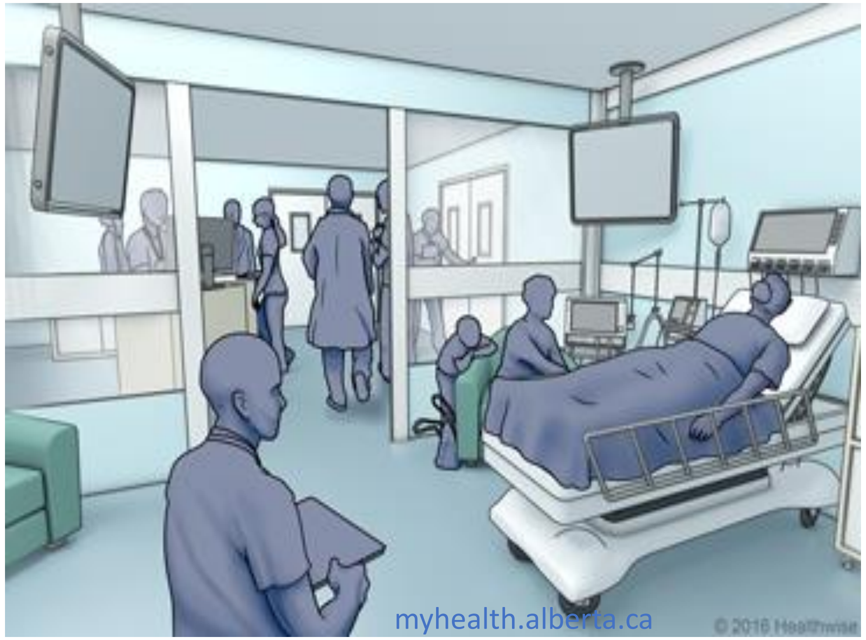
Predicting mortality of ICU patients

Sricharan Bandhakavi
Insight Health Data Science Fellow
Sep 2019

Mortality Prediction of ICU Patients - PhysioNet Computing in Cardiology Challenge 2012*



*https://physionet.org/challenge/2012

QUESTION

Can information collected during the first two days of an ICU stay predict in-hospital mortality?

OUTCOME

Logistic regression model based on ICU data predicts in-hospital mortality with

- 83% AUC
- 76% sensitivity, 74% specificity
- 32% precision

Model performs significantly better than current measures of mortality (SOFA/SAPS-I based models)

Data description

- Predictors (acute physiological measures as time-series data + general descriptors as non-time series data) along with survival outcomes on 4000 patients data

Acute physiological measures on up to 37 features over 48 hours ICU stay such as:

- Respiratory rate
- Heart rate
- Glucose level
- Blood urea nitrogen
- Coma score
- Additional health/physiological measures

\+

General patient level descriptors collected at time of ICU admission such as:

- Record ID
- Age (years)
- Gender (0: female, 1: male)
- Height (cm)
- ICU type (1: Coronoary care, 2: Cardiac surgery recovery, 3: Medical ICU, 4: Surgical ICU)

\+

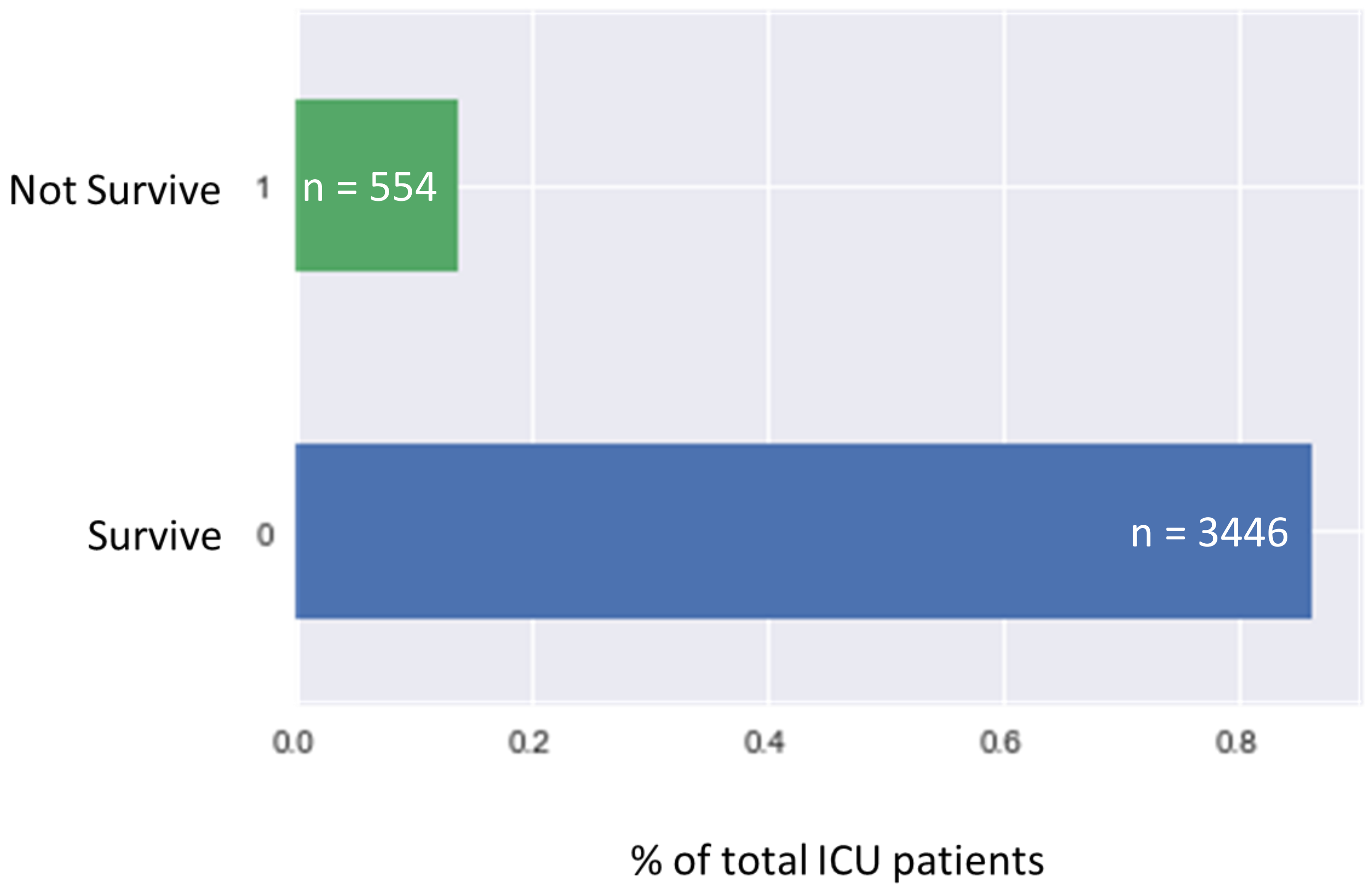
Patient Outcome

Survived or Not

Predictors

Outcome

Distribution of patient outcomes



Imbalanced outcomes in dataset

(addressed during modeling via class_weights “balancing”)

Feature extraction

Non-time-series data

- One-hot encoding of categorical features (Gender, ICU-type)
- Address missing values (eg: for gender created a new category of “missing”)
- Cap “outliers”

