Disease-Atlas: Navigating Disease Trajectories using Deep Learning

Bryan Lim
Department of Engineering Science
University of Oxford, UK
bryan.lim@eng.ox.ac.uk

Mihaela van der Schaar
Department of Engineering Science
University of Oxford, UK
mihaela.vanderschaar@eng.ox.ac.uk

Abstract

Joint models for longitudinal and time-to-event data are commonly used in longitudinal studies to forecast disease trajectories over time. While there are many advantages to joint modeling, the standard forms suffer from limitations that arise from a fixed model specification and computational difficulties when applied to high-dimensional datasets. We propose a deep learning approach to address these limitations, enhancing existing methods with the inherent flexibility and scalability of deep neural networks while retaining the benefits of joint modeling. Using data from the UK Cystic Fibrosis Trust, we show improvements in performance and scalability compared to traditional methods.

1 Overview

In this paper, we introduce a novel conception of the joint modeling framework using deep learning, capturing the relationships between survival probabilities, comorbidity risks, and biomarker trajectories at different stages of disease progression (Figure 1). The network architecture is divided into 3 components as described in Figure 2. Firstly, the shared layer uses a Long-short Term Memory network (LSTM) to incorporate patient history into forecasts, taking in both current longitudinal measurements $\mathbf{X}_t$ and static variables $\mathbf{V}$, and combining them with its internal memory state $\mathbf{m}_{t-1}$ to build a set of shared representations $\mathbf{h}_t$. Next, we group similar prediction tasks into separate task-specific fully connected layers, allowing for a greater degree of specialization. An input horizon τ is also fed into the biomarker and comorbidity risk layers, allowing for predictive distributions to be constructed for arbitrary horizons. Finally, the network produces parameters for predictive distributions in the output layer, using normal distributions for continuous variables, Bernoulli distributions for binary variables, and an exponential distribution for the survival. In addition to point forecasts, the network also produces uncertainty estimates using the Monte Carlo dropout procedure of [Gal and Ghahramani, 2016].

2 Results & Discussion

The performance of the Disease-Atlas was evaluated on data from the UK Cystic Fibrosis (CF) registry - containing measurements of a cohort of 10980 CF patients during annual follow-ups between 2008-2015, with a total of 87 variables associated with each patient across all years. In our investigations below, we consider a joint model for 2 continuous lung function scores (FEV1 and Predicted FEV1), 20 comorbidity and infection risks (treated as binary longitudinal observations) as well as death as the event of interest, simultaneously forecasting them all at each time step.

For survival predictions, we compared the performance of the Disease-Atlas against simpler neural network architectures – LSTMs and Multi-layer Perceptrons (MLPs) – and standard methods from biostatistics – landmarking [van Houwelingen and Putter, 2011] and joint models fitted with a two-step approximation [Wu, 2009]. In general, neural network architectures displayed a clear outperformance over landmarking and joint models, with the Disease-Atlas showing average improvements of 78% versus joint models in terms of the area under the precision-recall curve (AUPRC). Moreover, the Disease-Atlas also showed substantial gains over simpler LSTM and MLP models particularly over short horizons – improving LSTM one-year-ahead predictions by 21% – while also being able to flexibly forecast longitudinal trajectories over multiple horizons.

For longitudinal predictions, we zoom in on comparisons between the Disease-Atlas and joint models which accommodate longitudinal forecasts over various horizons. Once again, the Disease-Atlas demonstrated consistent outperformance - improving MSE forecasts by 56% on average and comorbidity predictions by up to 199% on average across all time steps. This underscores the ability of deep neural networks to learn complex interactions directly from data, and the benefits of a deep learning approach to joint modeling.

References

[Gal and Ghahramani, 2016] Yarin Gal and Zoubin Ghahramani. A theoretically grounded application of dropout in recurrent neural networks. In *Advances in Neural Information Processing Systems*, NIPS 2016, 2016.

[van Houwelingen and Putter, 2011] Hans van Houwelingen and Hein Putter. *Dynamic Prediction in Clinical Survival Analysis*. CRC Press, Inc., Boca Raton, FL, USA, 2011.

[Wu, 2009] Lang Wu. *Mixed Effects Models for Complex Data*. Chapman & Hall/CRC, Hoboken, NJ , USA, 2009.

