

PREDICTING ICU MORTALITY

Isaac Kim
Metis Fellow



METRICS:

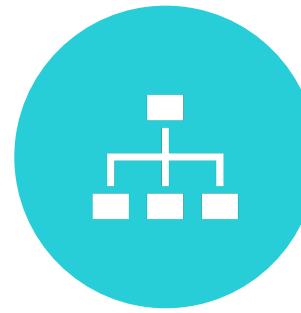
GCS

SAPS





COLLECT / CLEAN
DATA



CREATE
CLASSIFICATION
MODEL



IMPROVE PATIENT
CARE

OBJECTIVE

EXPLORATORY DATA ANALYSIS

4000 TRAIN,
4000 TEST

FEATURES:
PATIENT INFO,
LABS

TIME SERIES DATA:
BLOOD PRESSURE



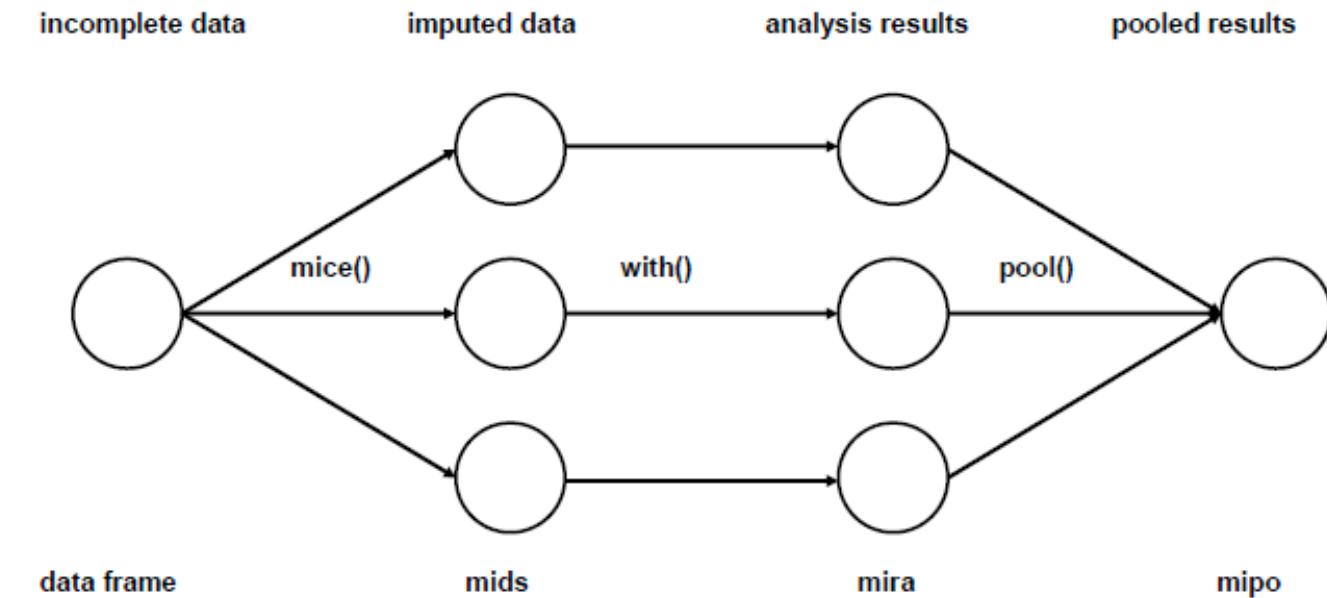
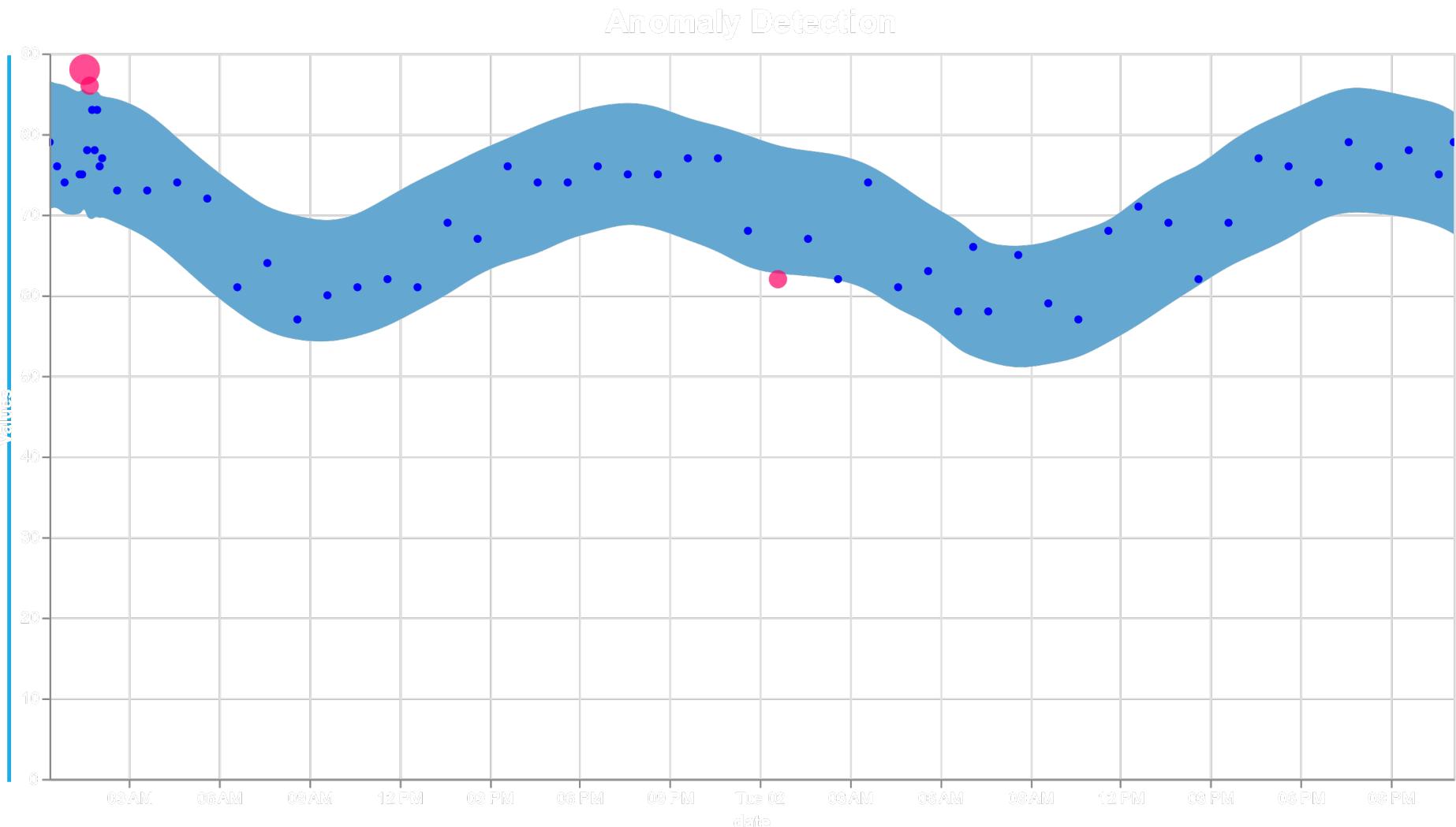


Figure 1: Main steps used in multiple imputation.