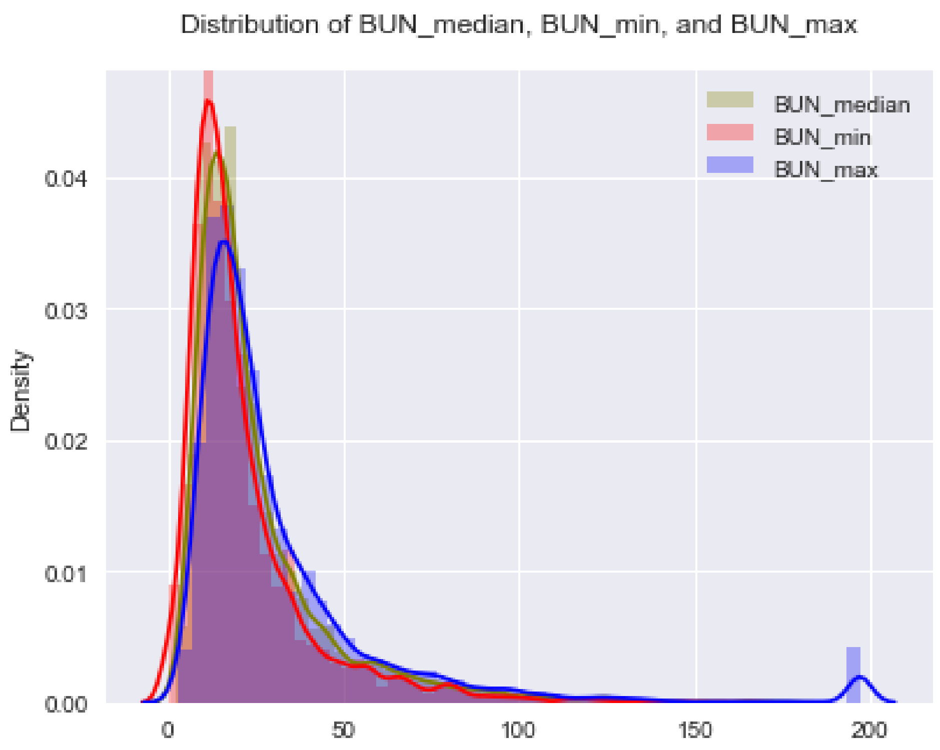
Time-series data

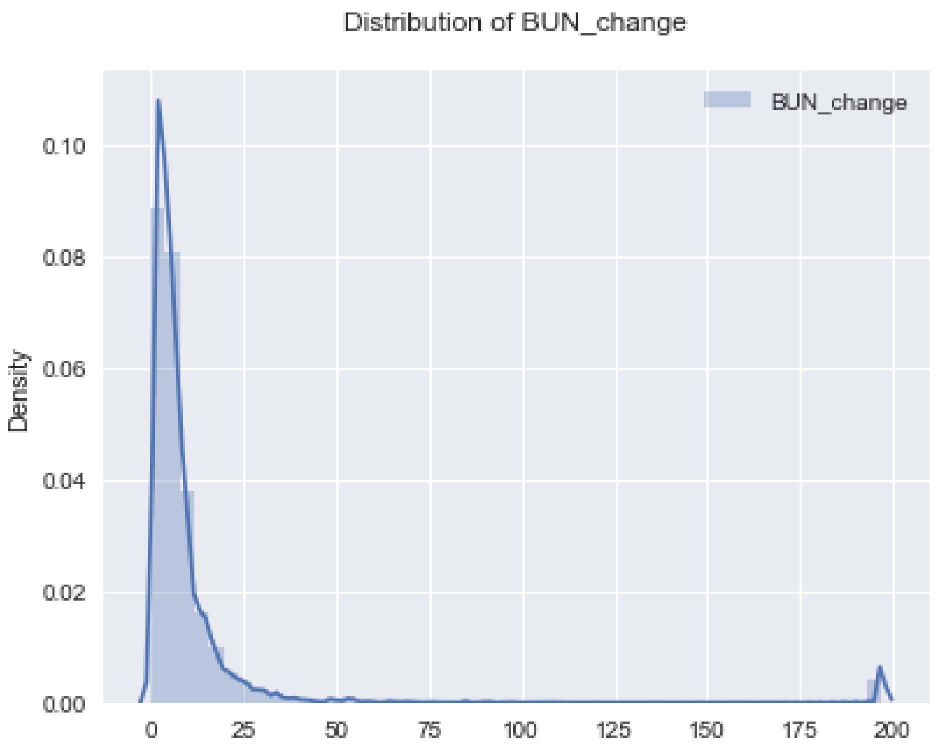
- For numeric features, extract minimum, maximum, median value, change (max-min)
- For categorical features, generate counts/patient
- Address missing values and cap “outliers”

Predictor Matrix

9 features + 145 features

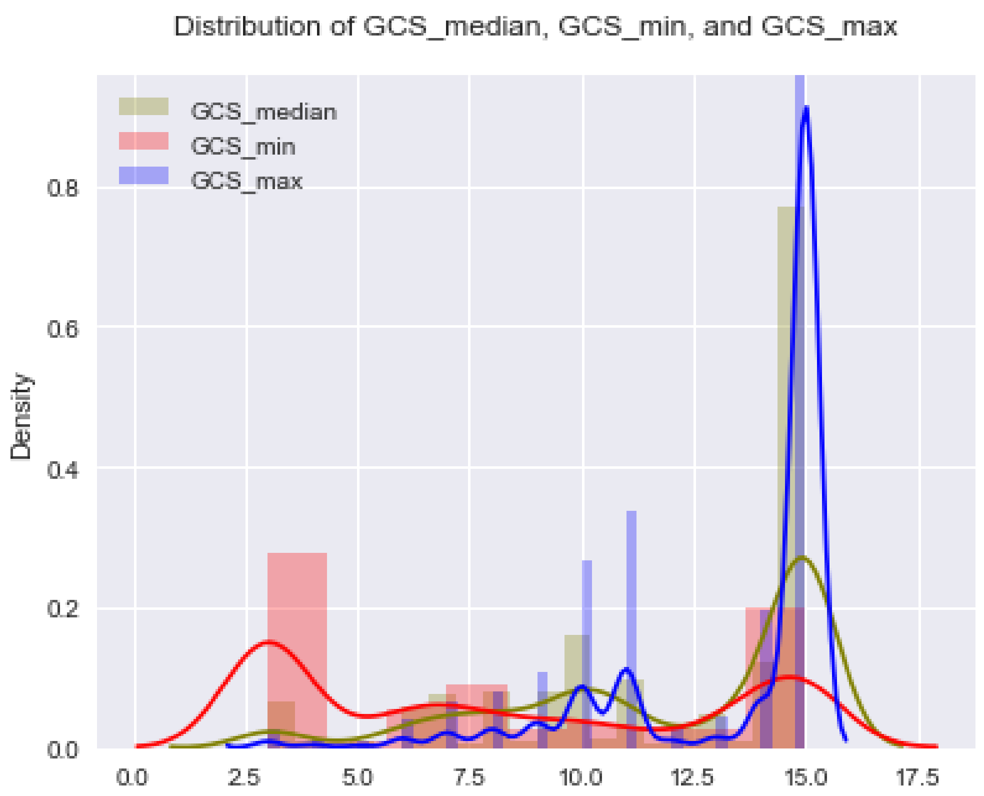
Example feature distributions/exploratory analysis - Patient distributions of Blood Urea Nitrogen extracted features

