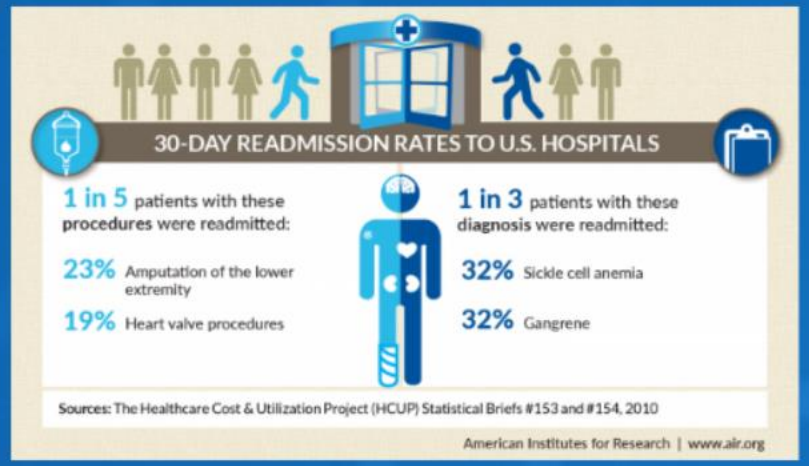
Interpretable Deep Learning Framework for Predicting all-cause 30-day Hospital Readmissions

CSCE 633 Final Project

[Click Here for Prezi Presentation]

Parvez Rafi
Arash Pakbin
Shiva Kumar Pentyala

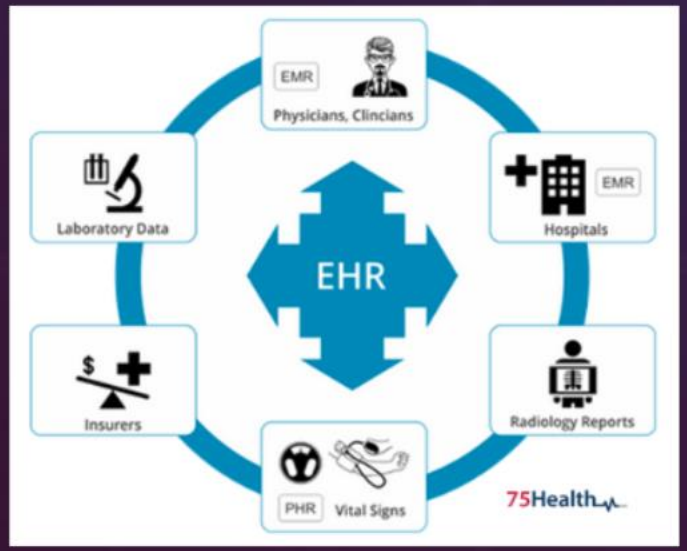
Introduction & Motivation



- 76% readmissions are avoidable.
- High cost of avoidable readmissions ($17 Billion)
- Hospital Readmissions Reduction Program (HRRP).

Problem Formulation

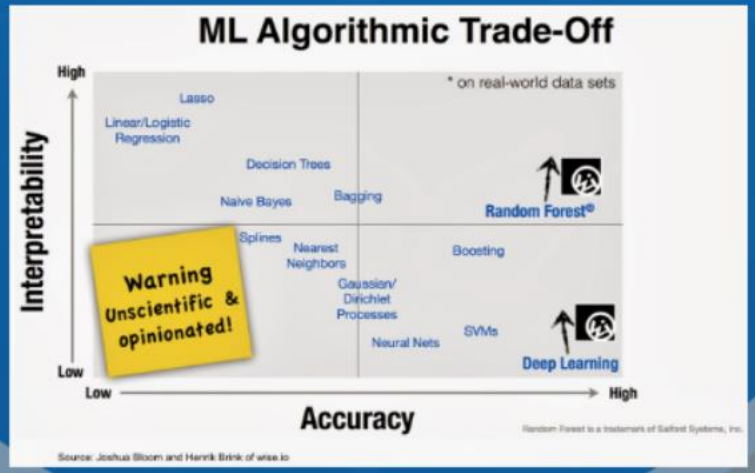
- Improved Decision Support Systems to reduce readmissions.
- Predictive ability of Electronic Health Record(EHR) data.
- Predicitng ICU readmissions using EHR.
- Deep Learning models like RNNs, good for sequential data.
- Slow clinical adoption of "black box" models.