MULTIPLE IMPUTATION BY CHAINED EQUATIONS

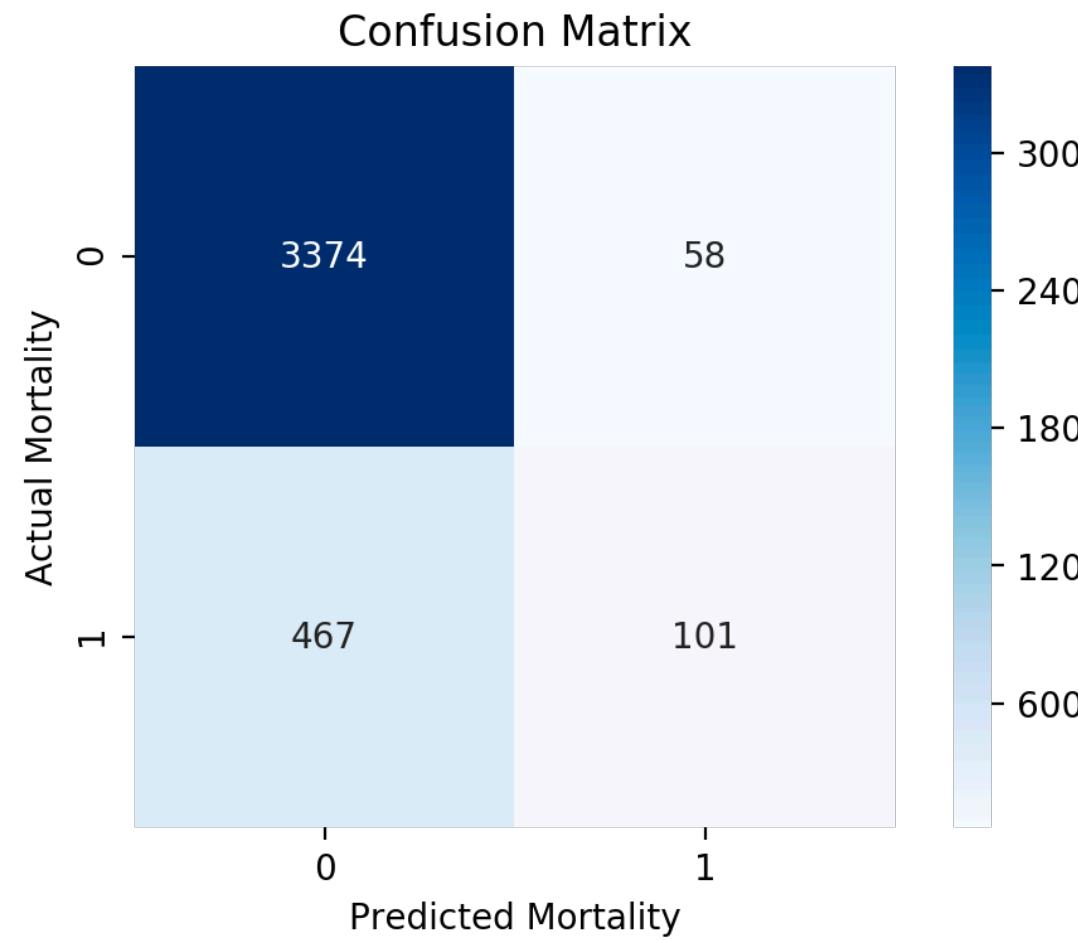
ANOMALY DETECTION WITH FBPROPHET

PROPHET

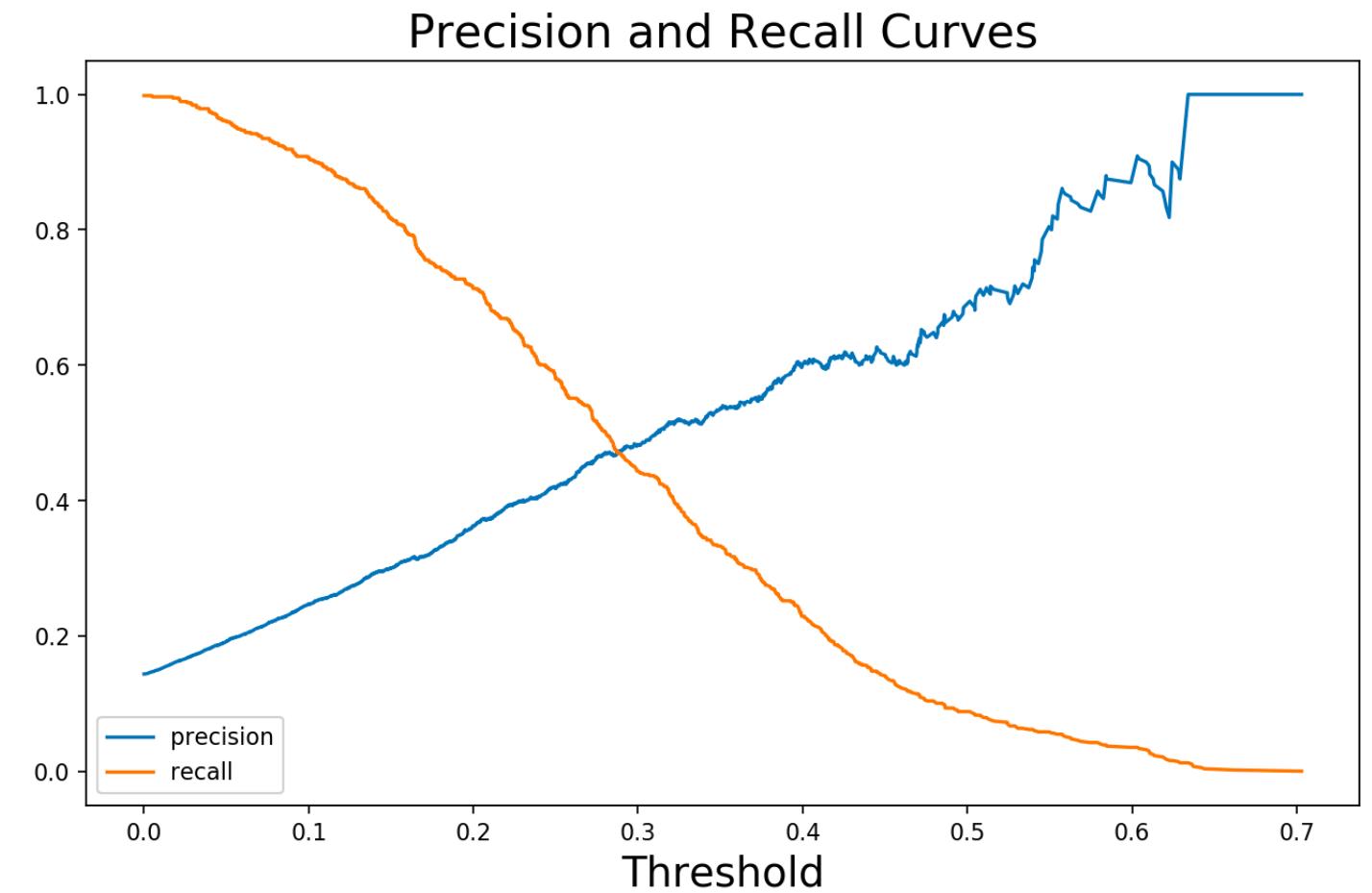


SCORING METRICS

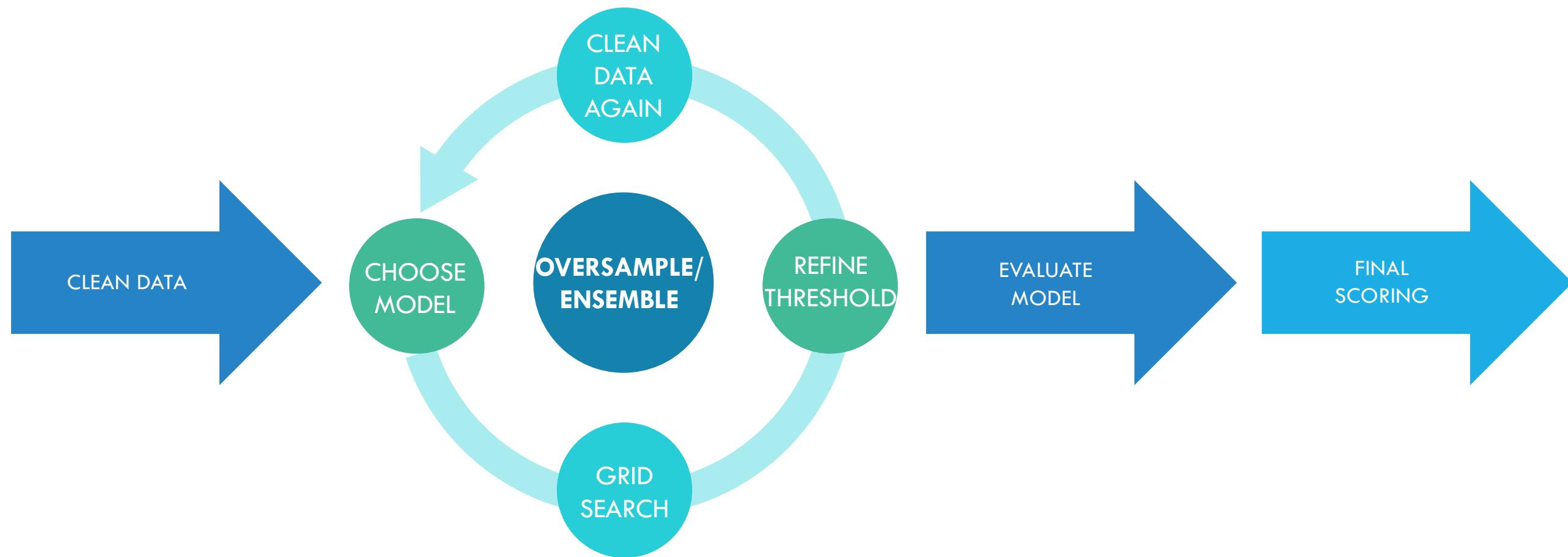
F1 Score

