Machine Learning Based Prediction of In-Hospital Mortality with Acute Myocardial Infarction

Julian S. Haimovich BS[1,2,3], Chenxi Huang, MS PhD[2,3], Shu-Xia Li PhD[2], Bobak J. Mortazavi PhD[2,3,4],and Harlan M. Krumholz MD SM[2,3]

[1]Albert Einstein College of Medicine, Bronx, NY; [2]Yale-New Haven Hospital Center for Outcomes Research and Evaluation, New Haven, CT; [3]Yale University School of Medicine, Department of Internal Medicine, New Haven, CT; [4]Department of Computer Science & Engineering, Texas A&M University, College Station, TX

Background

- ❖ Nearly 1 million patients in the United States are hospitalized with an acute myocardial infarction (AMI) each year, and between 3 and 8% of these patients do not survive to discharge
- ❖ Accurately predicting in-hospital outcomes for patients with AMI has the potential to:
  1. Aid in risk-stratification and management of patients presenting with AMI
  2. Improve retrospective analysis of hospital performance in the care of AMI patients
- ❖ Past modeling efforts (McNamara et al.) employed logistic regression with backward selection to produce a parsimonious variable set for predictions (C-stat = 0.87), but this study was limited by the inclusion of only a partial sample of the available cohort (22%) and patient variables (28%)

Objectives

- ❖ To determine if the application of machine learning techniques can improve prediction of in-hospital mortality in patients with AMI compared with previous models
- ❖ To compare the performance of different machine learning approaches

Data

- ❖ Patient data is taken from ACTION-GWTG registry, a national quality improvement registry for AMI collected from 655 participating hospitals over 10 years, and encompassing over 1 million patients
- ❖ Models are built on 96 patient variables available at time of presentation including history, risk factors, demographics, and initial laboratory values (except where otherwise noted for LR model)

Methods

<u>Four</u> modeling approaches are considered here:

1. <u>Logistic Regression (LR)</u>: fits a conventional logistic regression model using the same 9 parsimonious variables as included in the McNamara et al. study
2. <u>Logistic Regression with Lasso (Lasso)</u>: couples a conventional logistic regression approach with a cost function (Lasso), which results in a parsimonious set of variables that maximizes predictive capabilities
3. <u>Gradient Descent Boosting (XGBoost)</u>: leverages the creation of many weak decision trees to produce a final, accurate prediction via weighted majority vote ("boosting"); unlike logistic regression, XGBoost is able to account for non-linear, higher-order interactions among variables
4. <u>Meta-classifier Approach (Meta)</u>: uses an XGBoost model to combine the output of four models including Logistic Regression with Lasso, XGBoost, a Neural Network, and K-Nearest Neighbors.

