Improving the Prognosis of Intensive Care Unit Patients with Pneumonia

Boyang Tom Jin (tomjin@stanford.edu)

Introduction

PNEUMONIA

- Inflammation of the lung parenchyma typically caused by a bacterial or viral infection (1)
- Major types: hospital-acquired (HAP), ventilator-associated (VAP), and community-acquired (CAP) pneumonia (1, 2)
- HAP occurs up to 20 in 1000 hospital admissions; one-third of admissions to the intensive-care unit (ICU). Disease progression confounded by other underlying health conditions. (2)
- Treatment requires selecting the right choice of drugs in the face of antibiotic-resistant bacteria and unknown infectious agent source (2)

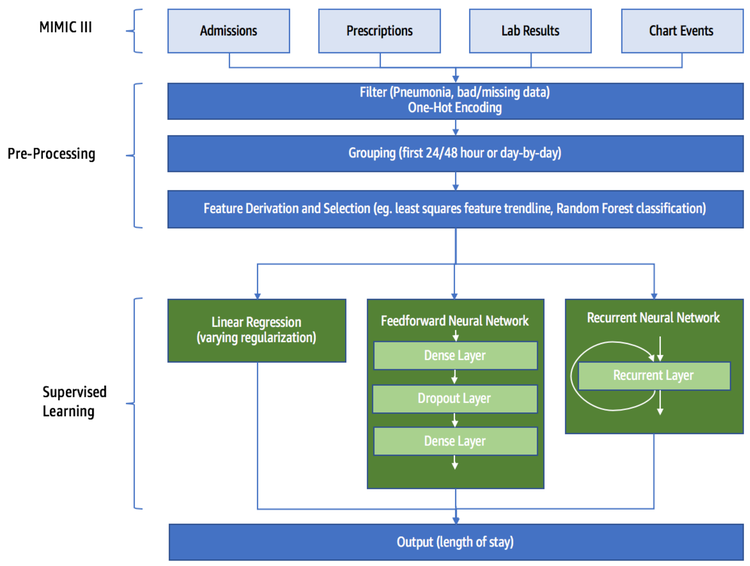
MOTIVATION

- HAP and VAP is a major cause of deaths (up to 27-51% mortality and 2x likelihood of death if in the ICU) (2)
- Burden on hospital resources, increasing average stays in the ICU from 6.4 to 21.6 days and adding on average $40,000 USD to hospitalization costs (2)

MIMIC-III Data

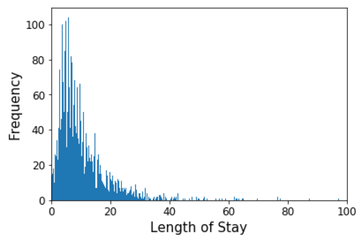
Beth Israel Lahey Health
Beth Israel Deaconess Medical Center

- Freely available de-identified electronic medical records of over 40,000 ICU patients between 2001 and 2012



PROCESSED DATA

- **4508** patients diagnosed with a form of pneumonia (via ICD code)
- Length of stay (LOS): **10.75 days** (mean)
- Number of initial features: **2474**
- Number of features after feature selection (Random Forest): **48**

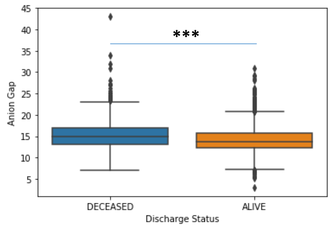


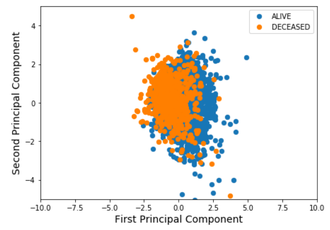
Clinical Features Influencing Outcomes

HOSPITAL EXPIRY

- Top five features selected using a Random Forest classifier with 1000 trees and bootstrapping
- Principal Component Analysis (PCA) on the selected features show greater variation in the population of patients who survive
- High levels of anion gap indicates acidosis, which could be a sign of poor respiratory exchange

Selected Feature
Length of Stay
Systolic Blood Pressure
Diastolic Blood Pressure
Age
Anion Gap





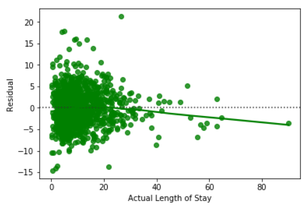
LENGTH OF STAY

Selected Feature	Bonferroni-Corrected P-Value
Systolic Blood Pressure	1.53×10^{-61}
Heart Rate	4.76×10^{-45}
Diastolic Blood Pressure	1.71×10^{-24}
O2	9.45×10^{-23}
IPRA2H Administered	8.61×10^{-15}
Age	2.28×10^{-13}
O2 Flow	2.44×10^{-12}
Positive End-Expiratory Lung pressure	4.35×10^{-12}
Has Medicare	5.43×10^{-12}

