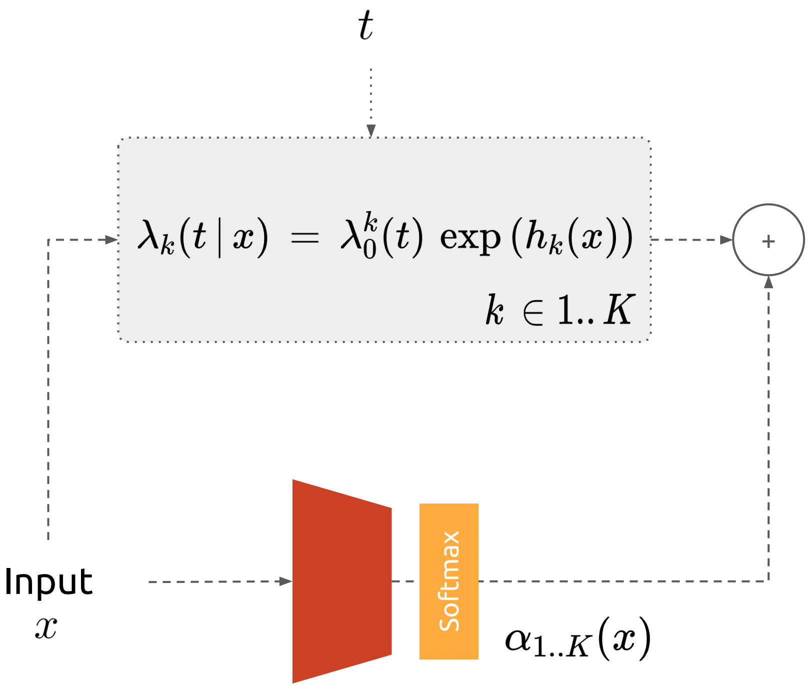
Neural Network
Interaction

DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network by J Katzman & al, 2018

Deep Cox Mixture



Deep Cox Mixture



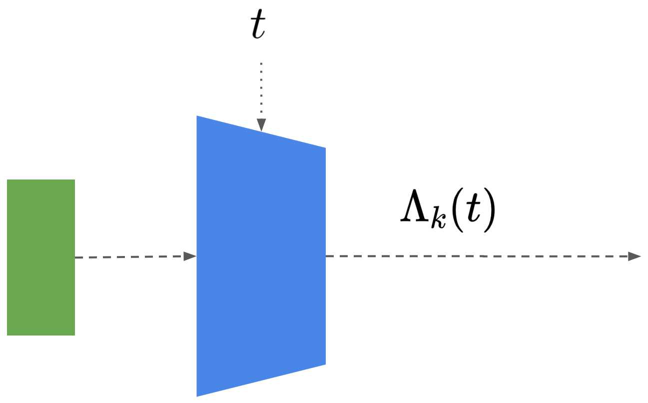
Outcome $$S(t \mid x) = \sum_k \alpha_k \exp(-\Lambda_k(t \mid x))$$
$$= \sum_k \alpha_k \exp\left(-\int_0^t \lambda_k(u \mid x) du\right)$$