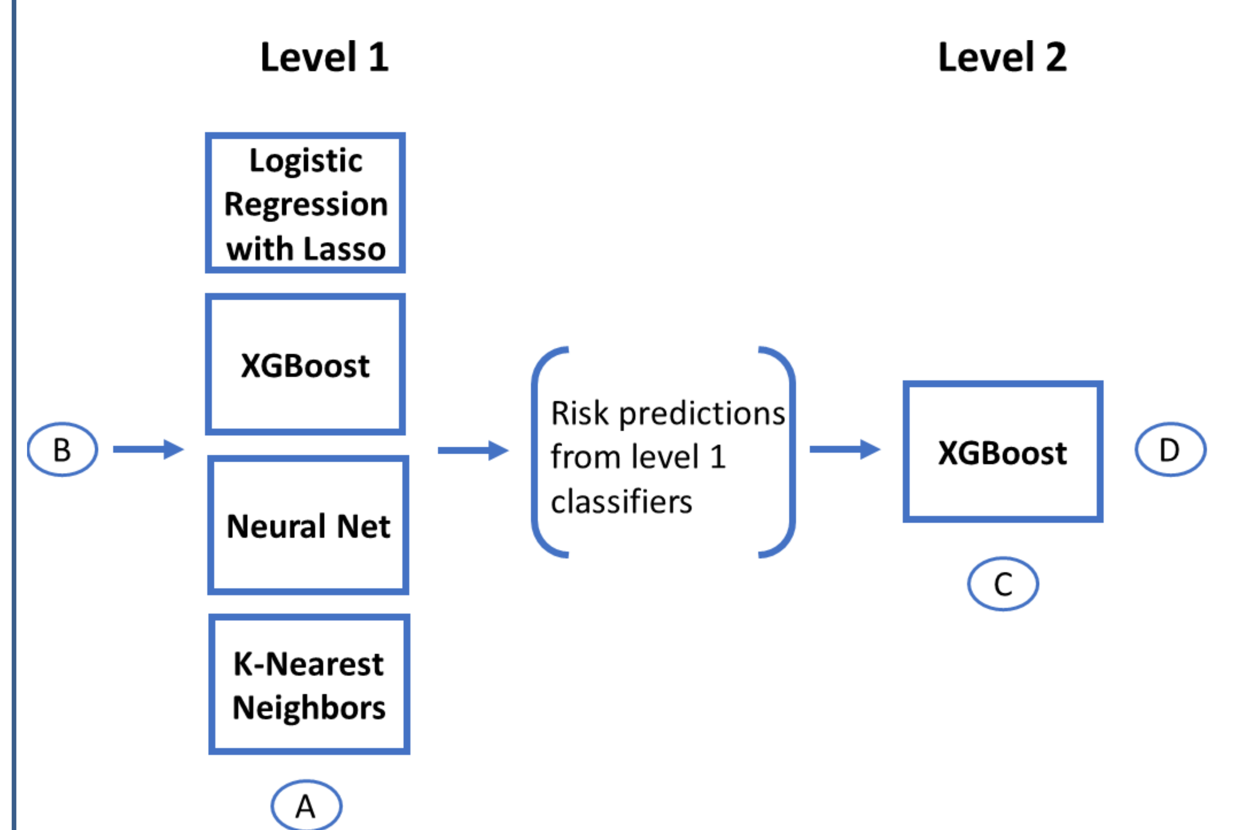
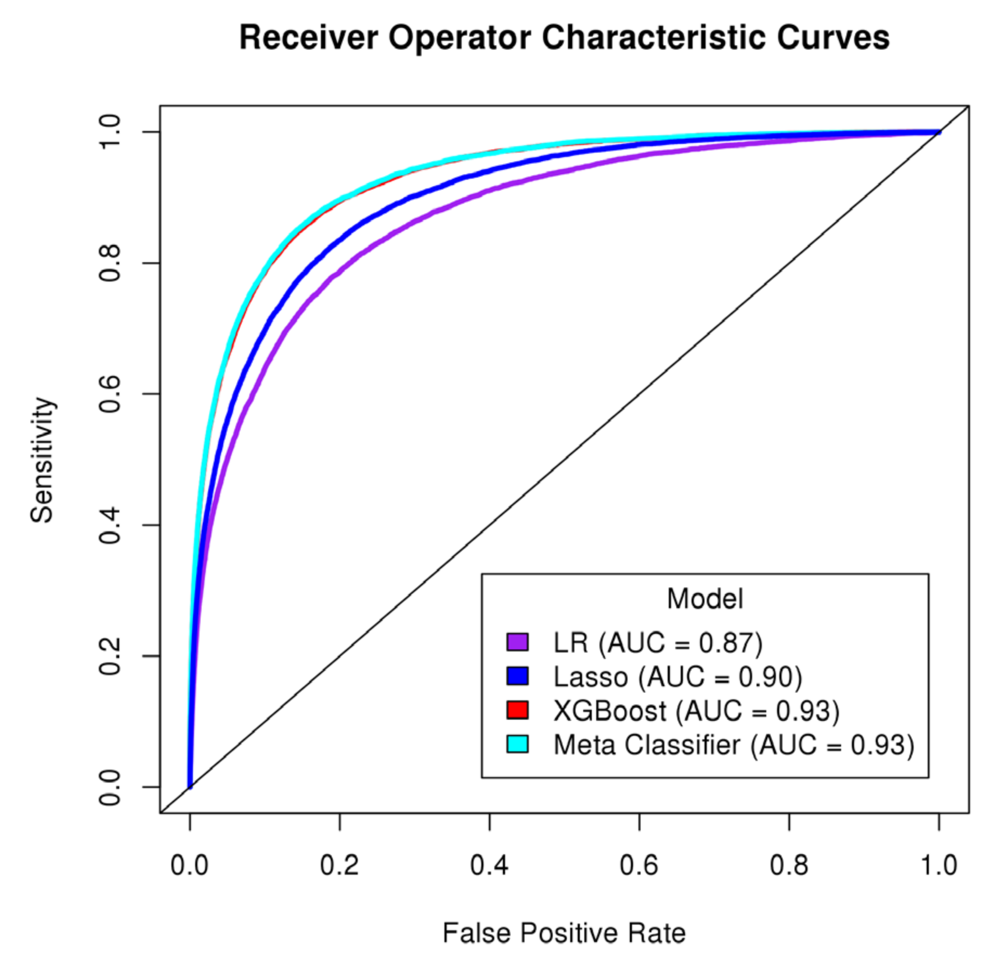