Literature Review

Literature Review

- Tradeoff: Interpretability vs Predictive accuracy (choi, et al, 2016).
- Interpretable Models using Lasso LR (Jovanovic, et al, 2016).
- LSTM RNN for risk prediction (Lipton, et al, 2016).
- Knowledge Distillation approach (Hinton, et al, 2015).
- Interpretable Deep Models for ICU outcome prediction (Che, et al, 2017).

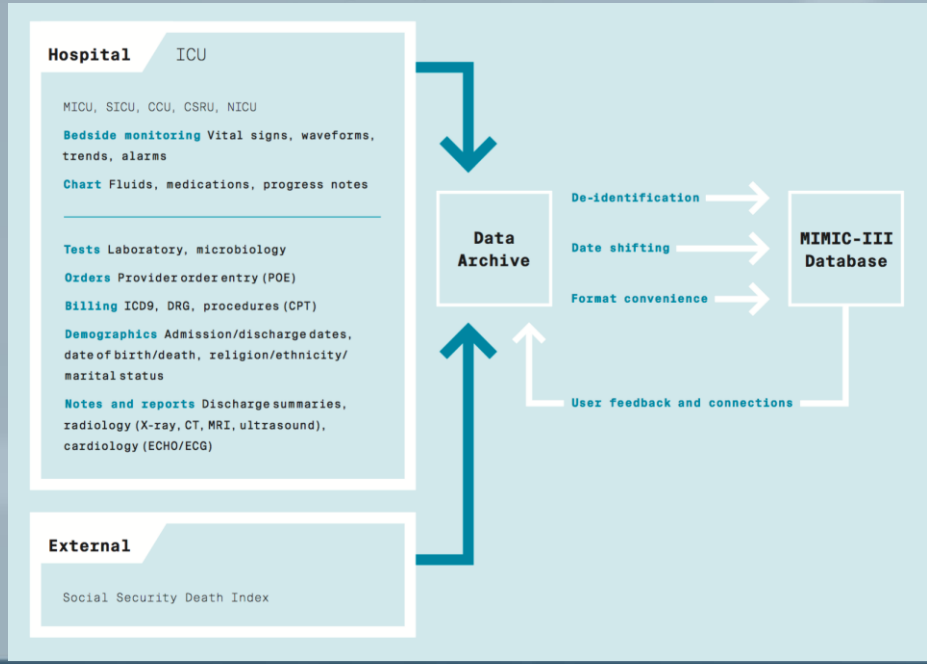


Data Description

MIMIC III

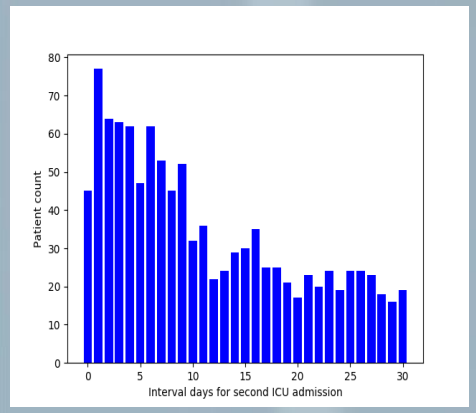
Medical Information Mart for Intensive Care

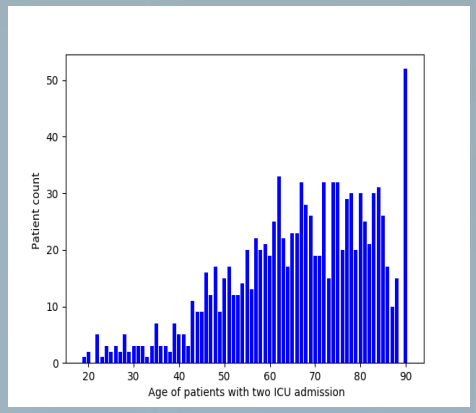
- EHR collected between 2001 and 2012.
- De-identified 58,000 hospital admissions for 38,645 adults and 7,875 neonates.
- Diverse and very large population of ICU patients.
- ~36 GB CSV files
- Observations: demographics, vitals, lab tests, Medications, Medical notes etc.
- Data indested to a PostgreSQL server.