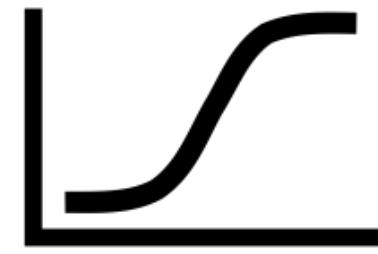


Adjusting Probability Threshold

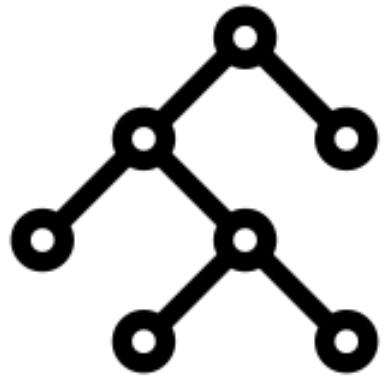


MACHINE LEARNING PROCESS





LOGISTIC
REGRESSION

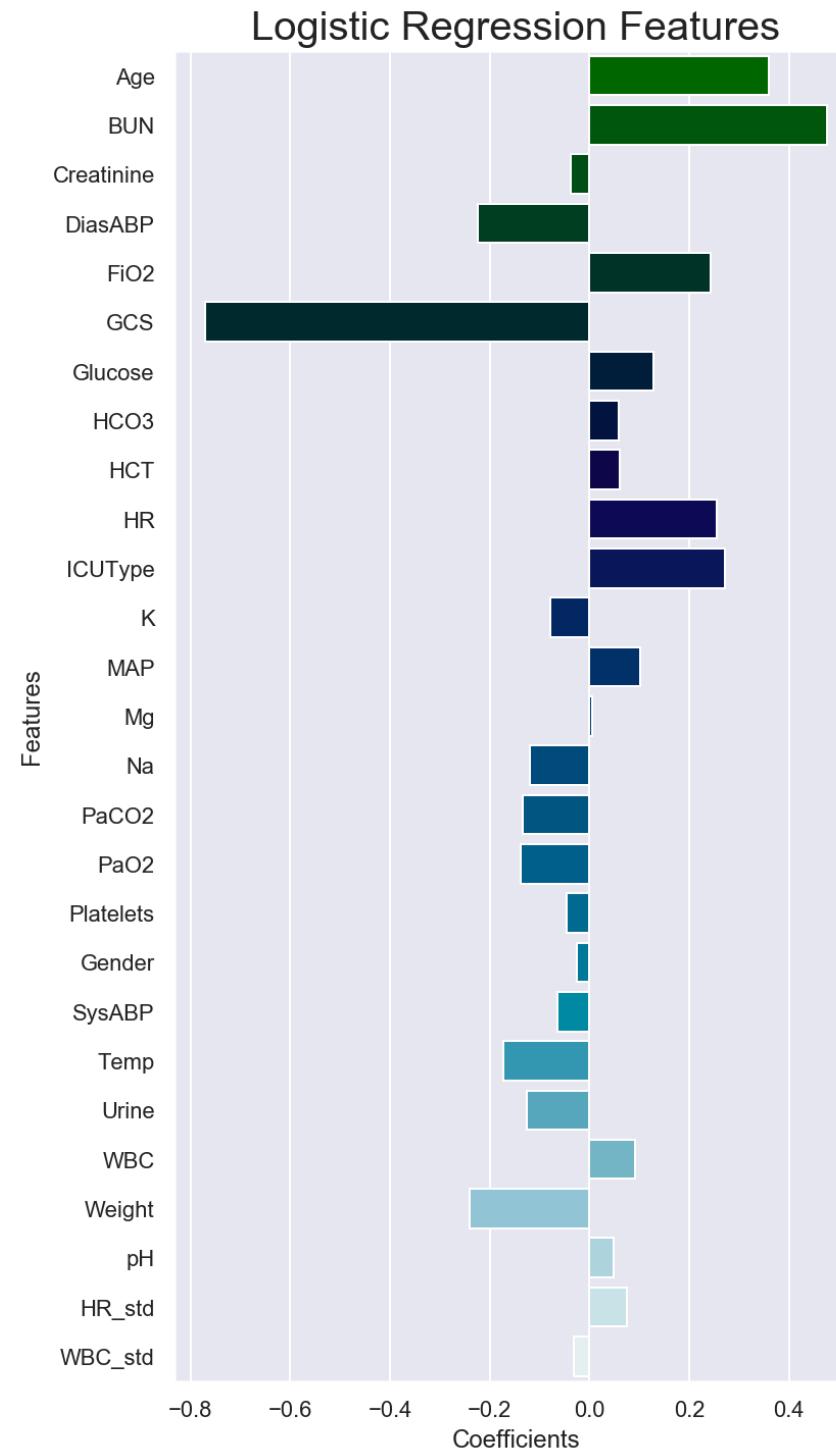


RANDOM FOREST

XGBoost

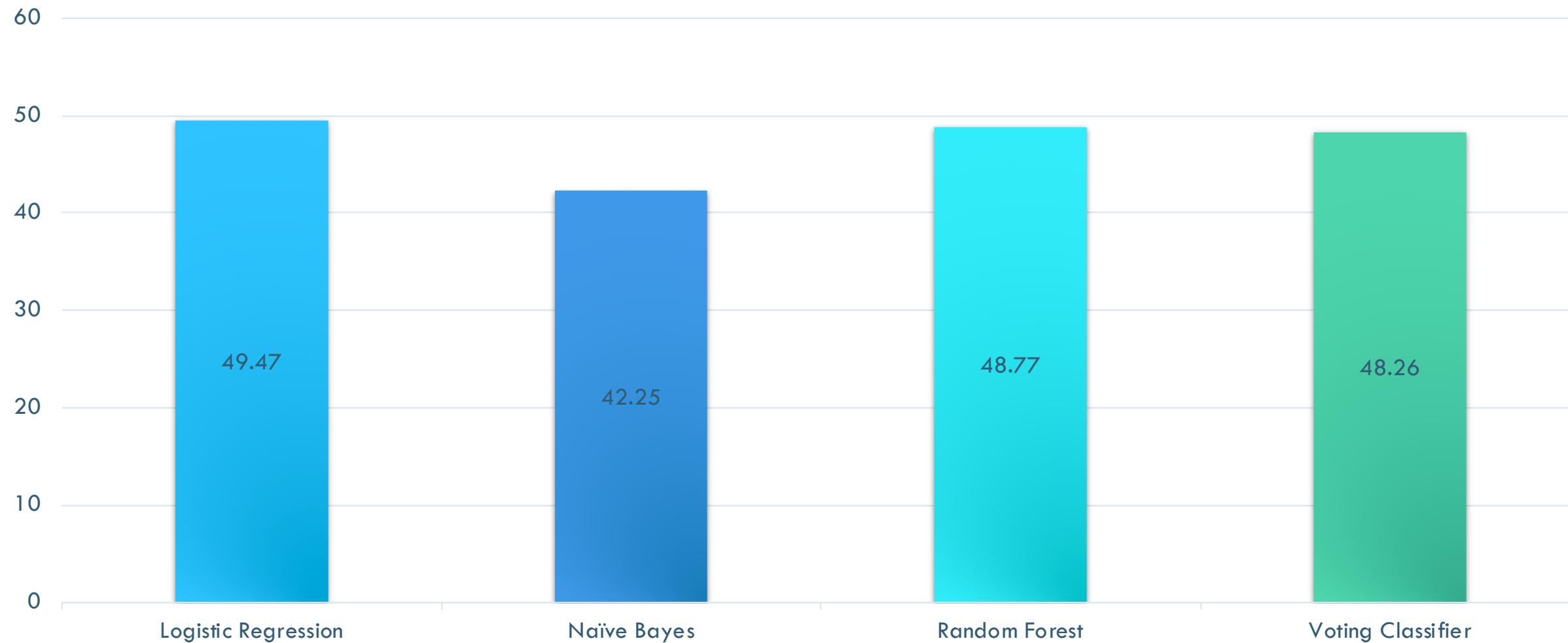
MODELS

KEY FEATURES



- Blood Pressure (Invasive)
- Age
- BUN (2.5-7.1 mmol/L)
- Glasgow Coma Scale: 3-15, higher is better

MODEL SCORES



THE BEST MODEL BY FAR

Test Set Score: 51.35

Final Score: 52.05

Grid Search CV Parameters:

