Improving Hospital Readmission Prediction with Longitudinal Medical Histories and Survival Targets

Sacha Davis[1], Russell Greiner[1,2]
[1]Department of Computing Science, [2]Alberta Machine Intelligence Institute

Motivating the Problem

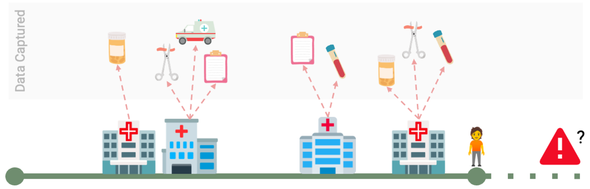
Hospital readmissions are one of the costliest problems facing health systems globally.



Features commonly used to predict readmissions are:

- Highly sparse and of poor quality OR
- Painstakingly hand-engineered

Rich information is contained in the poorly-behaved data captured alongside medical events.



How can we take advantage of longitudinal medical history data for predicting readmissions?

Furthermore, most problems are formulated as a binary readmission task at a particular time-point, and are therefore restricted to classification algorithms.

Binary Target



$$r_{30} = \begin{cases} 1, & \text{if readmitted within 30 days} \\ 0, & \text{otherwise} \end{cases}$$

Survival Target

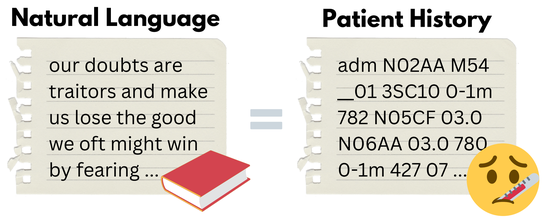


$$t = \text{Time until event}$$

$$\delta = \begin{cases} 1, & \text{if event is a readmission} \\ 0, & \text{event is a censoring} \end{cases}$$

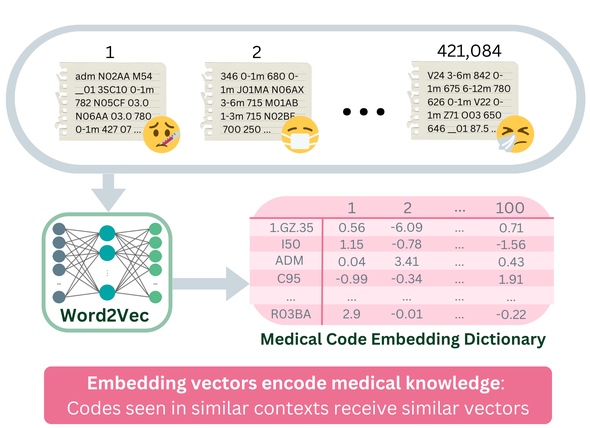
How can we use survival targets to provide more information and improve performance?

Representing Patients



Medical Code Embeddings

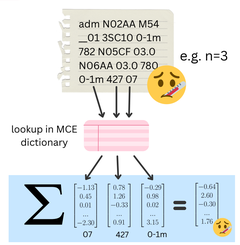
Population-wide medical usage from Alberta Health Services from Jan 01 2012 - Dec 31 2014:



Medical Code Embedding Dictionary

Embedding vectors encode medical knowledge:
Codes seen in similar contexts receive similar vectors

1 **Sum vectors** associated with **25** most recent medical codes



Agg-Seq Representation

2 **Collect** 8 features available directly from data source

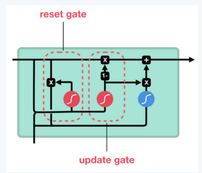
- Discharge disposition
- Length of hospital stay
- Admission acuity
- Comorbidity risk score
- ED use in last 6 months
- LACE score
- Age
- Sex

Bare Representation

Modeling Readmission-Free Survival

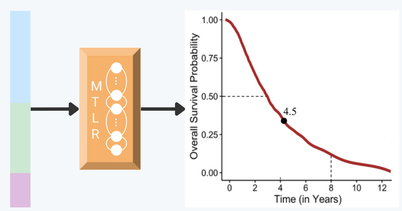
Models

Binary Target **Survival Target**

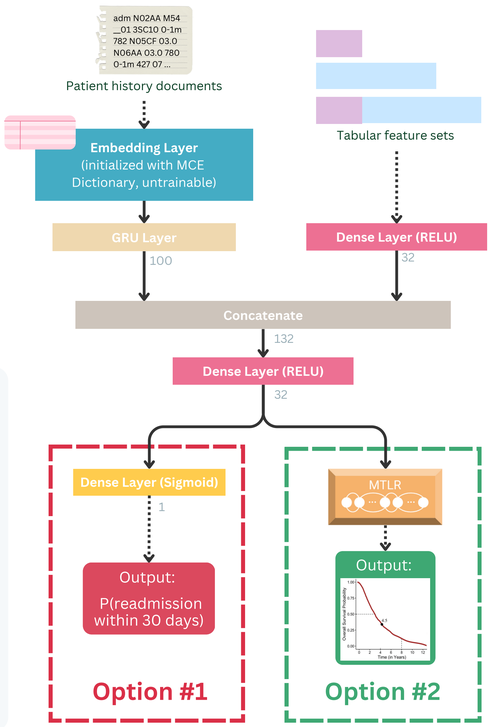


Gated Recurrent Units can ignore unhelpful medical codes when computing patient representations

MTLR models readmission-free survival probabilities using time-to-event information.



MTLR be combined with deep neural networks (N-MTLR) for representation learning.



Results

- **N-MTLR** achieves a **superior AUROC score** using survival targets
- Incorporating features based on **sequential medical histories** can **improve readmission prediction**
- Bare & AggSeq features **outperformed** Bare + 38 hand-engineered features (not pictured)