- Top 10 features as selected through univariate linear regression using the F-statistic with respect to the LOS
- Selected features are consistent with management and pathophysiology of pneumonia
  (eg. IPRA2H = Ipratropium bromide, a drug use to open congested airways)
- Other features requires further investigation (eg. Medicare patients associated with shorter stays)

Length of Stay Prediction: Linear Regression

Time Window	Regularization	Training Error	Validation Error	Testing Error
24	None	5.500	6.869	6.909
	L1	5.684	6.421	6.512
	L2	5.809	6.466	6.356
48	None	5.409	8.341	7.153
	L1	6.108	6.794	6.334
	L2	5.816	6.225	5.934



- Error is measured in mean absolute error (days)
- Plot shows the residual for the 48 hour window data using L2 regularization
- Testing error of **5.934 days** for 48 hour window is already better than the 15.27 days observed for the baseline.

Future Work

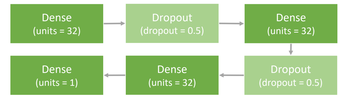
- Evaluate the performance of a convolutional neural network (CNN) on the time-series data in comparison to the LSTM model.
- Compare and contrast validation error on the same neural network models between different demographic data (eg. does gender play a role in predicting LOS?)

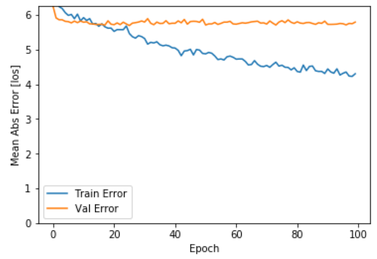
Length of Stay Prediction: Feedforward Neural Network

GOAL

Predict LOS as a function of ICU admission data. This model helps hospital resource planning and is a pseudo-predictor of disease severity.

NETWORK DESIGN





- Network was run on aggregated clinical data collected in the first 48 hour window
- Drop-out layer added to force network to rely more on population dynamics and reduce overfitting (3)
- Mean absolute validation error stabilized at **5.77 days** after approximately 10 epoches

Length of Stay Prediction: LSTM Recurrent Network

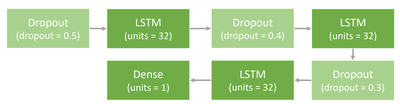
GOAL

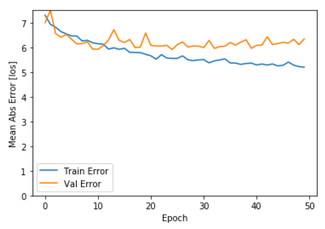
Improve hospital resource utilization by predicting remaining LOS given patient progression in the first five days in the ICU

Patient	Timestep 0	Timestep 1	...	Remaining LOS
P0000	X[0, 0] $\in \mathbb{R}^d$	X[1, 0] $\in \mathbb{R}^d$	...	24 days
P0001	X[0, 1] $\in \mathbb{R}^d$	X[1, 1] $\in \mathbb{R}^d$	...	12 days
...	...	...	...	...

- Input data has shape $\mathbb{R}^{n \times t \times d}$ where n is the number of patients, t was chosen to be 5 timesteps, d is the number of features

NETWORK DESIGN





- LSTM mean absolute validation error stabilized at **6.174 days** after approximately 10 epochs

Conclusion

- Selected features are related to respiratory function (eg. O_2 saturation, expiratory lung pressure), consistent with previous studies (4).
- Anion gap warrants further attention in its role in predicting mortality
- Best LOS prediction (based on selected features) was feedforward with mean absolute validation error of **5.77 days** (compare with 15.27 days baseline and 4.1 days oracle)
- LSTM performance is suboptimal to feedforward and was challenging to tune. Possible reasons:
  - Noisy or missing data (lab and vital sign data not consistently available)
  - Underlying health conditions makes learning generalizable patterns hard

References

1. Mazzei, M. Community-Acquired Pneumonia (CAP). Salem Press Encyclopedia of Health, 2019.
2. Barbier F, Andremont A, Wolff M, Bouadma L. Hospital-acquired pneumonia and ventilator-associated pneumonia: Recent advances in epidemiology and management. Current Opinion in Pulmonary Medicine. 2013.
3. Baldi, P. & Sadowski, P. Understanding dropout. in Advances in Neural Information Processing Systems. 2013.
4. Menéndez R, Cremades MJ, Martínez-Moragón E, Soler JJ, Reyes S, Perpiña M. Duration of length of stay in pneumonia: Influence of clinical factors and hospital type. Eur Respir J. 2003.

Video Link: https://t.ly/pGY9G