Neural networks versus Logistic regression for 30 days all-cause readmission prediction

Ahmed Allam, George Thoma, Michael Krauthammer

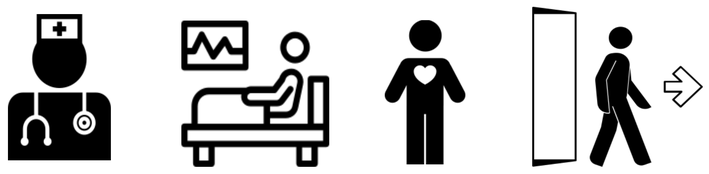
Acknowledgements

Background

Act 1: Admission

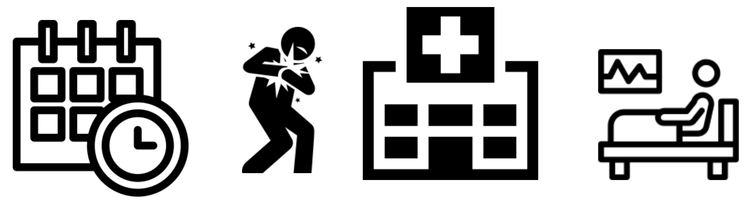
Patient admitted to hospital for Heart Failure (HF) event

Act 2: Discharge



Patient getting treated for HF and later discharged

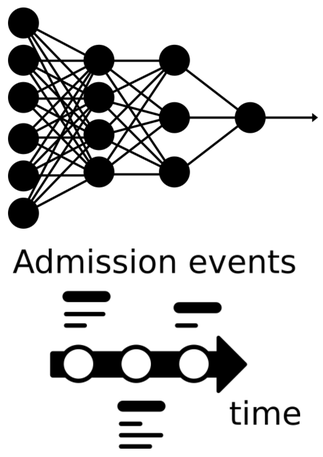
Act 3: Readmission



Within **30 days** from discharge, patient gets readmitted **AGAIN** to hospital

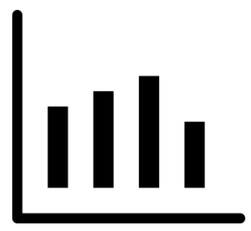
Aims

1. Explore the systematic application of neural network-based models versus logistic regression for predicting 30 days all-cause readmission after discharge from a HF admission
2. Examine the additive value of patients' hospitalization timelines on prediction performance

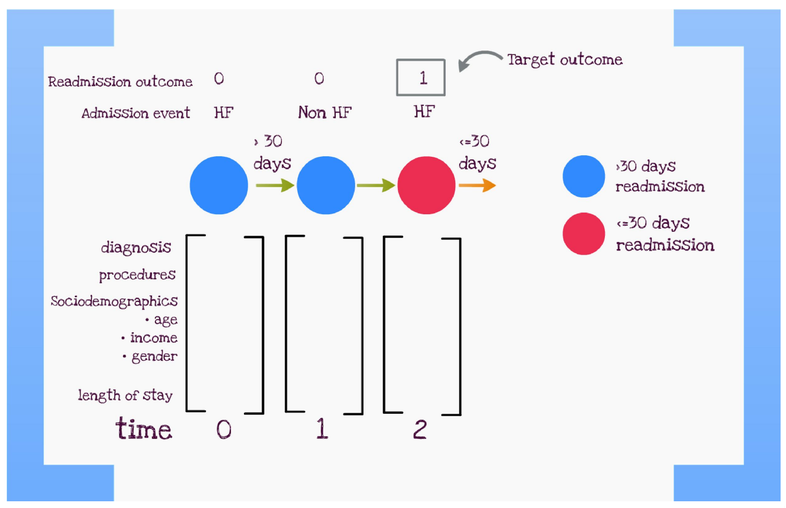


Dataset

- Nationwide readmission database from Healthcare Cost and Utilization Project (HCUP)
- Includes patients discharges from **21 states** accounting for **49.1%** of US hospitalizations over **1 year** period.

