Explainable Intensive Care Discharge Prediction

University of Bristol

Ryan Porter

Raul Santos-Rodriguez Jeff Clark Alex Hepburn

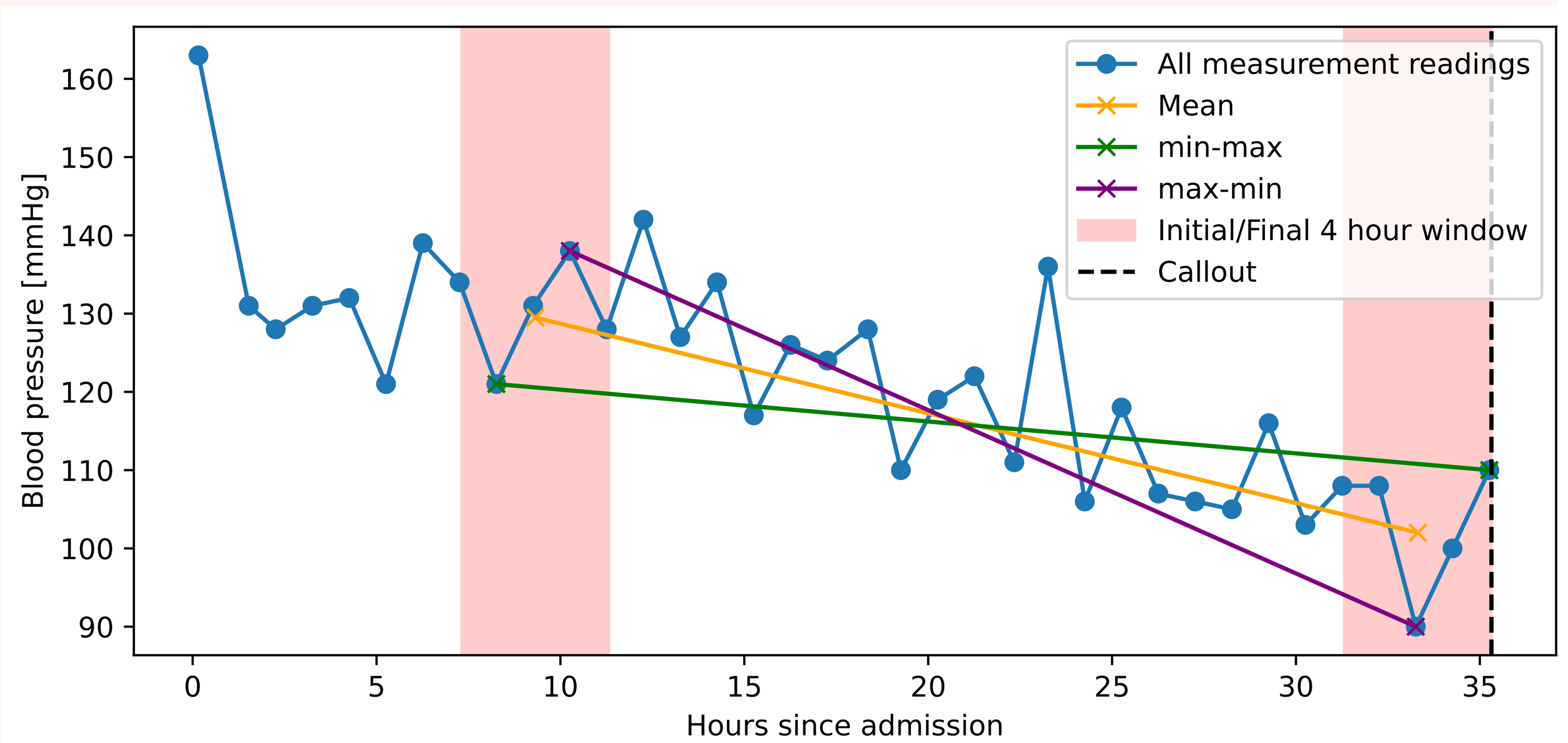
Motivation

- Intensive Care Units (ICU) are commonly a choke point for patient flow in hospitals - The Bristol Royal Infirmary ICU has been near 100% capacity since the COVID-19 pandemic.
- ICU beds are required for many types of patients. Their availability is a determining factor in the scheduling of many urgent surgeries and treatments.
- Currently, patients are called out for discharge by Doctors, upon assessment under extremely time-constrained conditions.
- There is a wealth of data collected on vitals measurements and lab results for patients in ICU.
- This data contains information about the recovery of ICU patients, however Doctors are far too busy to take the majority of it into consideration when making discharge decisions.
- Could machine learning methods digest this data and use it to support the discharge decision making process by making predictions of discharge outcomes?
- Can predictions from ML methods be explained and trusted in the medical context?