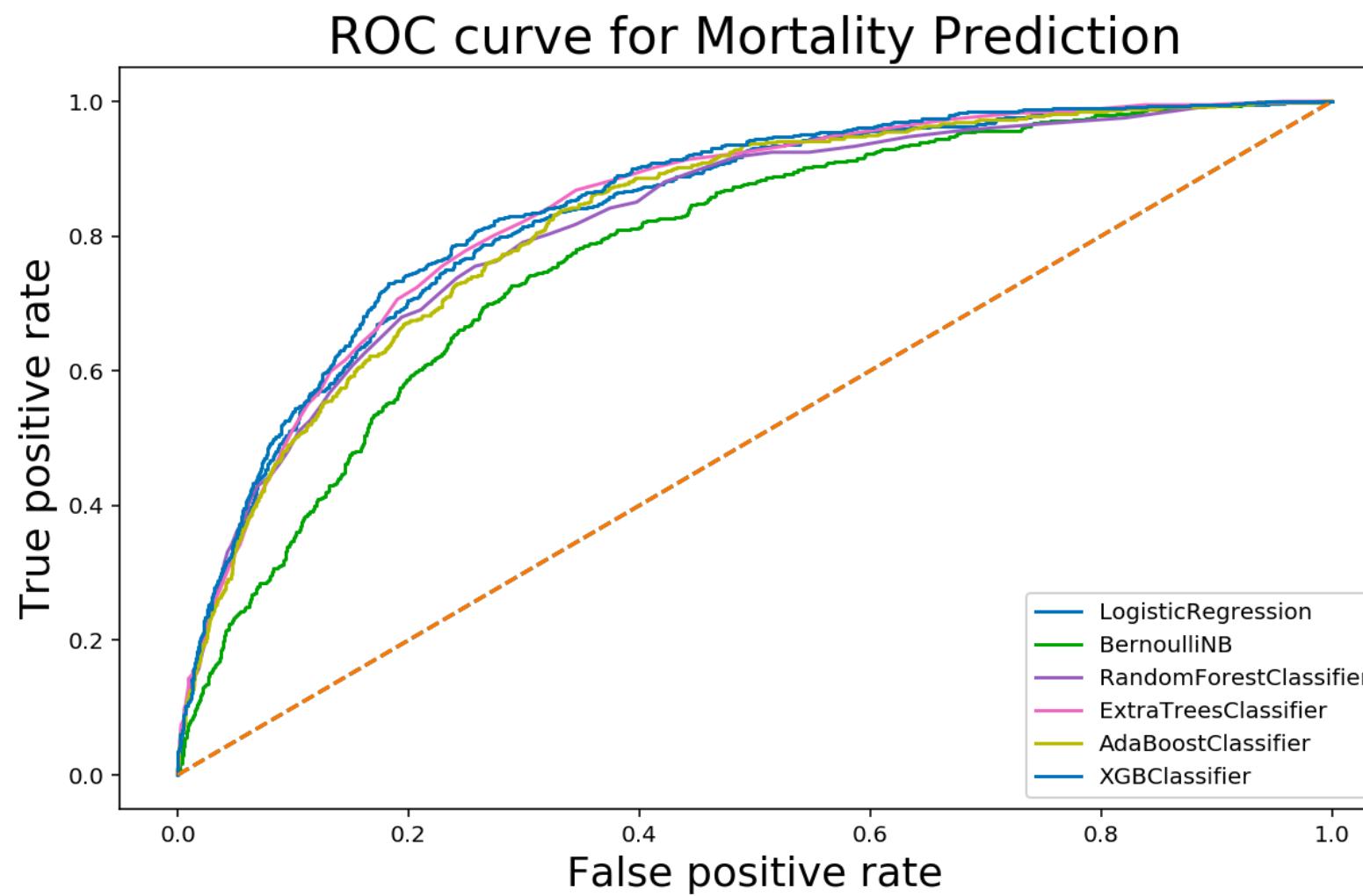
colsample_bytree: 0.8 gamma: 1

learning_rate: 0.1 max_depth: 3

subsample: 0.9 threshold: .241



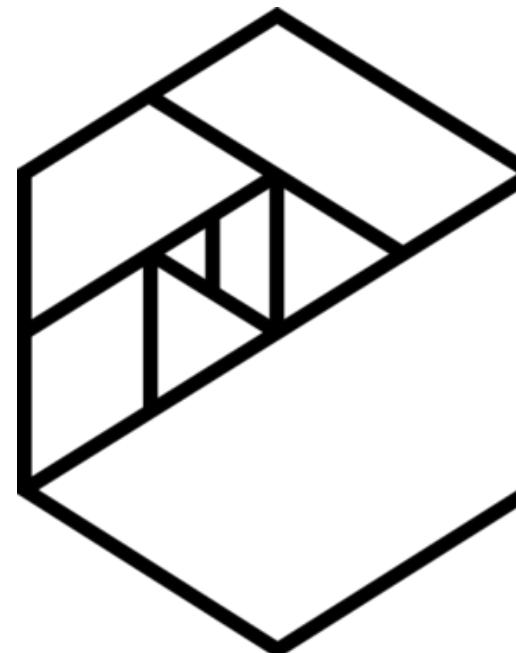
COMPARING AUC



Logistic Regression	.832
Bernoulli Naïve Bayes	.776
Random Forest	.820
Extra Trees	.837
Ada Boost	.824
XGBoost	.846

RECOMMENDATIONS

- Model can be tuned to favor precision/recall
- Make sure hospitals take Lactate readings
- Research ways to find better/uniform measurements



METIS

THANK YOU

GITHUB.COM/ISAACKIM0537/ISAAC_KIM_PROJECT_03

ISAAC.KIM.D@GMAIL.COM

APPENDIX



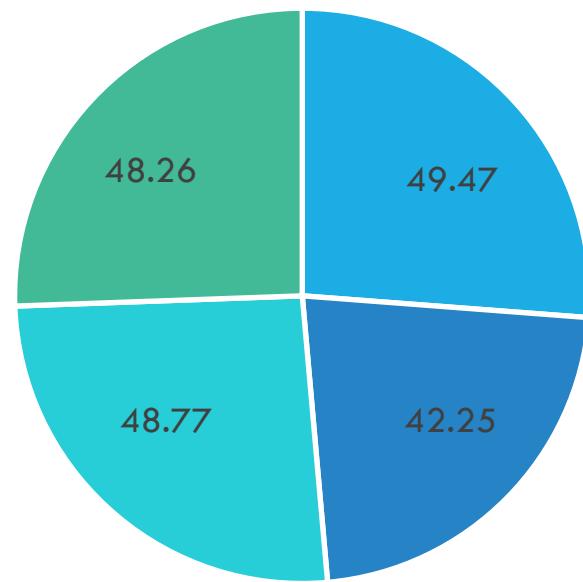
GOING FORWARD

- Need more data
- Better cleaning
- Domain expertise
- Anomaly Detection
Methods revisited