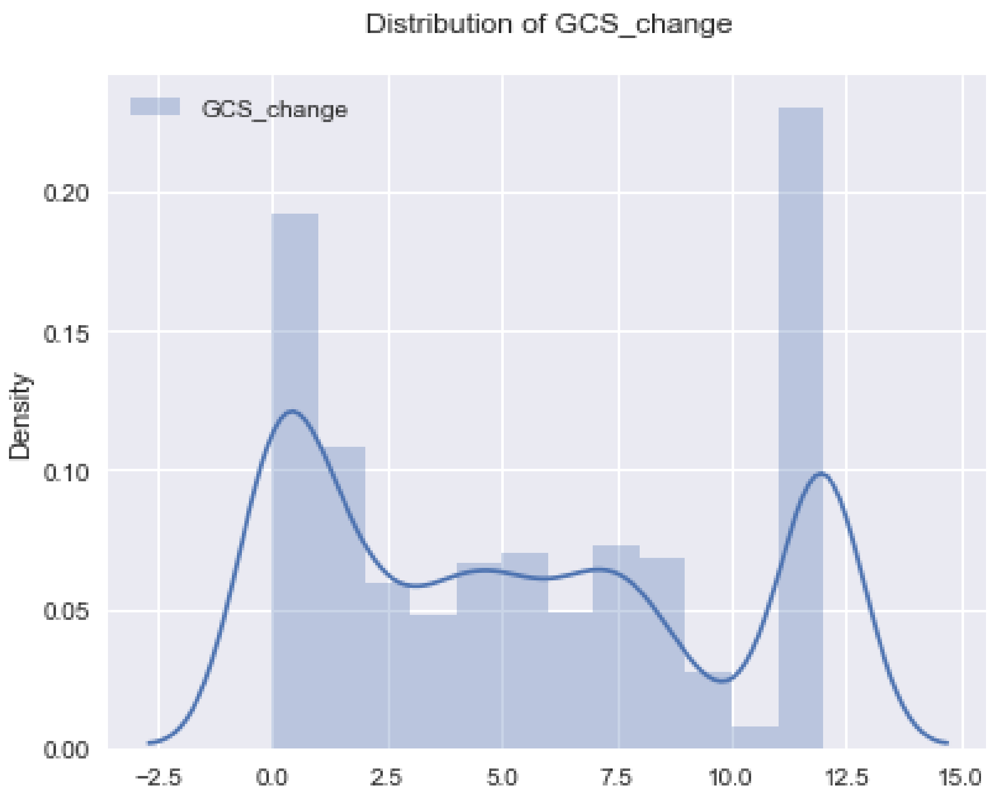




- Nearly overlapping distributions of BUN across median, min, and max.
- Small % of patients show large BUN_change

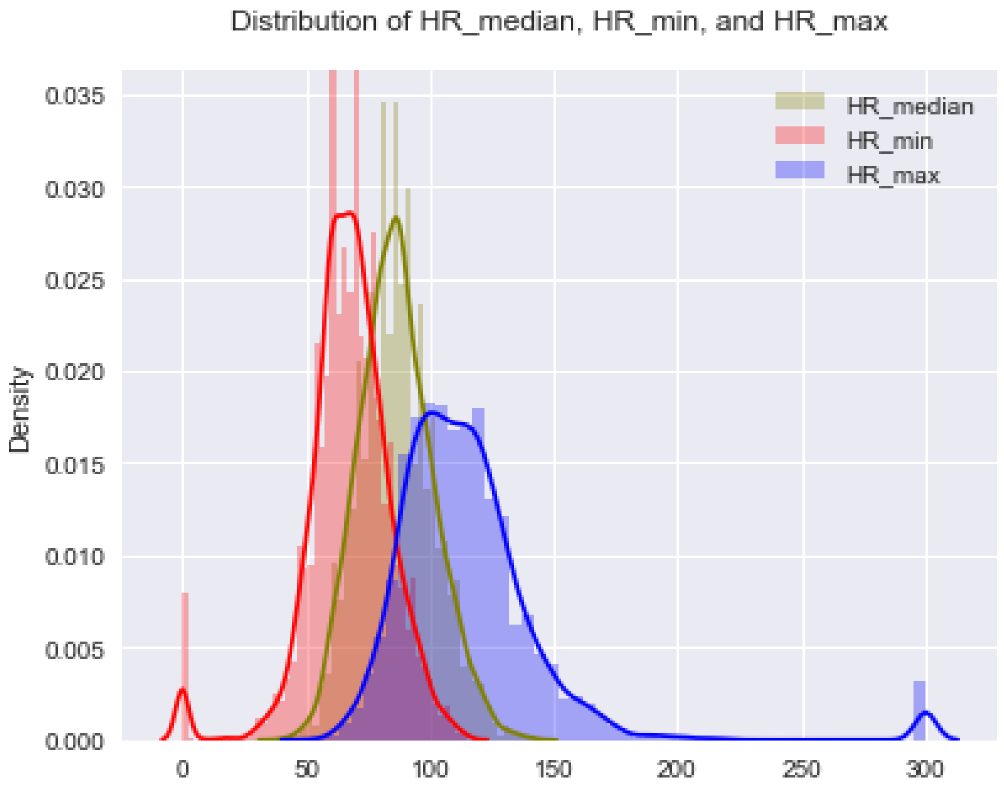
Example feature distributions/exploratory analysis - Distribution of extracted Gleason Coma Scores



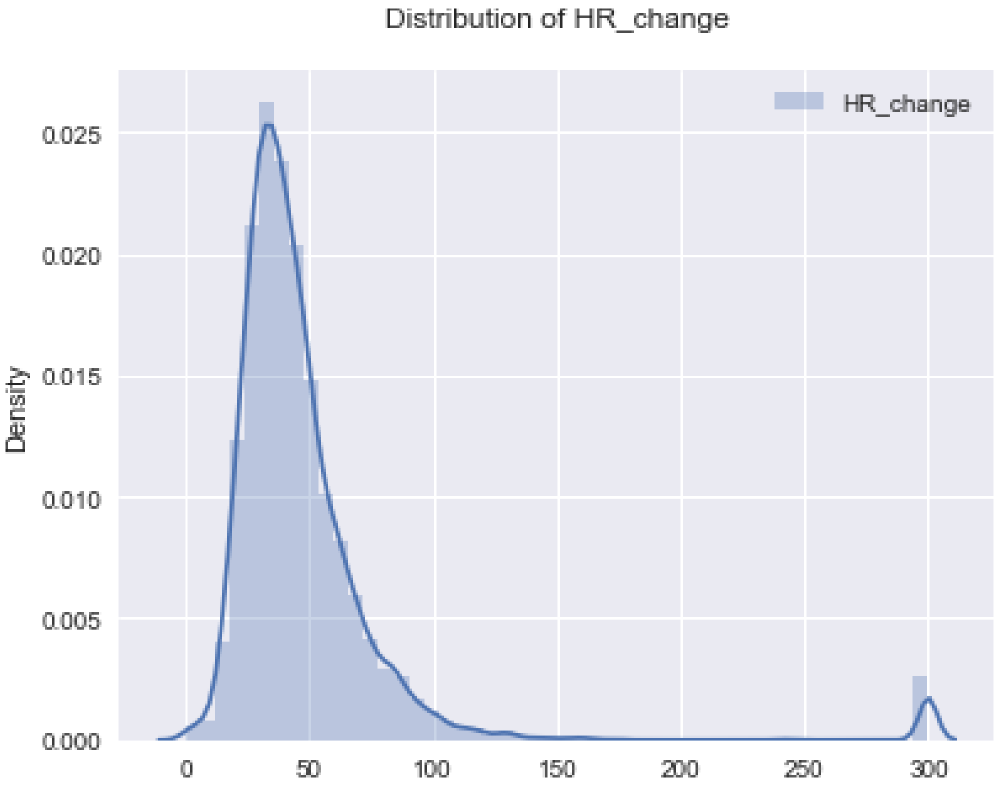


- Relatively different distributions of GCS_min vs GCS_max.