Picking up where it was left off

- Physiological measurements from the last 4 hours can be used to predict whether discharging patients at the current time would result in a positive or negative outcome.
- A positive outcome is defined as discharge with no further readmission to ICU, whereas a negative outcome entails either readmission or in-hospital death.
- This is achieved via supervised learning and has been demonstrated on the popular and widely available MIMIC-III dataset.
- Can this method be adapted to predict if a discharge in the future will be successful?
- Could variables that account for the recovery or deterioration of patients over the course of their ICU stay improve the accuracy of these predictions?

Capturing recovery/deterioration

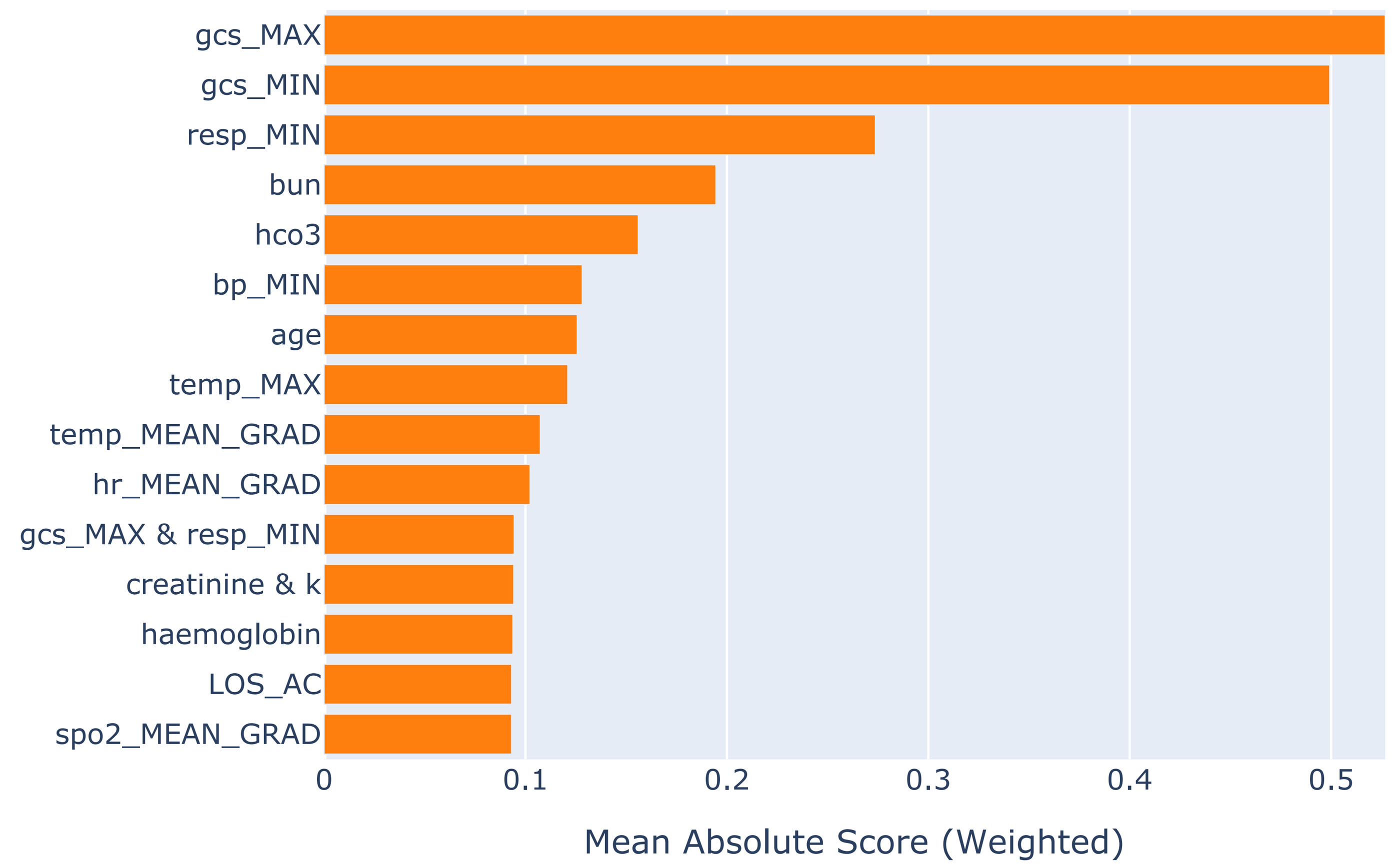


Blood pressure readings from a randomly selected ICU patient. Three gradient features are calculated from the initial* and current data windows; mean, min-max and max-min.