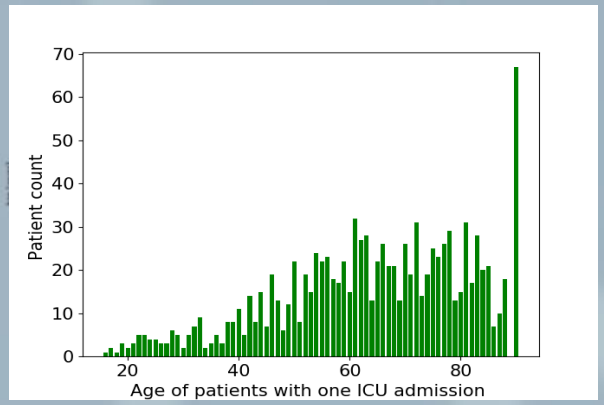


Trends in the Data

- 11.9% adult patients readmitted to ICUs within 30 days.



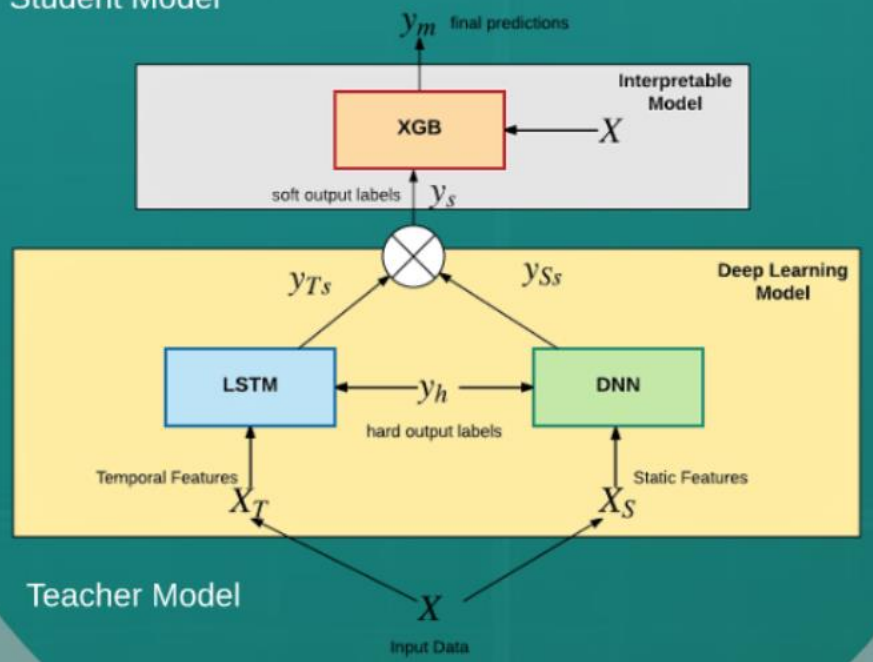




https://opkula.wordpress.com/teaching-philosophy/

Proposed Model

Interpretable Mimic Learning



Data Processing & Feature Engineering

Model Development

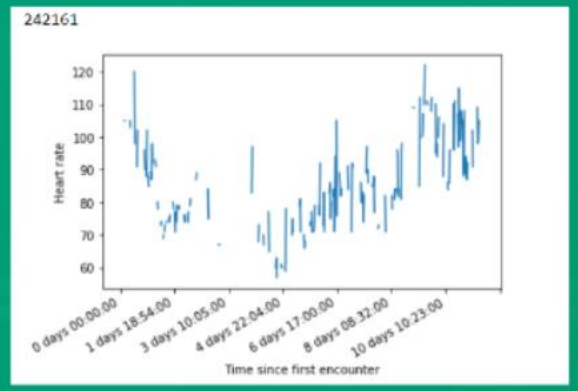
Modelling Sequential Features

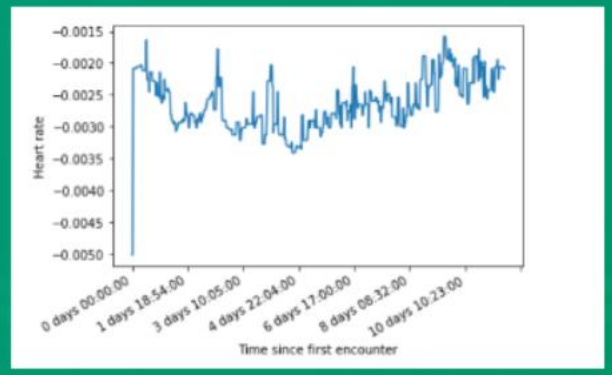
- Examples:blood cultures, ABP [Diastolic]ABP [Systolic], Albumin, Heart Rate, Respiratory rate

- Irregularly spaced time sequences

	itemid	29	50	51	52	79	87	90	92	113	114	...	228100	228299	228305	228394	228395	228396
icustay_id	charttime																	
200021	00:00:00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
	00:07:00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
	00:15:00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
	00:55:00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
	01:00:00	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN

- Selected 17 temporal features.
- Modelling missing data: i. Forward filling ii. 0-imputation





- Feature Normalization

Modelling Static Features

- Examples: Demographics, Insurance, Lab Test Results

```
   subject_id  hadm_id  icustay_id first_careunit last_careunit        age  \
0           3   145834      211552           MICU           MICU  76.526788
1           4   185777      294638           MICU           MICU  47.845044
2           6   107064      228232           SICU           SICU  65.940670
3           9   150750      220597           MICU           MICU  41.790226
4          11   194540      229441           SICU           SICU  50.148292

  gender marital_status insurance  IsReadmitted
0      M        MARRIED  Medicare             0
1      F         SINGLE   Private             0
2      F        MARRIED  Medicare             0
3      M            NaN  Medicaid             0
4      F        MARRIED   Private             0
```

- Generated Severity Scores:

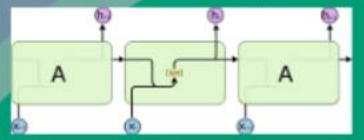
	icustay_id	sirs	oasis	sofa	sapsii	saps