Multi Layer Perceptron

Deep Cox Mixtures for Survival Regression by C Nagpal & al, 2021

Proposed Model

Neural Survival Clustering

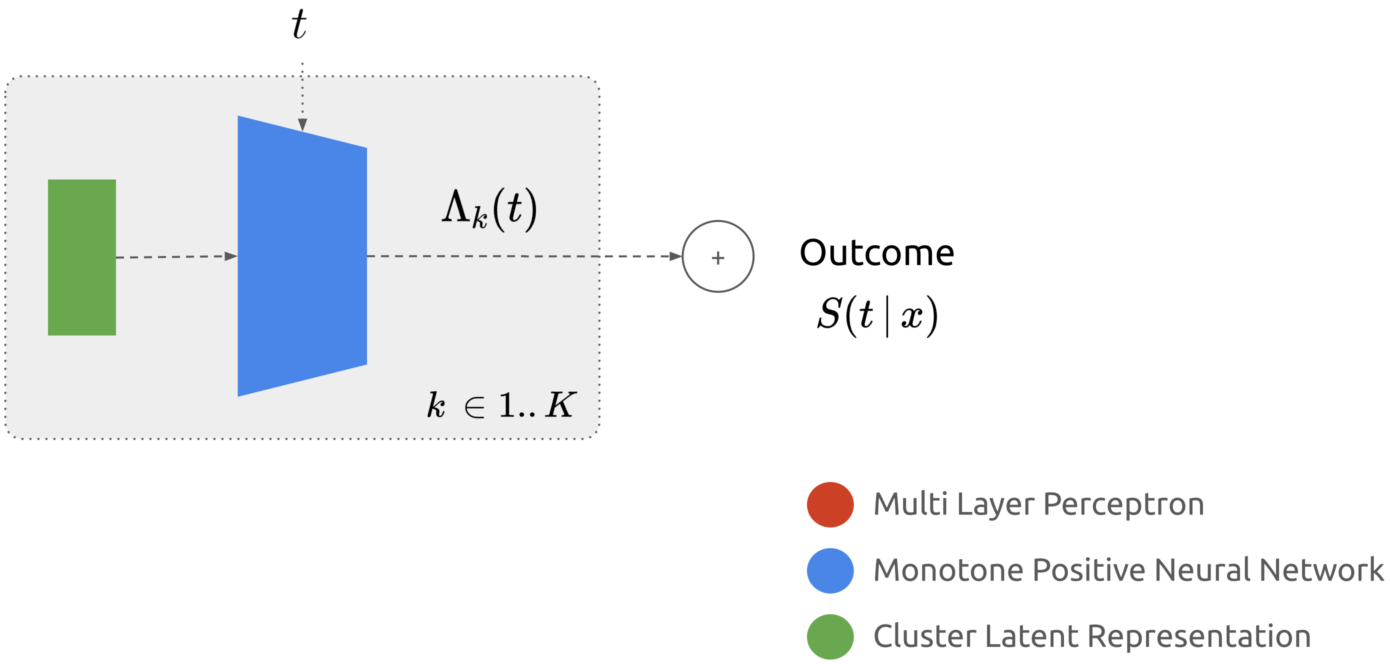


Multi Layer Perceptron

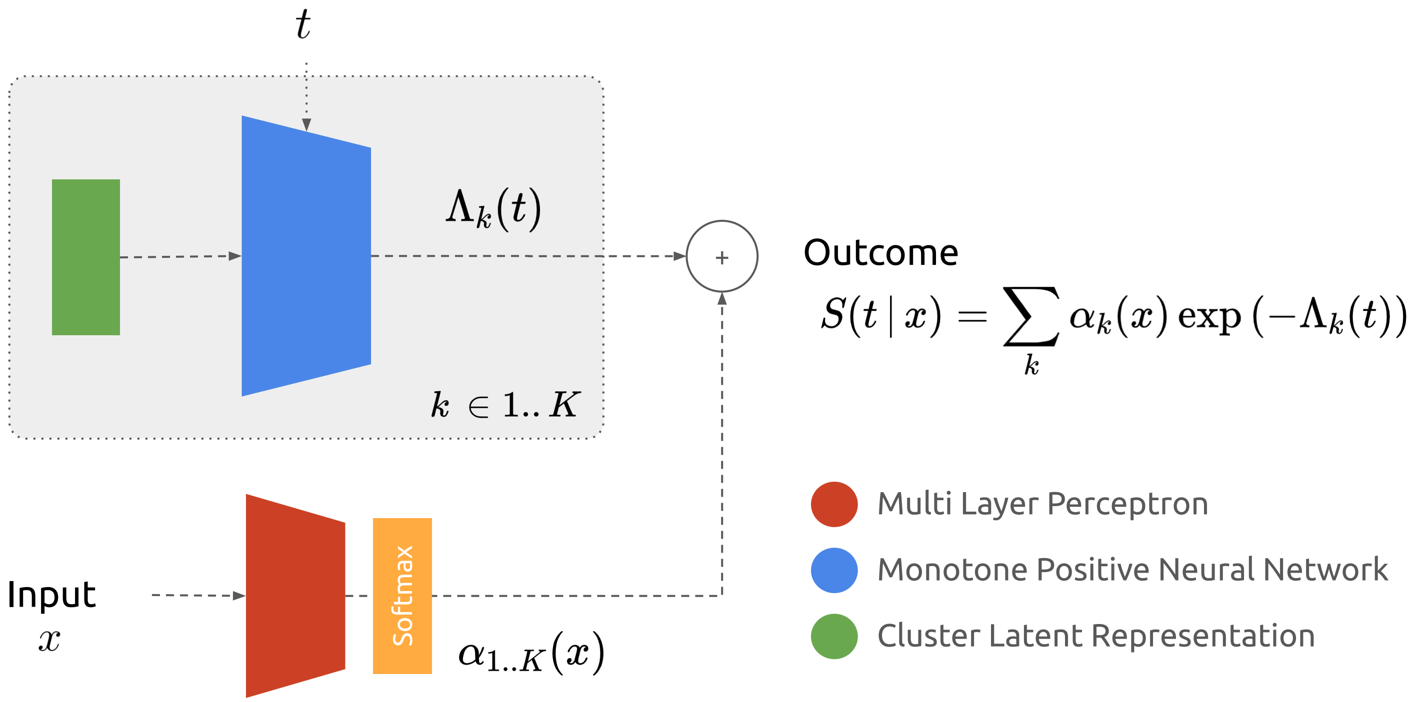
Monotone Positive Neural Network

Cluster Latent Representation

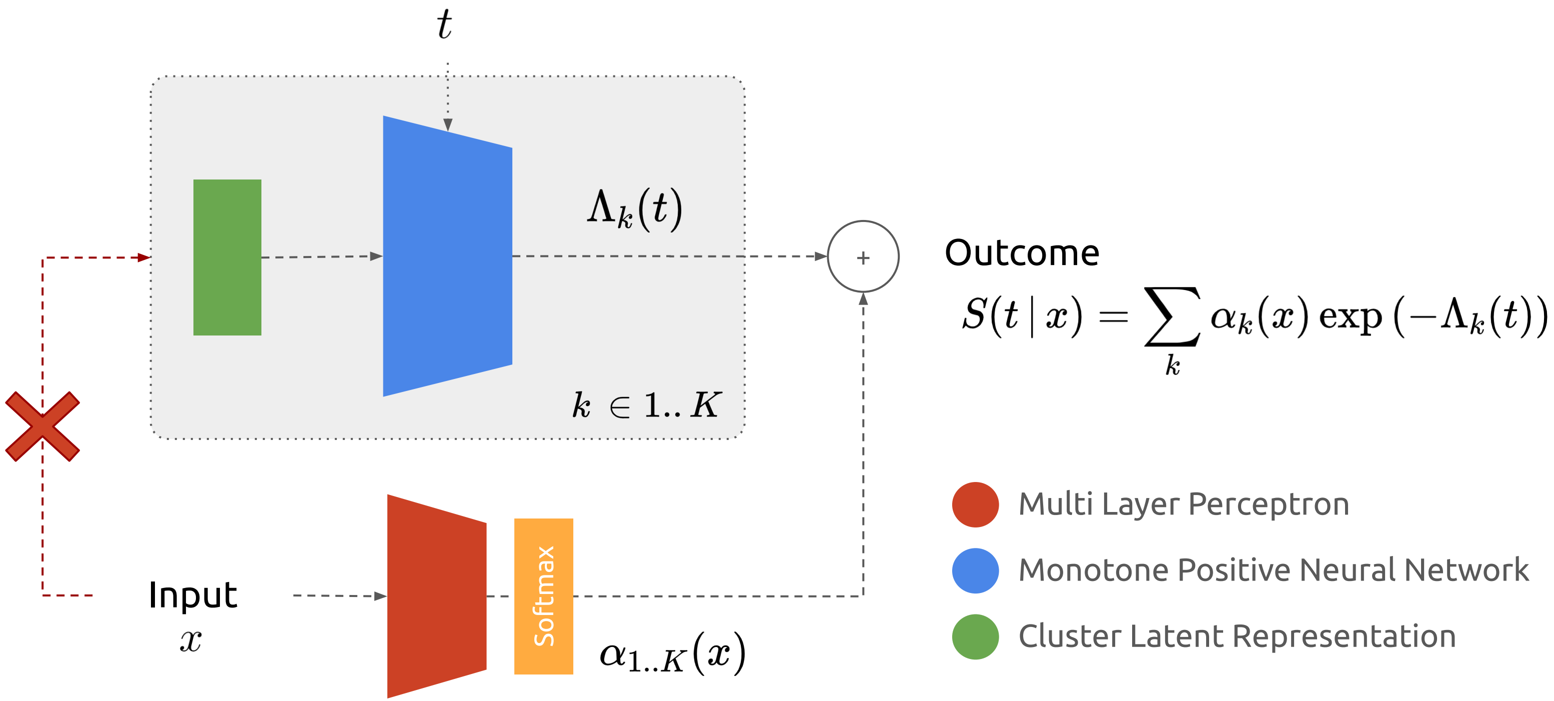