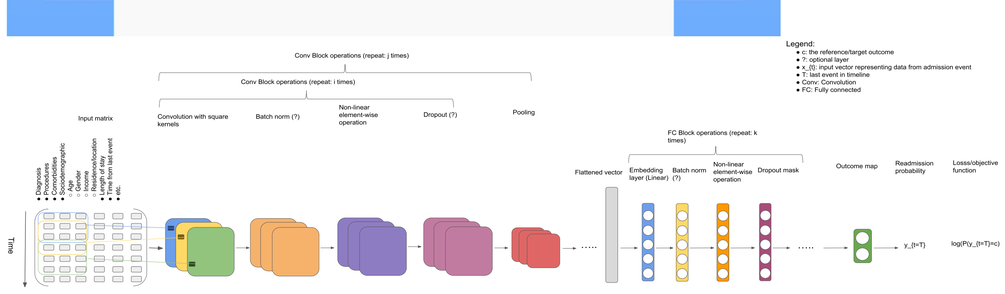


- number of patients: **272,778**
- number of events: **512,842**
- number of HF events: **343,328** (**66.6%**)
- number of readmissions: **81,087** (**23.6%**)
- number of readmissions for last HF: **45,183** (**16.6%**)
- age: **72.89** (**14**) with range [**18-90**]
- gender: **49%** female

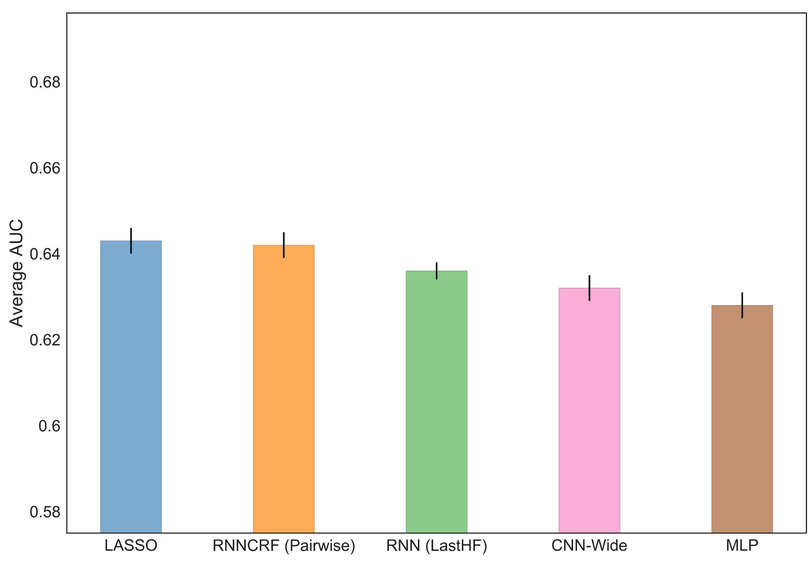


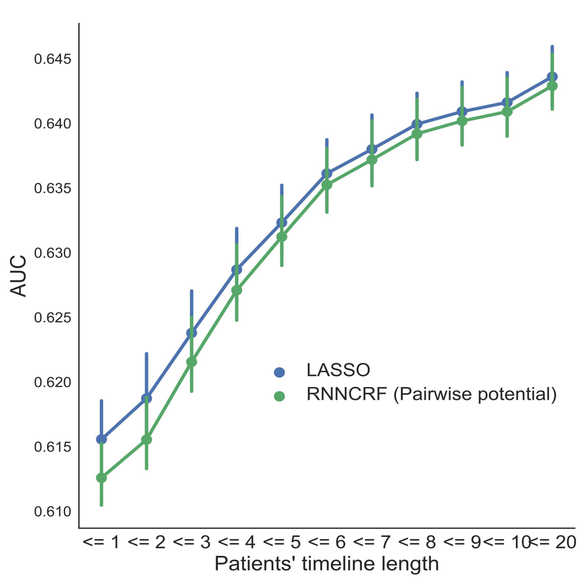
Model Zoo

- Recurrent neural network (RNN)
- RNN with scheduled sampling (RNNSS)
- RNN with conditional random fields (RNNCRF)
- Neural CRF
- CRF only
- Convolutional neural network (CNN)
- CNN wide model [kim,2014]
- Multilayer perceptron (MLP)
- Logistic regression (L1 regularization) (LASSO)



Results





Wrap-up

➔ A combination of RNN with CRF is the best neural model

➔ LASSO model performs equally to RNNCRF

➔ Using patients' timeline improves prediction performance of neural models

Code repository

Presentation