0	200001	0.75	0.600000	0.272727	0.322034	0.454545
1	200003	1.00	0.500000	0.272727	0.254237	0.545455
2	200006	0.50	0.457143	0.045455	0.169492	0.340909

- Selected 52 out 753 lab items
- Feature Normalization

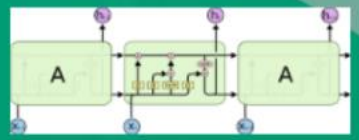
Model Development

Long Short Term Memory (LSTM)

- For Seqential Features



Vanilla RNN



LSTM RNN

```
Layer (type)                 Output Shape              Param #
=================================================================
lstm_1 (LSTM)                (None, 200, 256)          322560
_________________________________________________________________
dropout_1 (Dropout)          (None, 200, 256)          0
_________________________________________________________________
time_distributed_1 (TimeDist (None, 200, 1)            257
=================================================================
Total params: 322,817
Trainable params: 322,817
Non-trainable params: 0
_________________________________________________________________
```

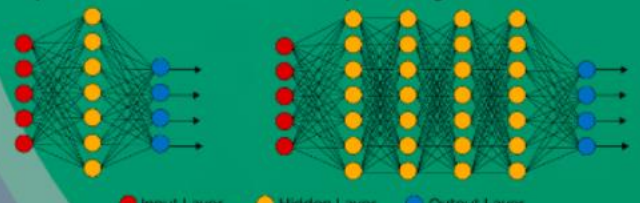
Model Parameters

Deep Neural Network

- For non-temporal features

Simple Neural Network

Deep Learning Neural Network




```
Layer (type)                 Output Shape              Param #
=================================================================
dense_1 (Dense)              (None, 80)                6480
_________________________________________________________________
dense_2 (Dense)              (None, 40)                3240
_________________________________________________________________
dense_3 (Dense)              (None, 10)                410
_________________________________________________________________
dropout_1 (Dropout)          (None, 10)                0
_________________________________________________________________
dense_4 (Dense)              (None, 1)                 11
=================================================================
Total params: 10,141
Trainable params: 10,141
Non-trainable params: 0
_________________________________________________________________
```