Example feature distributions/exploratory analysis - Distribution of extracted heart rate features (unadjusted)



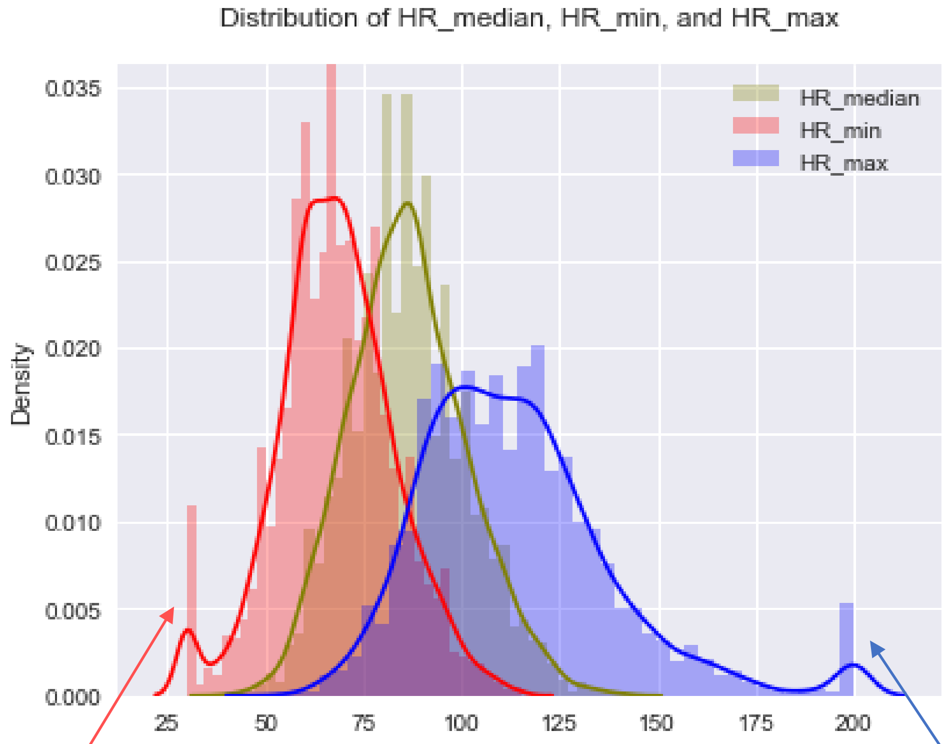
beats per minute



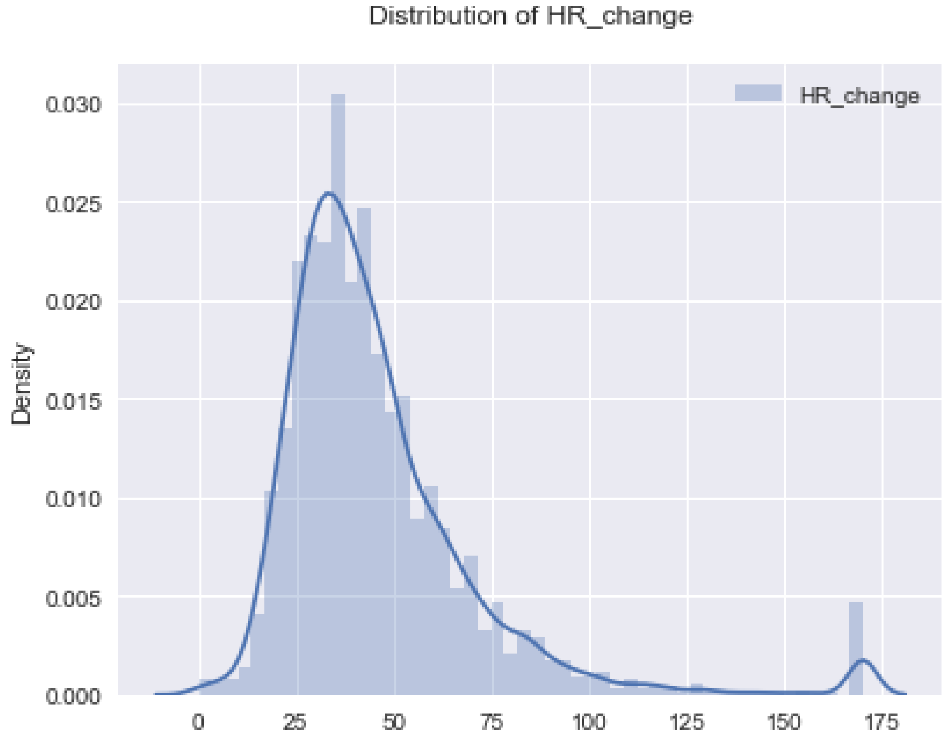
beats per minute

- Relatively different distributions of HR_min vs HR_max.
- Small % of patients show large change in heart rate during ICU stay

Example feature distributions/exploratory analysis - Distribution of extracted heart rate features (with outlier “capping”)