FUN STUFF

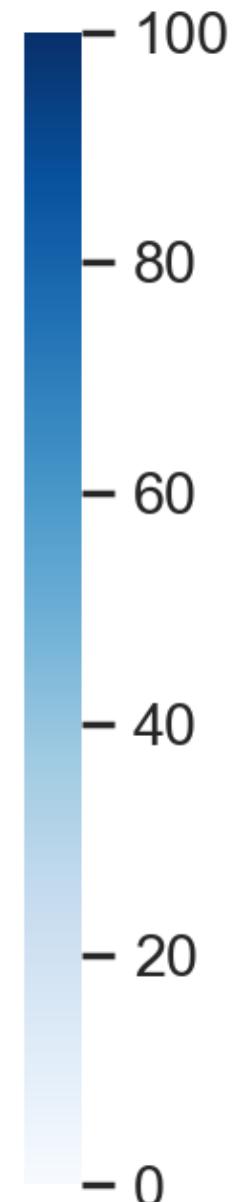
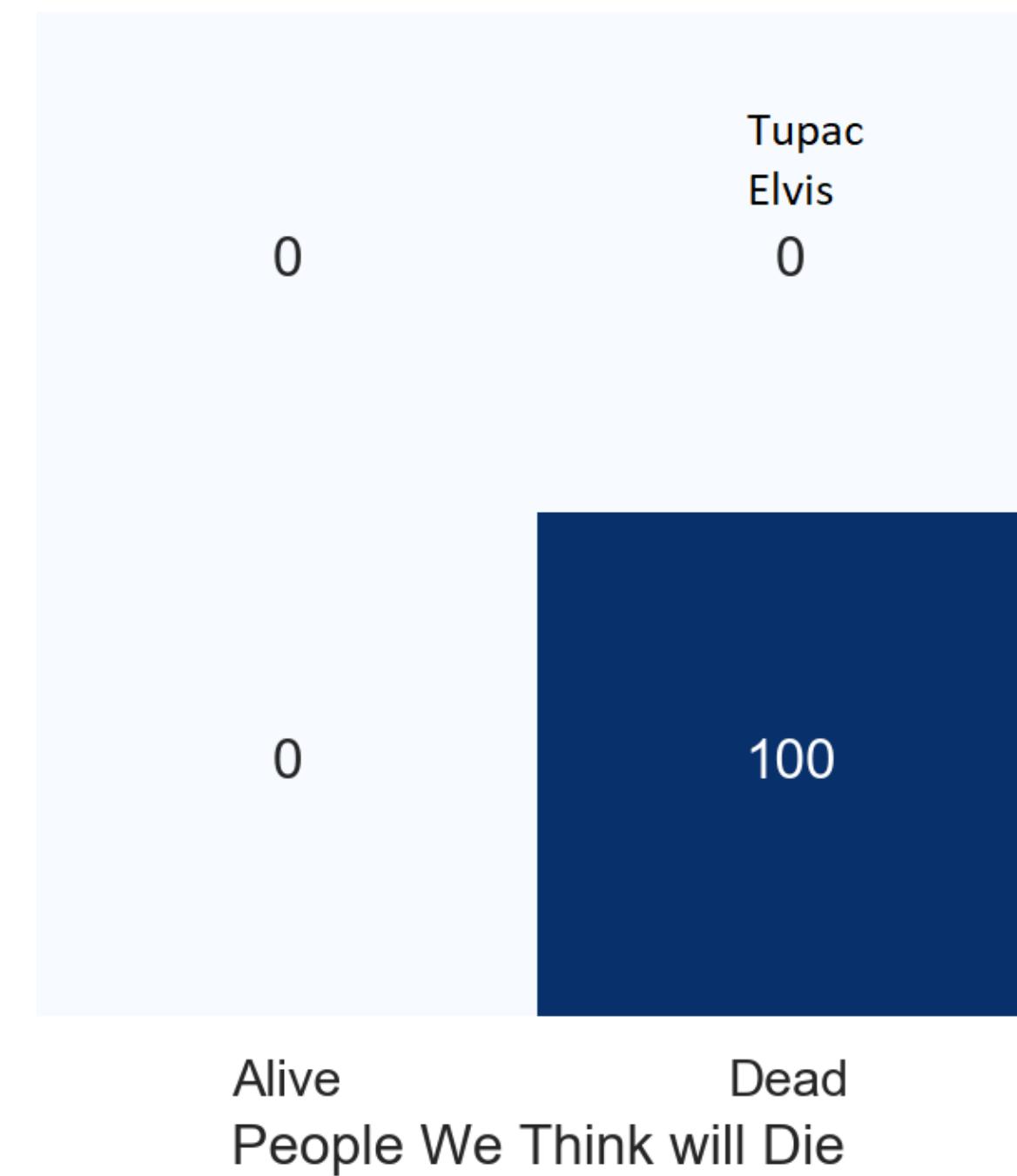
Model Performance