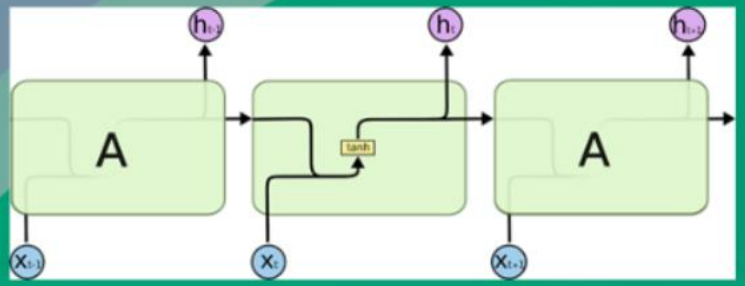
Gradient Boosting Machines

- For interpretable Models
- Feature Selection

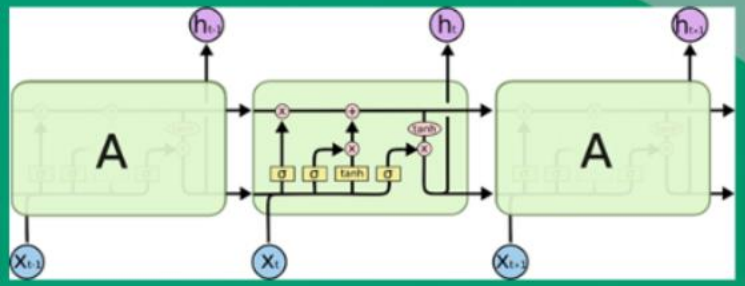
src: https://blog.bigml.com/2017/03/14/introduction-to-boosted-trees/

Long Short Term Memory (LSTM)