Neural Survival Clustering



Neural Survival Clustering



Neural Survival Clustering



Experimental Results

1. Setting
2. Predictive Performance
3. Clustering

Setting

Dataset	Number of observations	Number of Features	Number of Events	Number of Right Censored
SUPPORT	9,105	30	6,201 (68.1 %)	2,904 (31.9%)
METABRIC	1,904	9	1,103 (57.9 %)	801 (42.1%)
SYNTHETIC	25,000	3	16,385 (65.5 %)	8,615 (34.5 %)

Setting

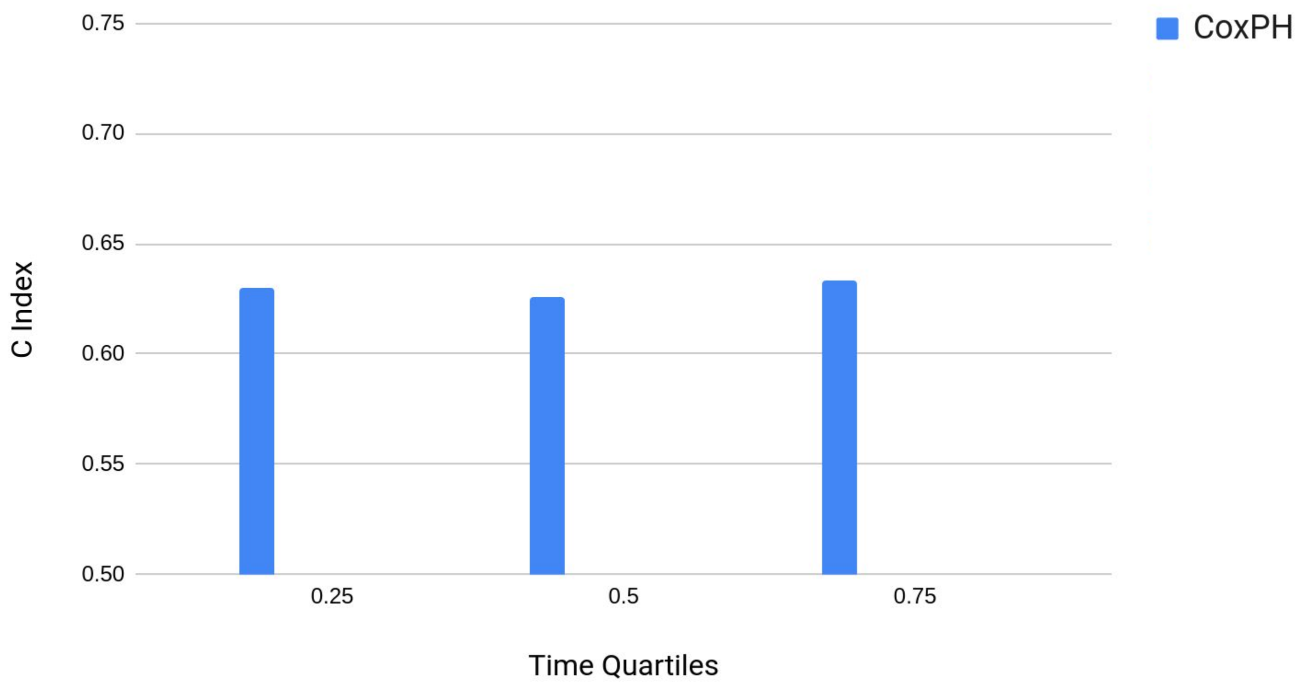
Dataset	Number of observations	Number of Features	Number of Events	Number of Right Censored
SUPPORT	9,105	30	6,201 (68.1 %)	2,904 (31.9%)
METABRIC	1,904	9	1,103 (57.9 %)	801 (42.1%)
SYNTHETIC	25,000	3	16,385 (65.5 %)	8,615 (34.5 %)