beats per minute

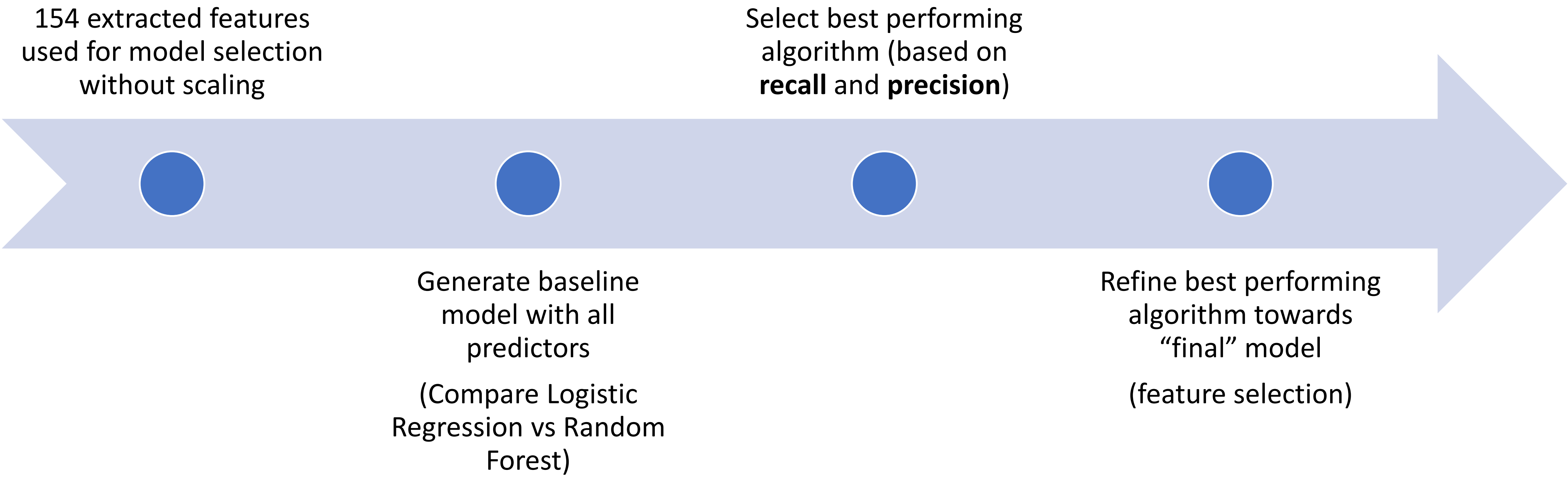


beats per minute

Outliers capped at
<2.5%ile
of distribution

Outliers capped at
>97.5%ile
of distribution

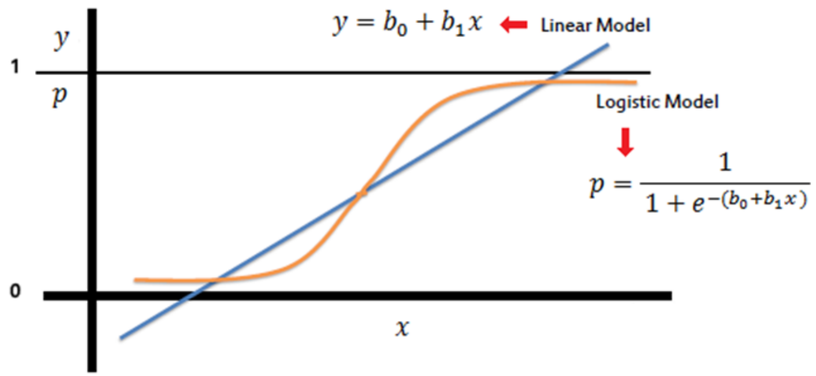
Predictive modeling – high level schema



Logistic Regression vs Random Forest

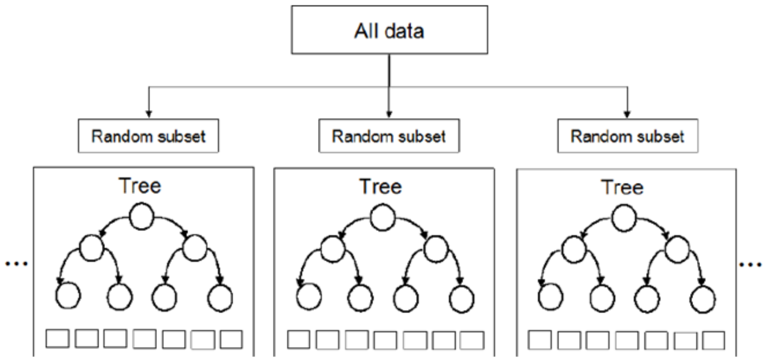
Logistic Regression

- Classification based on the probability of odds given a combination of predictor variables

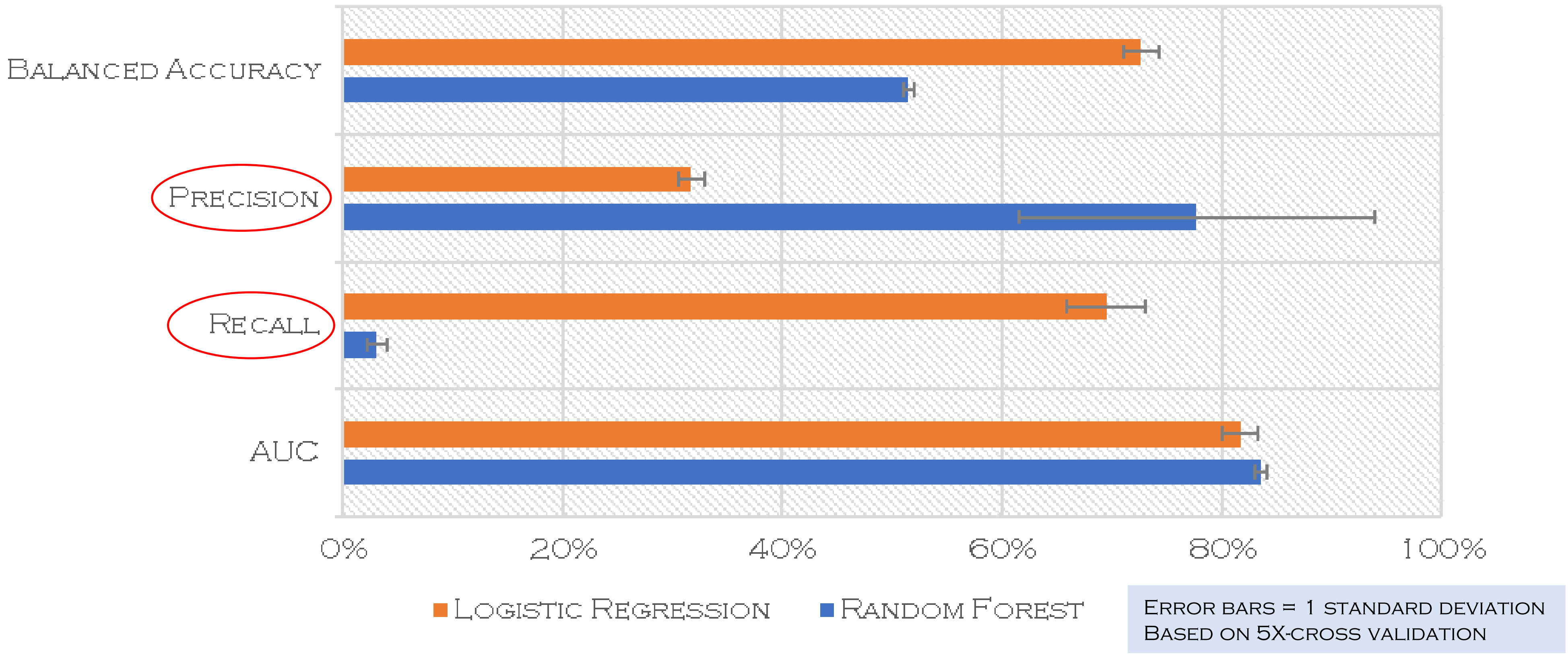


Random Forest

- Classification based on an ensemble of decision trees trained on different subsets of features and observations