- For Seqential Features



Vanilla RNN



LSTM RNN

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_1 (LSTM)                (None, 200, 256)          322560
_________________________________________________________________
dropout_1 (Dropout)          (None, 200, 256)          0
_________________________________________________________________
time_distributed_1 (TimeDist (None, 200, 1)            257
=================================================================
Total params: 322,817
Trainable params: 322,817
Non-trainable params: 0
_________________________________________________________________
```

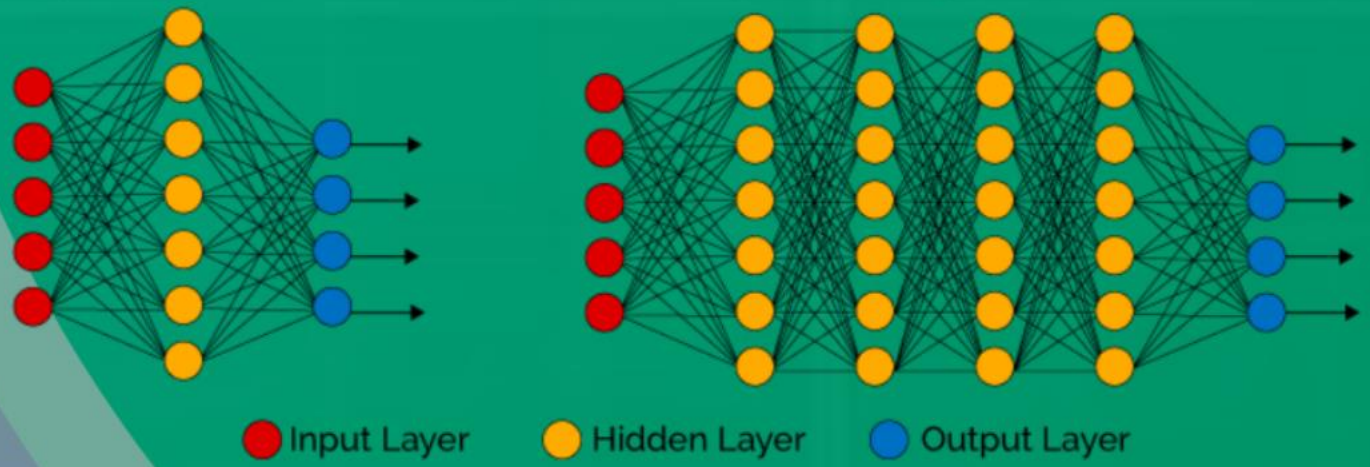
Model Parameters

Deep Neural Network

- For non-temporal features

Simple Neural Network

Deep Learning Neural Network



Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 80)	6480
dense_2 (Dense)	(None, 40)	3240
dense_3 (Dense)	(None, 10)	410
dropout_1 (Dropout)	(None, 10)	0
dense_4 (Dense)	(None, 1)	11

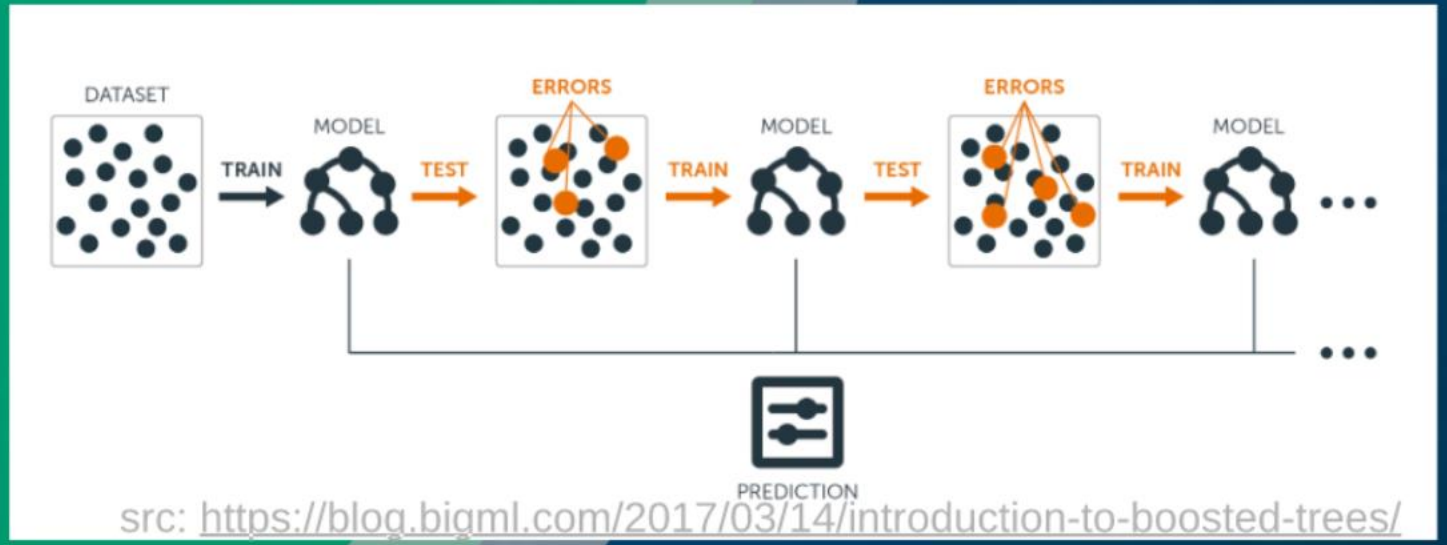
Total params: 10,141
Trainable params: 10,141
Non-trainable params: 0

Gradient

- For interpret
- Feature Sele

XG Boost/ GBM

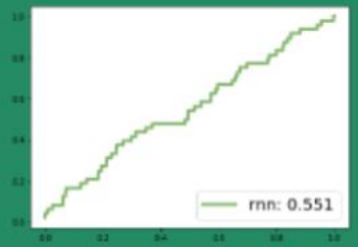
- For interpretable Models
- Feature Selection



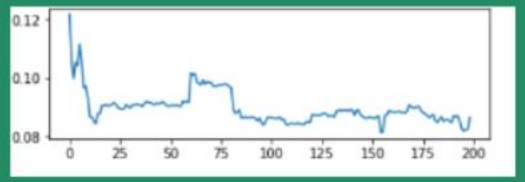
src: https://blog.bigml.com/2017/03/14/introduction-to-boosted-trees/


Perfomance of LSTM
Performance of DNN
Interpretation

Perfomance of LSTM

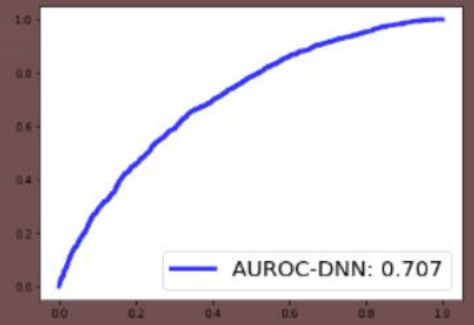


- Model dependent on Maxlen of sequence passed to LSTM
- Performance Data Dependent



Variation of 30-day ICU readmission probability for a patient

Perfomance of DNN



- Deep Neural Networks performed better compared to LSTM
- AUROC decreased when the hyperparameter(epochs) increased beyond 60