Metrics

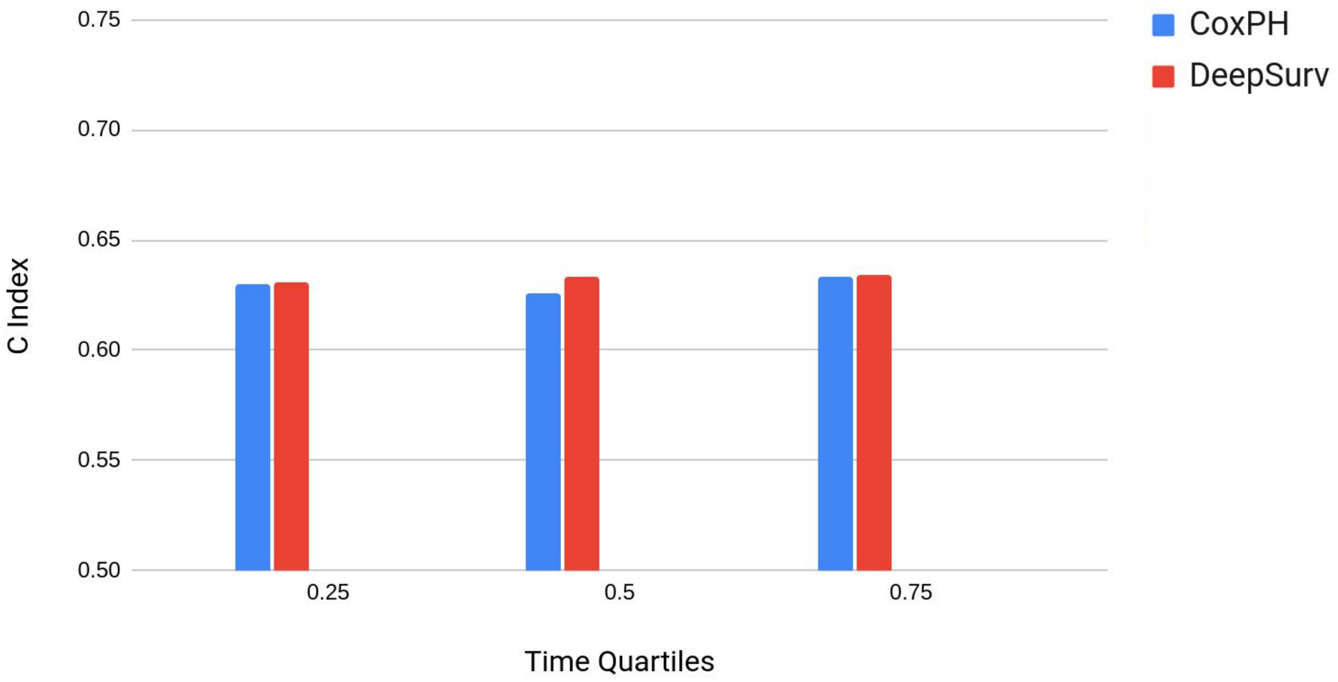
- Time Dependent C Index
- TIme Dependent Brier Score