Figure 1. Computational approach. Level 1 classifiers consist of four independent models each trained on the same initial 40% training sample (A). The next 40% training sample (B) is then input into the Level 1 classifiers, resulting in one risk estimate from each Level 1 model. These four risk estimates are then used to train the Level 2 XGBoost classifier (C). A final sample (D) is used to test the performance of the Level 1 and Level 2 classifiers.

Results

❖ <u>**Figure 2: Receiver-Operator Characteristic Curves for LR, Lasso, XGBoost, and Meta models**</u>



❖ <u>**Table 1: Summary of model performance for LR, Lasso, XGBoost, and Meta models**</u>

	LR	Lasso	XGBoost	Meta
ROC AUC (C-statistic)	**0.872**	**0.900**	**0.929**	**0.930**
PR AUC	0.36	0.42	0.55	0.55
F-score	0.41	0.45	0.53	0.53
Sensitivity	0.41	0.48	0.54	0.54
Specificity	0.97	0.97	0.98	0.98
PPV	0.41	0.42	0.51	0.52
NPV	0.97	0.97	0.98	0.98
Brier Score Decomposition				
Reliability ($x10^{-6}$)	15.9 +/- 4.5	43.0 +/- 9.9	7.1 +/- 2.8	1.3 +/- 1.8
Resolution ($x10^{-3}$)	7.4 +/- 0.2	7.4 +/- 0.1	9.6 +/- 0.2	9.8 +/- 0.2
Uncertainty	0.044	0.044	0.044	0.044
Overall	0.38	0.037	0.035	0.034

Results

❖ <u>**Table 2: Shift table comparison of individual risk estimates from Lasso and XGBoost/Meta models**</u>

		Lasso risk		
		Low (< 1%)	Moderate (1-5%)	High (> 5%)
XGBoost risk	Low (< 1%)	0.2% (88,777)	0.5% (43,080)	0.4% (677)
	Moderate (1-5%)	1.8% (3,233)	2.2% (41,069)	3.4% (13,301)
	High (> 5%)	9.5% (258)	11.8% (6,473)	26.1% (30,630)
Meta risk	Low (< 1%)	0.2% (89,567)	0.5% (47,061)	0.4% (565)
	Moderate (1-5%)	2.0% (2,418)	2.2% (36,636)	2.9% (11,263)
	High (> 5%)	9.7% (310)	11.8% (6,925)	24.8% (32,780)

Table 2. Each cell represents a cohort of patients whose individual Lasso risk falls within the Lasso range and whose individual XGBoost/meta classifier risk falls within the given XGBoost/meta risk range. Event rate is given as a percentage for each cohort, and the sample size is shown in parentheses.

Conclusions

- ❖ Machine learning based approaches outperform conventional logistic regression in predicting in-hospital mortality with AMI, and therefore have the potential to both enhance hospital-specific risk adjustment for retrospective profiling, and improve risk-stratification of AMI patients
- ❖ Amongst the machine learning methods, non-linear models such as XGBoost and the meta-classifier outperform the linear Lasso model in predicting in-hospital mortality with AMI

References

1. Benjamin, Emelia J., et al. "Heart Disease and Stroke Statistics. 2017 Update: A Report From the American Heart Association." *Circulation*, vol. 135, no. 10, 2017, doi:10.1161/cir.0000000000000485.
2. McNamara, Robert L., et al. "Predicting In-Hospital Mortality in Patients With Acute Myocardial Infarction." *Journal of the American College of Cardiology*, vol. 68, no. 6, 2016, pp. 626–635., doi:10.1016/j.jacc.2016.05.049.