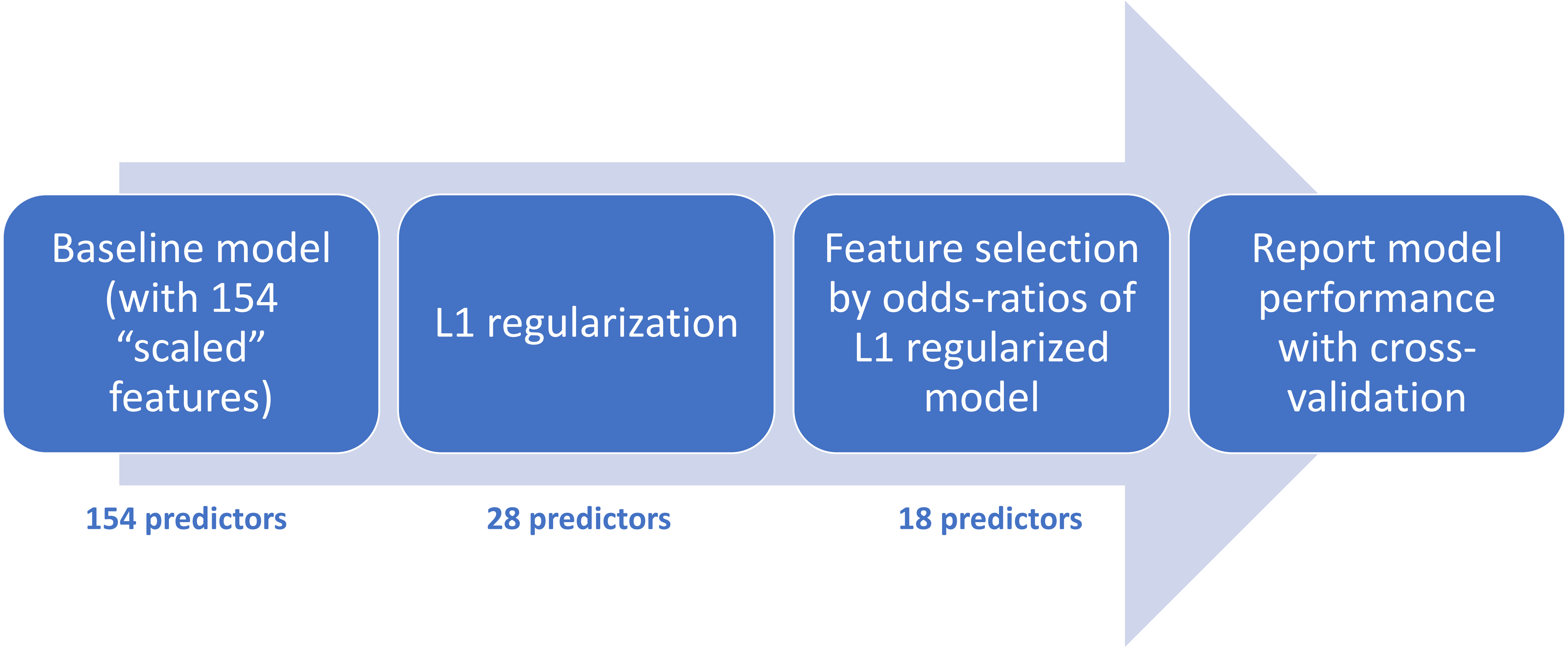


Poor recall with out-of-the-box Random Forest compared to Logistic Regression

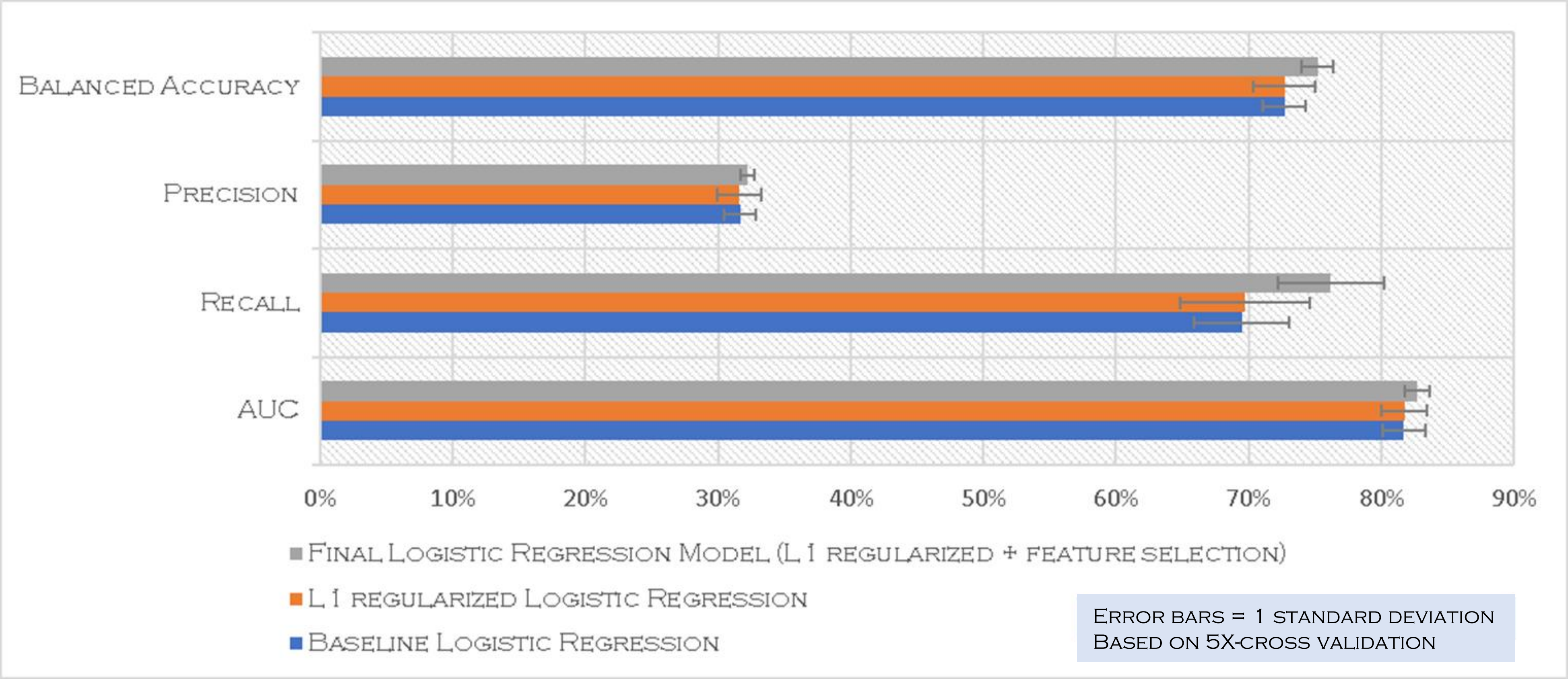


→*Logistic regression algorithm chosen for further refinement towards "final" model*