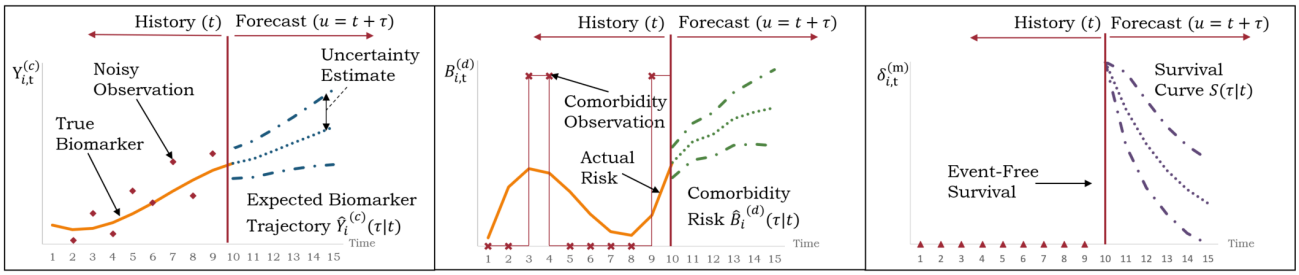

Figure 1: Disease-Atlas Predictions over Time

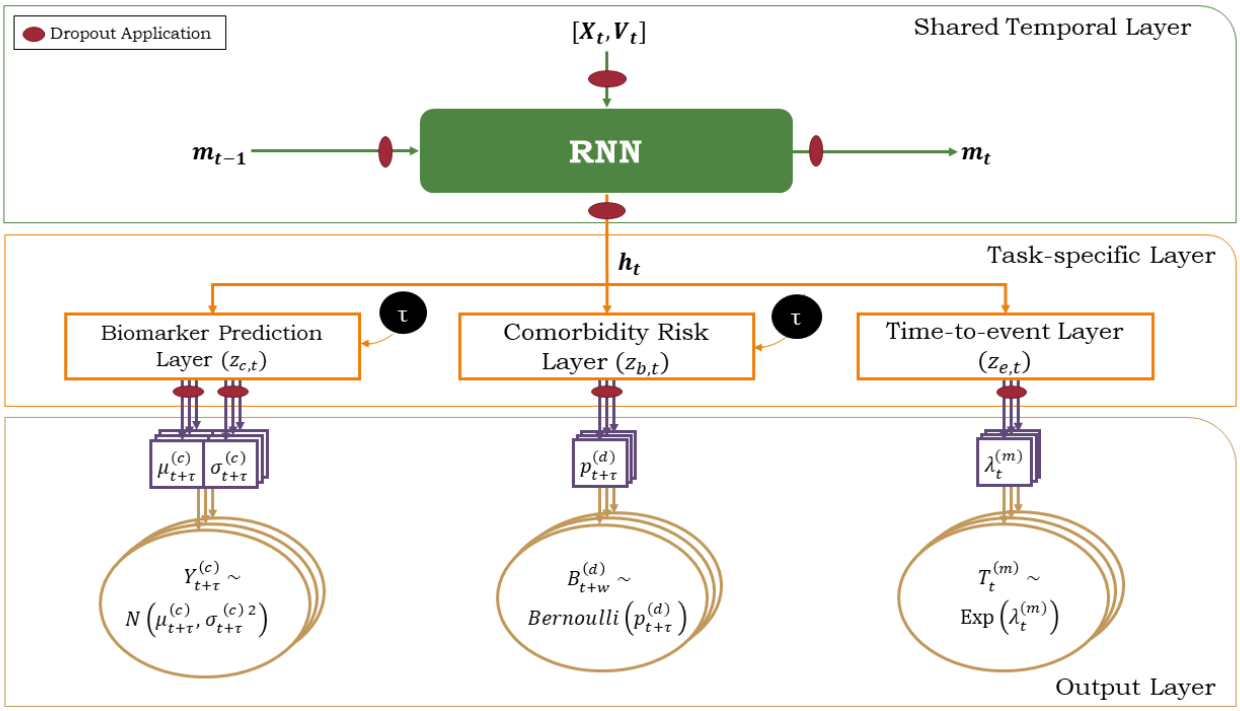

Figure 2: Disease-Atlas Network Architecture