Experiment: 5 fold cross-validation with inner split for hyperparameter tuning

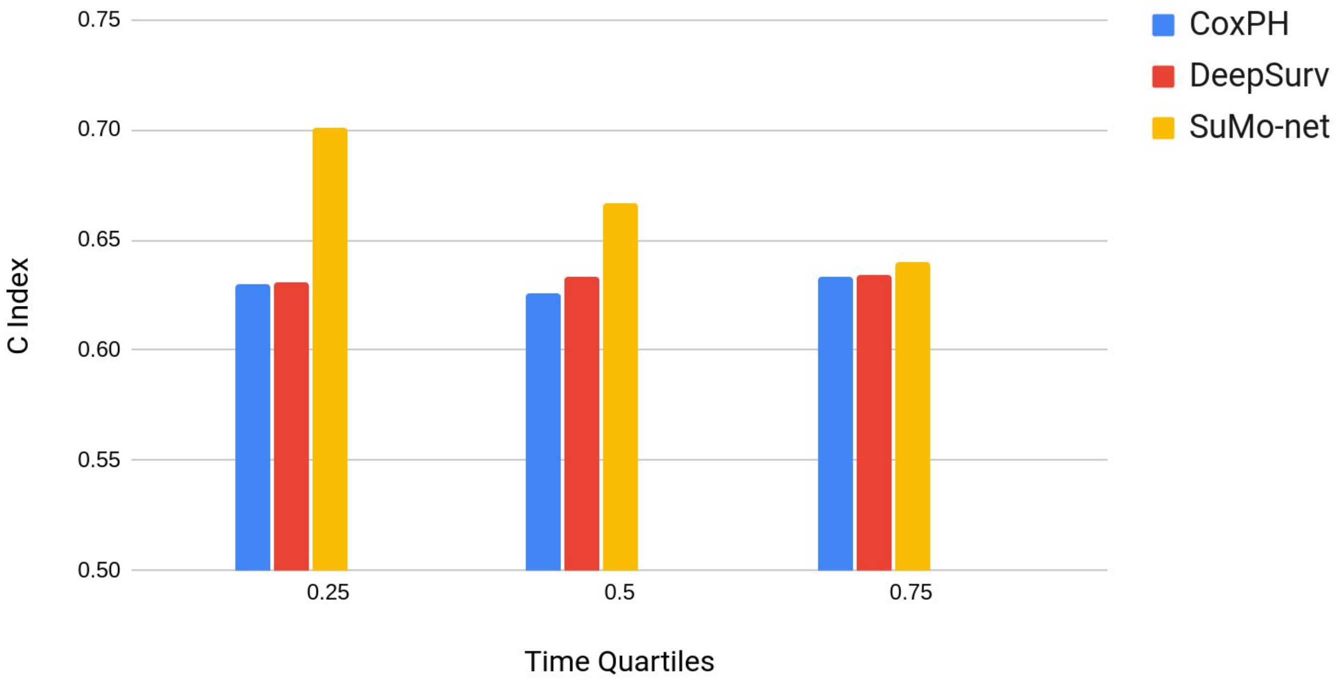
Predictive Performance



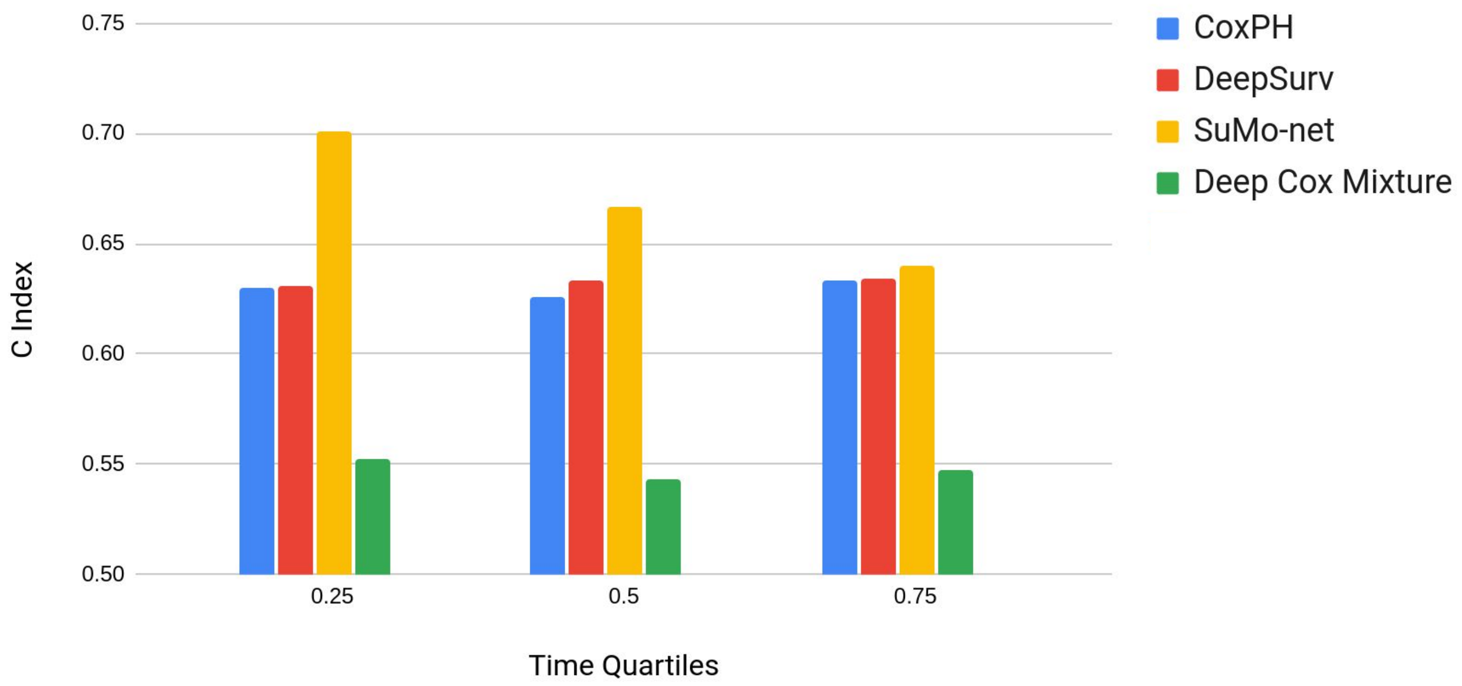
Predictive Performance



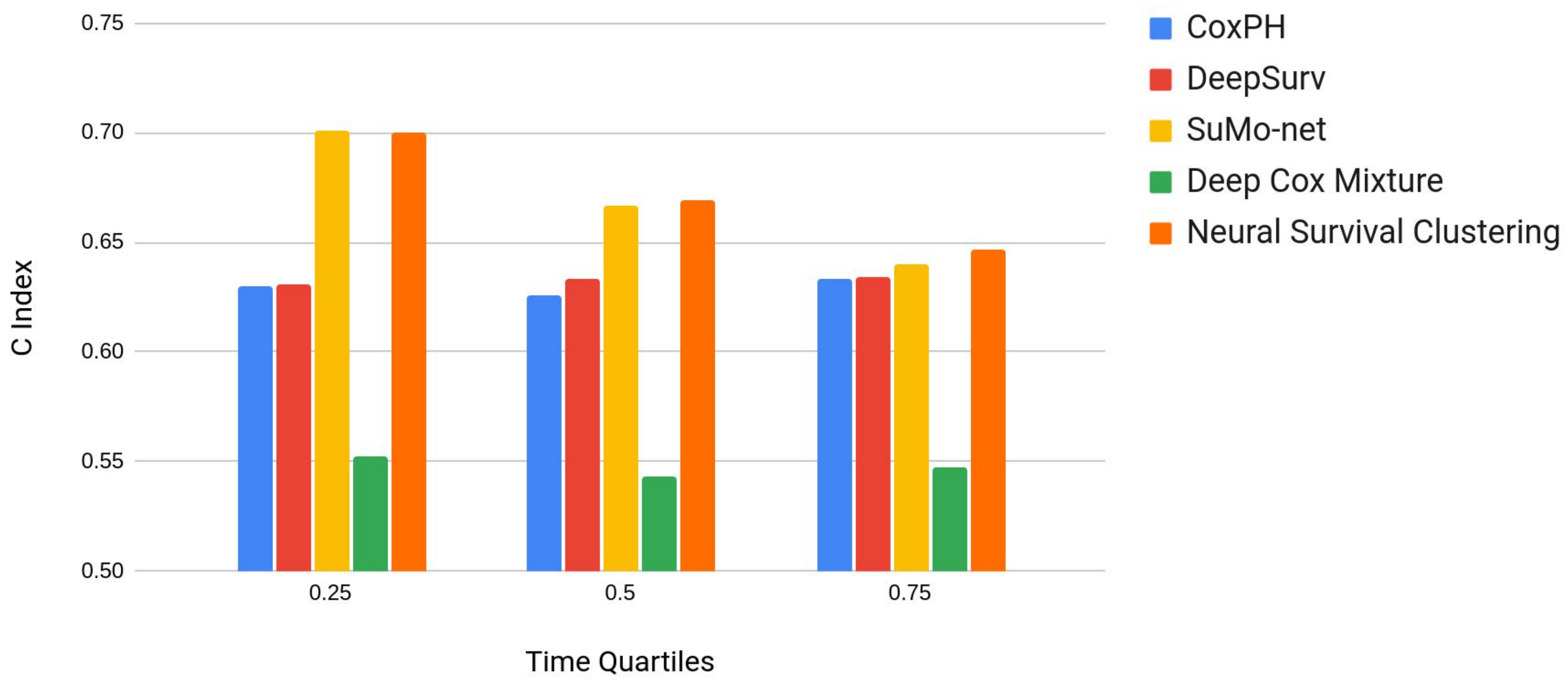
Predictive Performance