```
# Fit the model
model.fit(X_tr, y_tr, epochs=60, batch_size=20)

Epoch 1/60
38032/38032 [==============================] - 2s - loss: 0.3629 - acc: 0.8814
Epoch 2/60
38032/38032 [==============================] - 2s - loss: 0.3523 - acc: 0.8822
Epoch 3/60
38032/38032 [==============================] - 2s - loss: 0.3474 - acc: 0.8822
Epoch 4/60
38032/38032 [==============================] - 2s - loss: 0.3446 - acc: 0.8822
Epoch 5/60
38032/38032 [==============================] - 2s - loss: 0.3438 - acc: 0.8822
Epoch 6/60
38032/38032 [==============================] - 2s - loss: 0.3437 - acc: 0.8822
Epoch 7/60
38032/38032 [==============================] - 2s - loss: 0.3425 - acc: 0.8822
Epoch 8/60
38032/38032 [==============================] - 2s - loss: 0.3417 - acc: 0.8822
Epoch 9/60
```

Perfomance of XGB

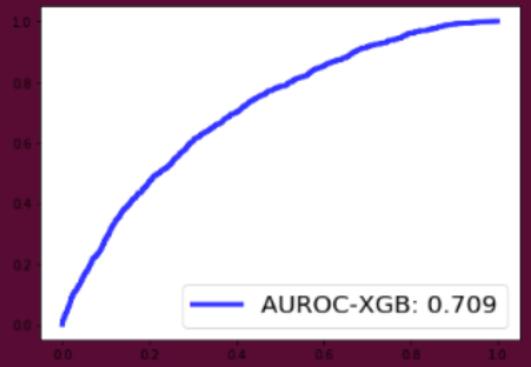
XGBoost: Interpretable Model

Grid-Search

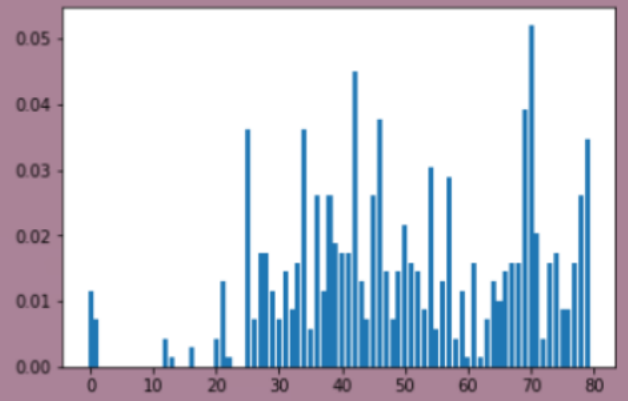
```
colsample_bytree: 0.7
learning_rate: 0.05
max_depth: 6
min_child_weight: 11
missing: -999
n_estimators: 5
nthread: 4
objective: 'binary:logistic'
seed: 1337
silent: 1
subsample: 0.8
```

Best set of parameters through grid-search cross-validation

AUROC



Important Predictors



Feature Scores

RDW
Bicarbonate
PTT
Creatine Kinase (CK)
Base Excess
pO2
saps
Urea Nitrogen
Eosinophils
Sodium, Whole Blood
Albumin
sapsii


That's all Folks!