Logistic regression model refinement schema

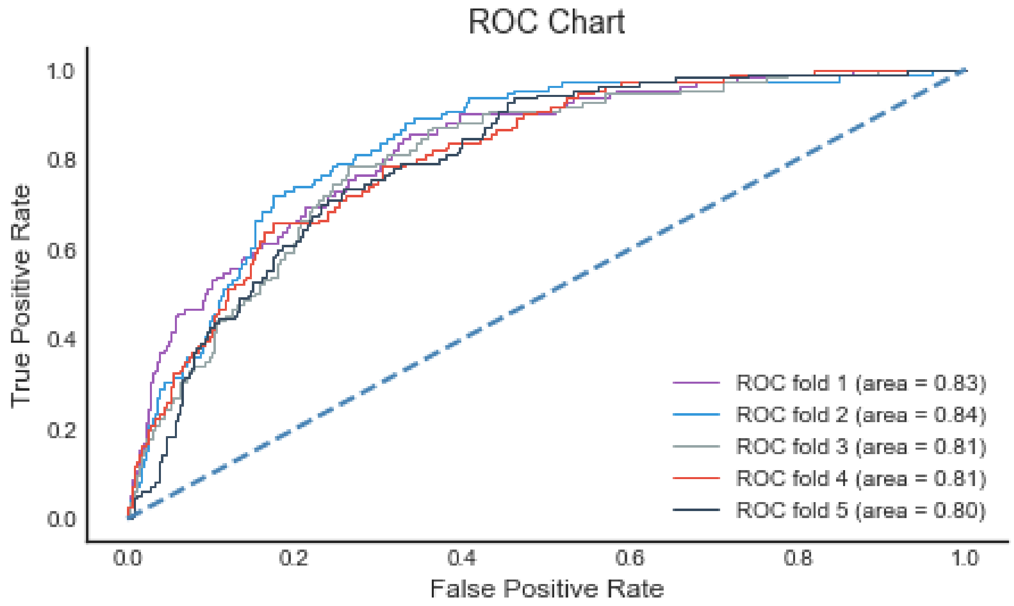


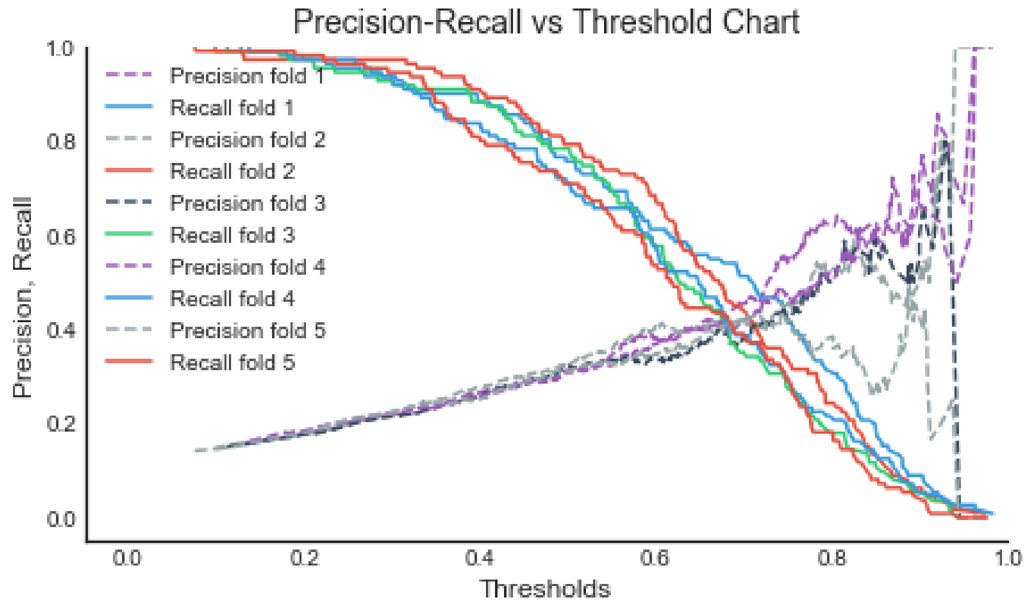
Comparison of performance metrics for all (logistic regression) models



→*Equivalent performance in final logistic regression model (based on 18 predictors/features) vs more complex models*

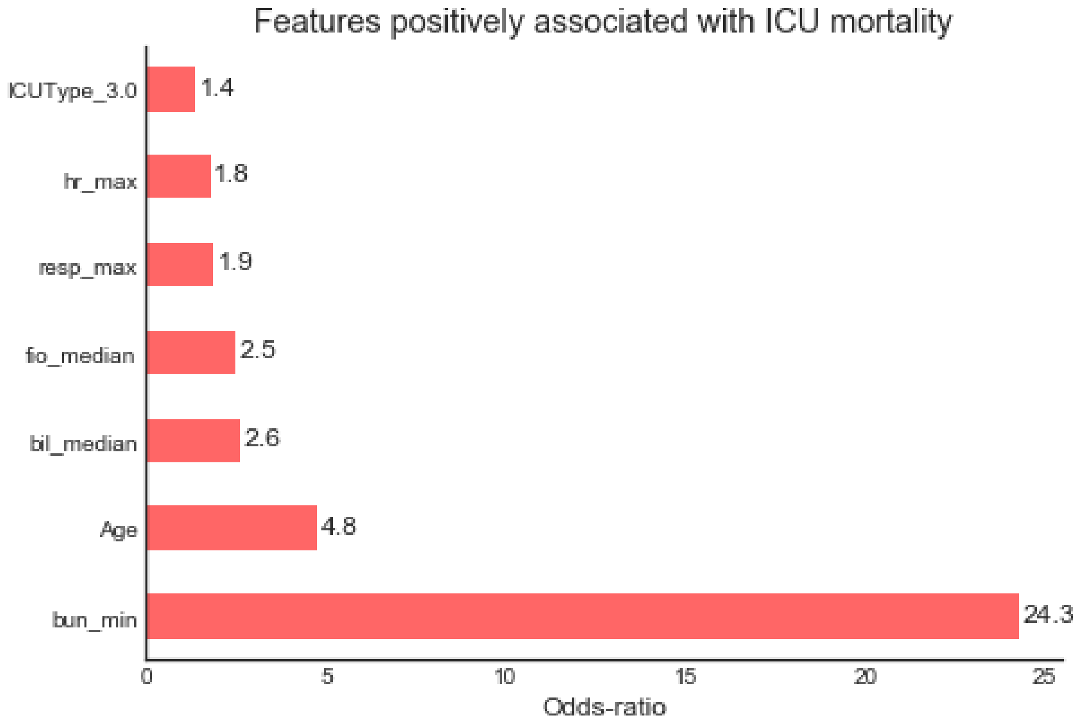
ROC and Precision-Recall vs Threshold Chart

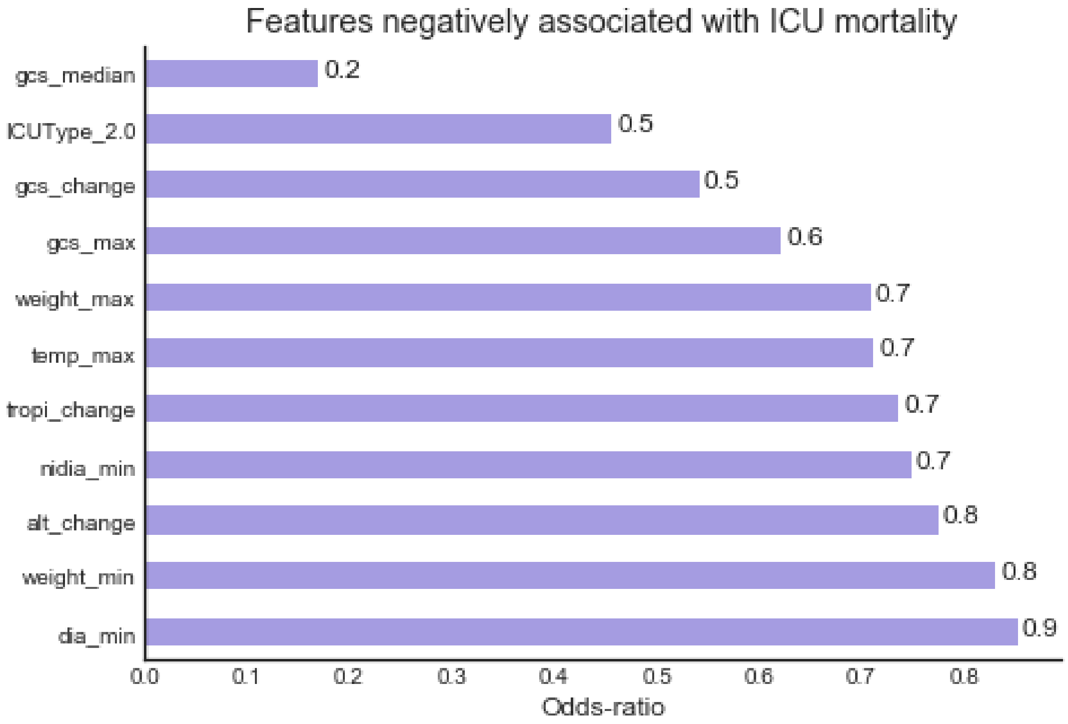




→ *Opportunity exists to improve model sensitivity via reducing probability threshold < 0.5 (at cost of precision)*

“Final” model predictors of in-hospital mortality for ICU patients and odds-ratios*