■ Logistic Regression ■ Naive Bayes
■ Random Forest ■ Voting Classifier

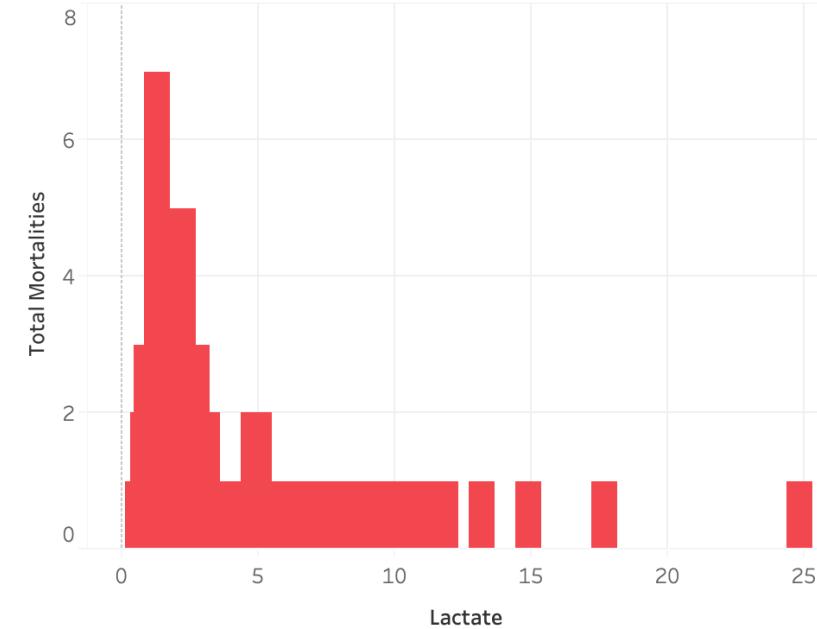
People Who Actually Die
Alive Dead

Confusion Matrix

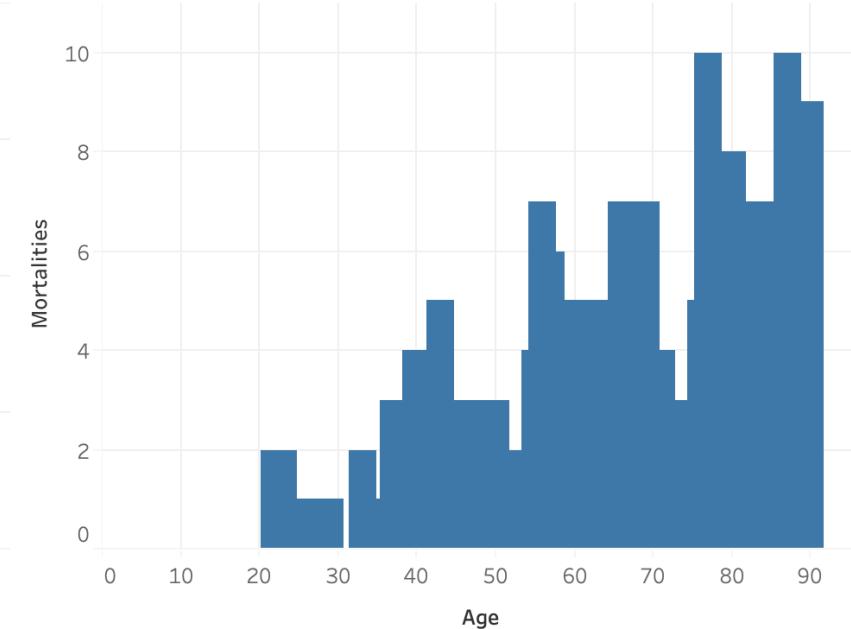


KEY FEATURES

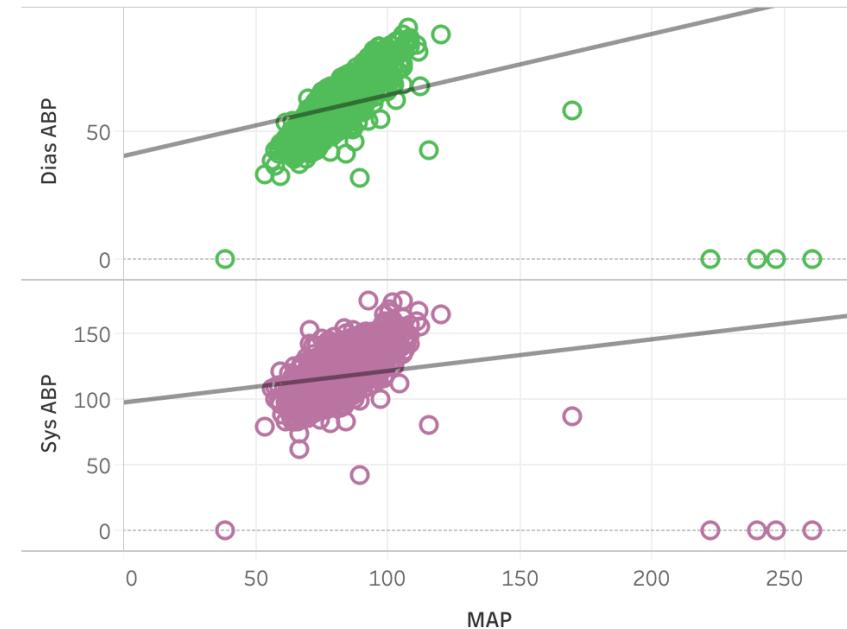
Mortalities by Lactate Measure



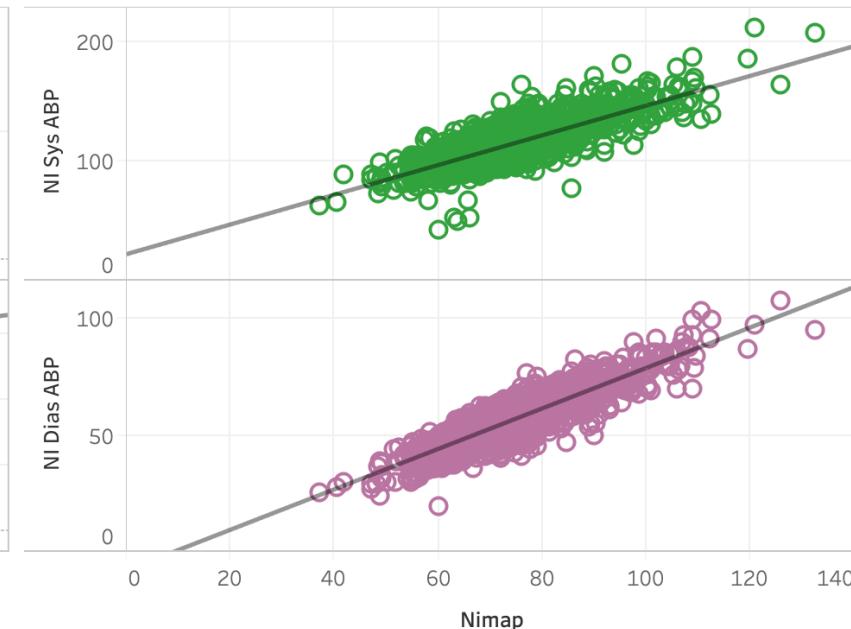
Total Mortalities by Age



Blood Pressure Measurements

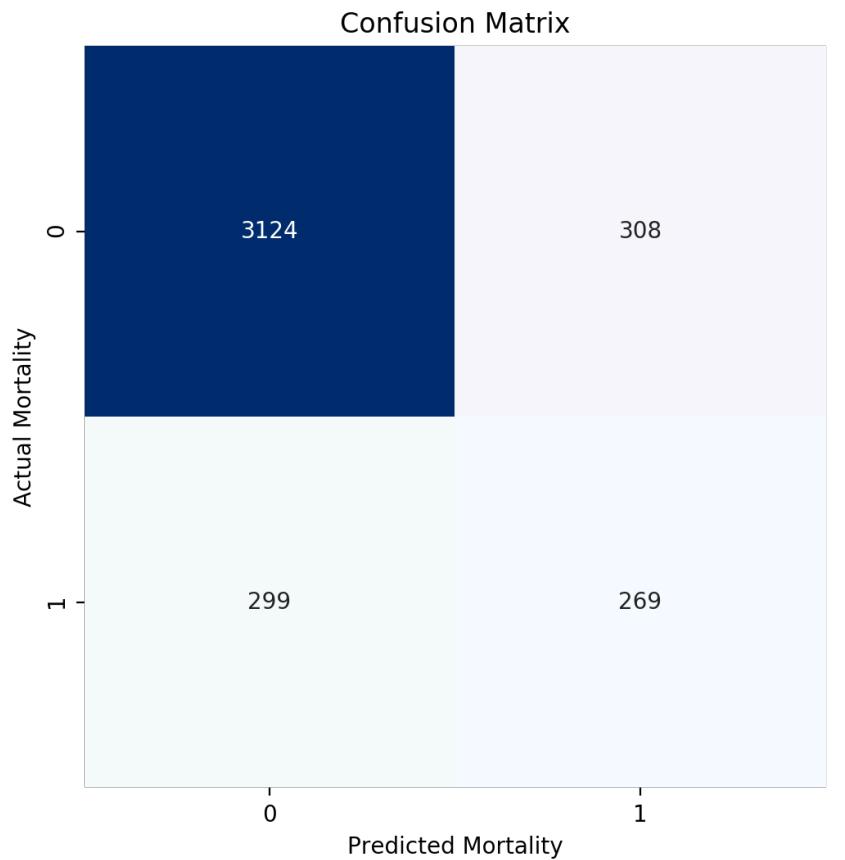


Non-Invasive BP

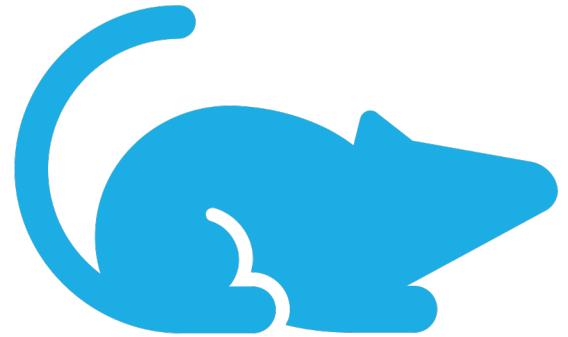
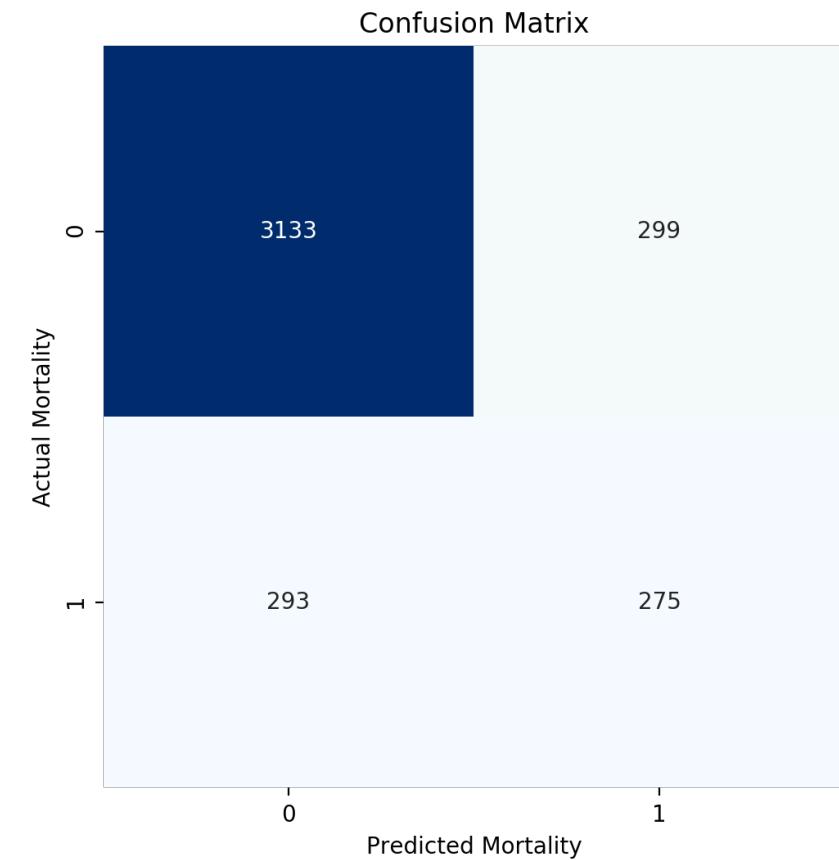


MICE

Mean Imputation: 46.8



MICE Imputation: 48.0



PROPHET

Final Score 48.77 with Random Forest

Importance = Deviation Score

Fourier Series for Modeling

Changepoints (Lag) Laplace Transform

Math heavy, not as accurate as Arima, but much lighter computationally

FEATURES

- Albumin (g/dL)
 - ALP [Alkaline phosphatase (IU/L)]
 - ALT [Alanine transaminase (IU/L)]
 - AST [Aspartate transaminase (IU/L)]
 - Bilirubin (mg/dL)
 - BUN [Blood urea nitrogen (mg/dL)]
 - Cholesterol (mg/dL)
 - Creatinine [Serum creatinine (mg/dL)]
 - DiasABP [Invasive diastolic arterial blood pressure (mmHg)]
 - FiO2 [Fractional inspired O₂ (0-1)]
 - GCS [Glasgow Coma Score (3-15)]
 - Glucose [Serum glucose (mg/dL)]
 - HCO3 [Serum bicarbonate (mmol/L)]
 - Hct [Hematocrit (%)]
 - HR [Heart rate (bpm)]
 - K [Serum potassium (mEq/L)]
 - Lactate (mmol/L)
 - Mg [Serum magnesium (mmol/L)]
 - MAP [Invasive mean arterial blood pressure (mmHg)]
 - MechVent [Mechanical ventilation respiration (0:false, or 1:true)]
 - Na [Serum sodium (mEq/L)]
 - NIDiasABP [Non-invasive diastolic arterial blood pressure (mmHg)]
 - NIMAP [Non-invasive mean arterial blood pressure (mmHg)]
 - NISysABP [Non-invasive systolic arterial blood pressure (mmHg)]
 - PaCO2 [partial pressure of arterial CO₂ (mmHg)]