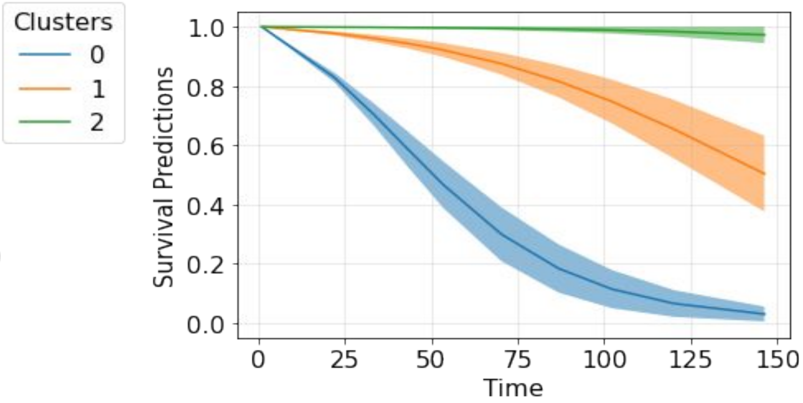
Predictive Performance



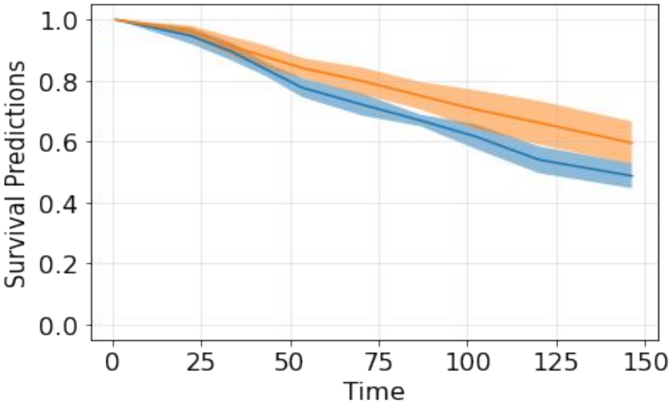
Predictive Performance



Clustering



Neural Survival Clustering



Deep Cox Mixture

Conclusions

- More **interpretable** survival distributions as **not depend** on input covariates

- **No parametric assumptions** on the survival distributions

- End-to-end **optimization** of population clustering

- Competing **predictive performance** with **state-of-the-art** methodologies

Github

Contact
vincent.jeanselme@mrc-bsu.cam.ac.uk
@JeanselmeV