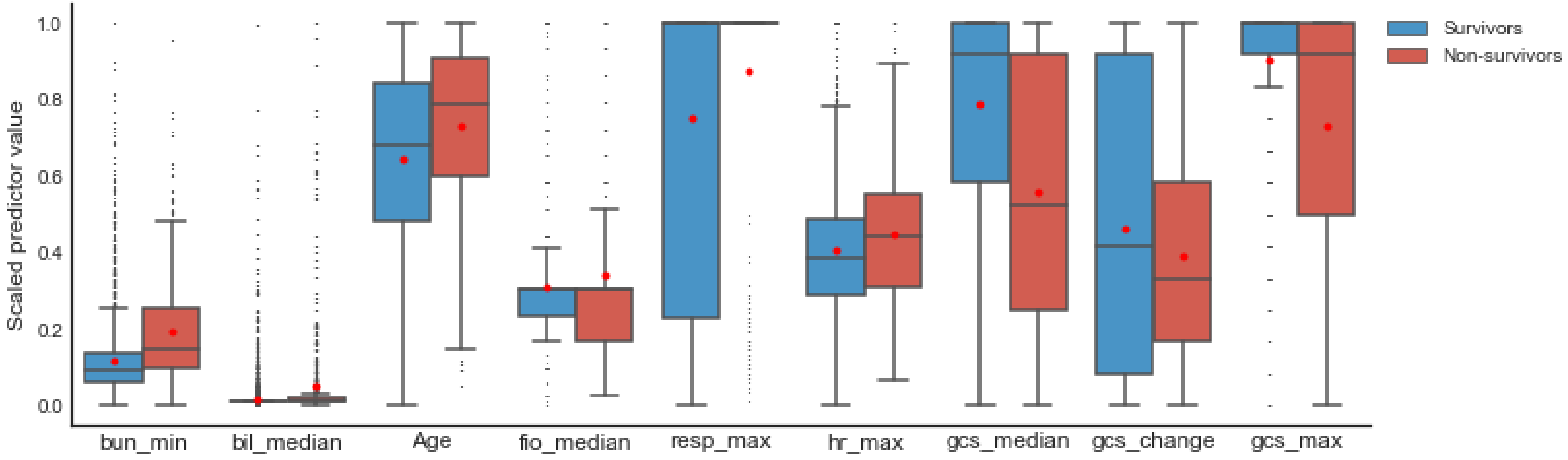




*for predictors scaled between 0 – 1.

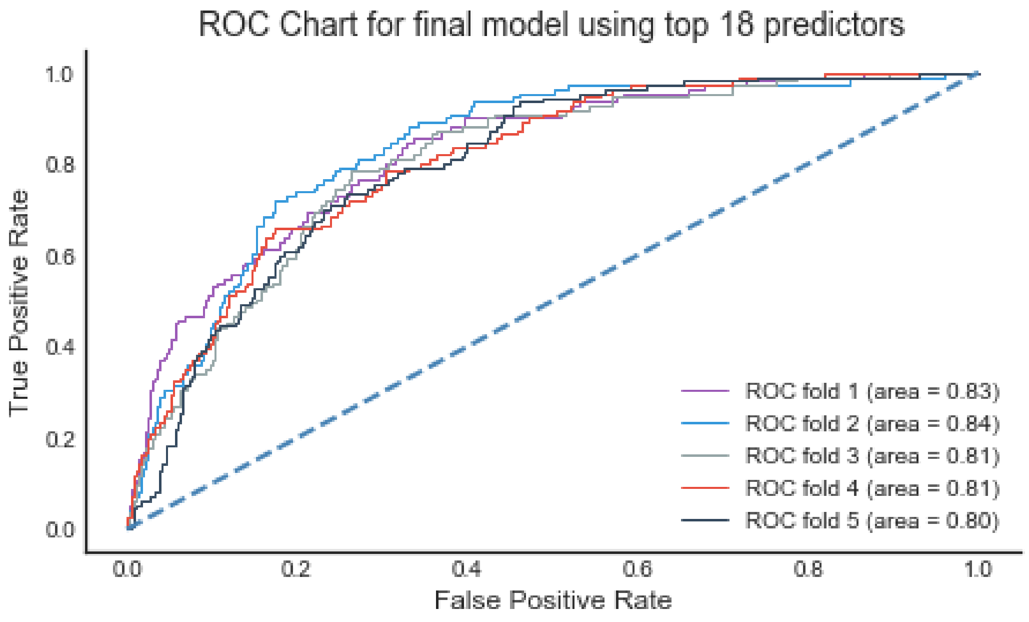
→*Thus, an odds-ratio of 24.3 for BUN_min (Blood Urea Nitrogen minimum) indicates increase in odds of on-hospital mortality as BUN_min increases from lowest to highest value of its measured range.*

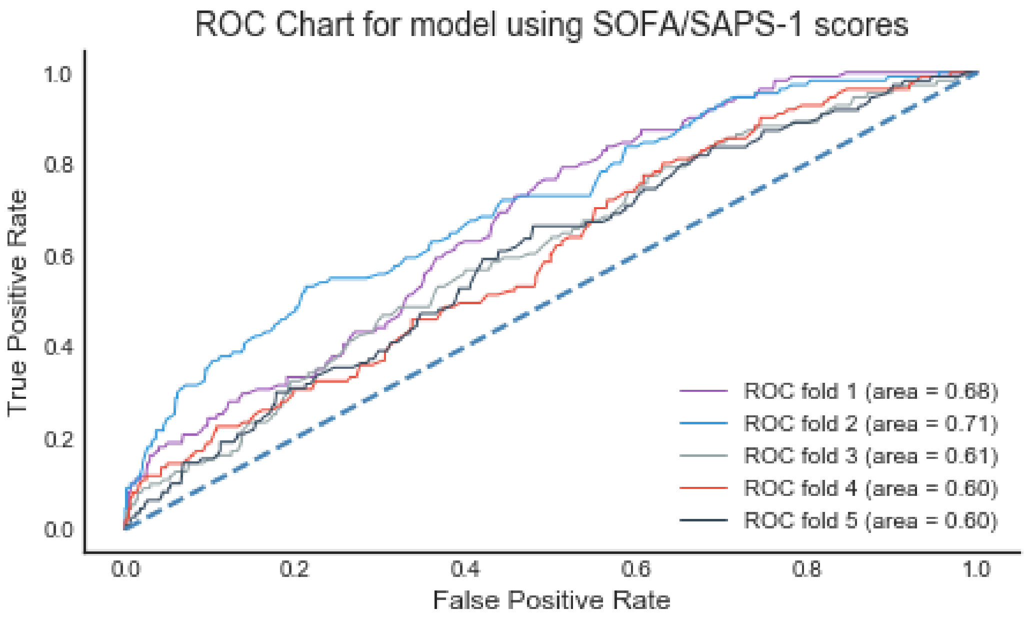
Distributions of top predictors in survivors vs non-survivors



→Positive predictors of mortality have generally higher values in non-survivors (bun_min, bil_median, age, fio_median) and negative predictors have mortality have generally lower values in non-survivors.

Comparison of “final” model to “benchmark” model using SOFA/SAPS-1 scores





→*Across multiple test-train subsets, “final” model generated in this study outperforms models developed using SOFA/SAPS-1 scores*

Summary/Conclusions

A parsimonious logistic regression model based on 18/154 total features was developed to predict in-house mortality of ICU patients. Most important predictors are:

- Blood urea nitrogen levels
- Gleason Coma Score level
- Age
- Bilirubin levels

Model shows superior performance to benchmark models using current measures of mortality. It may be possible to further improve model performance by:

- Addition of more features (eg: past medical history)
- Addition of more data
- Increase of model sensivity (at cost of precision) via lowering of probability thresholds