 - PaO2 [Partial pressure of arterial O₂ (mmHg)]
 - pH [Arterial pH (0-14)]
 - Platelets (cells/nL)
 - RespRate [Respiration rate (bpm)]
 - SaO2 [O₂ saturation in hemoglobin (%)]
 - SysABP [Invasive systolic arterial blood pressure (mmHg)]
 - Temp [Temperature (°C)]
 - TropI [Troponin-I (μg/L)]
 - TropT [Troponin-T (μg/L)]
 - Urine [Urine output (mL)]
 - WBC [White blood cell count (cells/nL)]
 - Weight (kg)*
-

FEATURES

*Lactate dropped due to low number of observations

	0	1
0	Age	0.35907750790894677
1	BUN	0.47609556820477267
2	Creatinine	-0.0382167117486476
3	DiasABP	-0.22460778583943242
4	FiO2	0.24202299549038958
5	GCS	-0.7706742717629923
6	Glucose	0.12911837902169404
7	HCO3	0.05802712245767662
8	HCT	0.06157763587568083
9	HR	0.2559296491437221
10	ICUType	0.2718112802246158
11	K	-0.07915416192485707
12	Lactate	0.10148817081317264
13	MAP	0.005590229759995209
14	MechVent	-0.11935859296081534
15	Mg	-0.134211074828018
16	NIDiasABP	-0.13716468837977272
17	NIMAP	-0.046784657432028726
18	NISysABP	-0.02539144472454623
19	Na	-0.0645591197976834
20	PaCO2	-0.17181209557383095
21	PaO2	-0.12577331985945175
22	Platelets	0.09110751593501927
23	SysABP	-0.2402938900598164
24	Temp	0.04912439024735025
25	Urine	0.07560244197974413
26	WBC	-0.03054197986989252

CONTEST RANKING

SAPS = Simplified Acute Physiology Score

My Score: .518 XGB at threshold = .29

<https://physionet.org/content/challenge-2012/1.0.0/>

Participant	Score
Alistair Johnson, Nic Dunkley, Louis Mayaud, Athanasios Tsanas, Andrew Kramer, Gari Clifford	0.5353
Luca Citi, Riccardo Barbieri	0.5345
Srinivasan Vairavan, Larry Eshelman, Syed Haider, Abigail Flower, Adam Seiver	0.5009
Martin Macas, Michal Huptych, Jakub Kuzilek	0.4928

Sample and random predictors

SAPS-I (in m-code)	0.3125
SAPS-I (in C)	0.3097