- By calculating gradients for hourly vital measurements, such as blood pressure, their rate of change over the course of the ICU stay can be used to inform predictions.
- Both the magnitude and direction of the gradient may be meaningful in medical context.
- Prior studies suggest the most informative data is in the initial and current time periods.
- Time of day effects - natural variations in physiological measurements due to biological mechanisms (e.g. circadian rhythms). Can appear informative when actually benign.
- To avoid these effects the initial data window is shifted such that it is spaced a multiple of 24 hours from the current data window.
- The addition of gradient-based features to supervised learning models hypothesised to improve performance, while remaining highly interpretable.

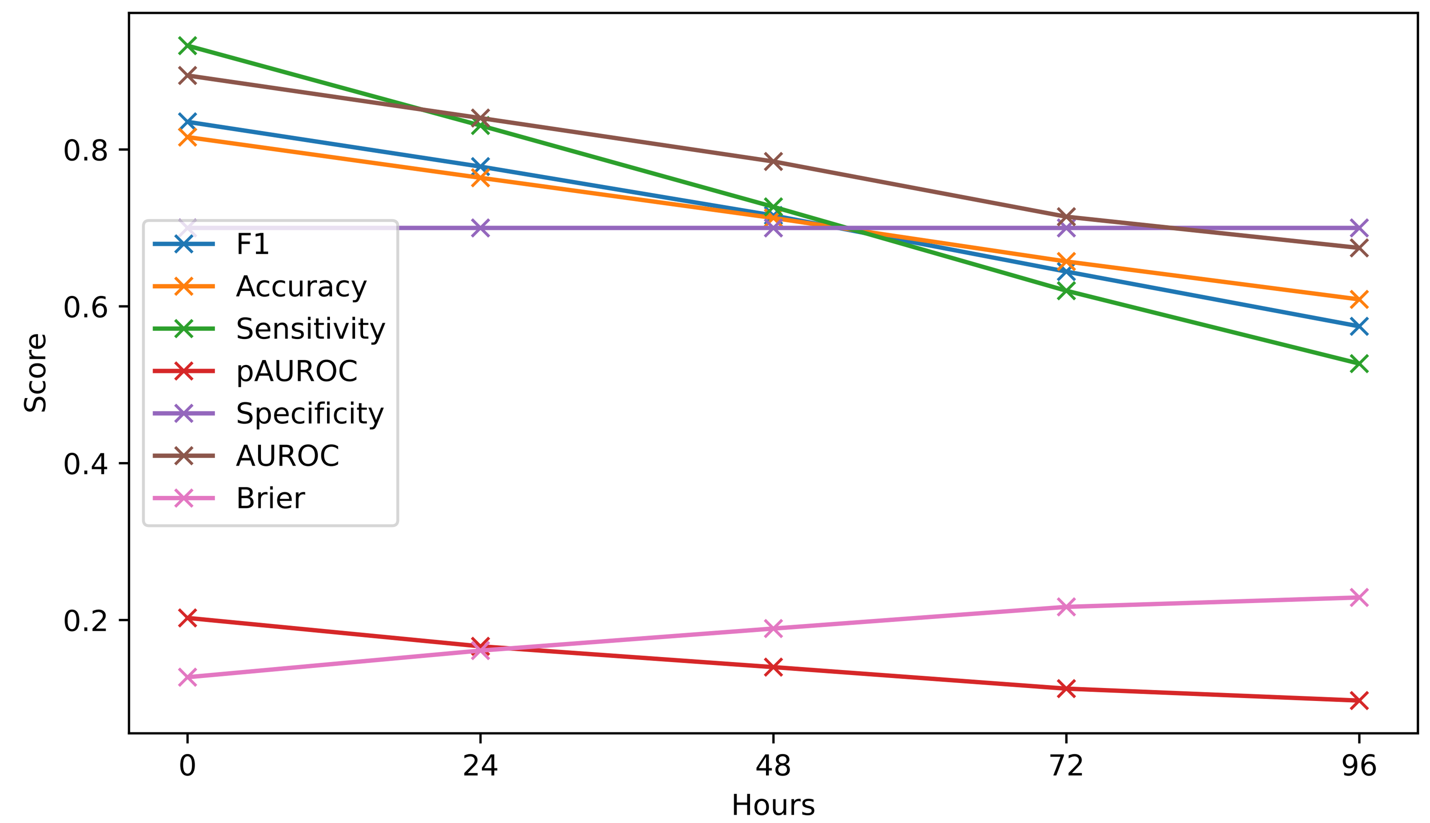
Predictions with Explainable Boosting Machines

Explainable Boosting Machines (EBM) are a highly performing **glass box** supervised machine learning method, meaning that they are highly explainable and editable. This makes them ideal for ICU discharge predictions, since the reasoning behind their decisions can be reviewed with clinical expertise. By training an EBM to classify patients as currently ready for discharge or not with the additional gradient features a model was produced that achieved an F1-score of ~ 0.84 (on MIMIC). This was not an overall improvement on the prior methods. However, some of the gradient features ranked in the top 15 according to feature importance, suggesting that they held predictive power.



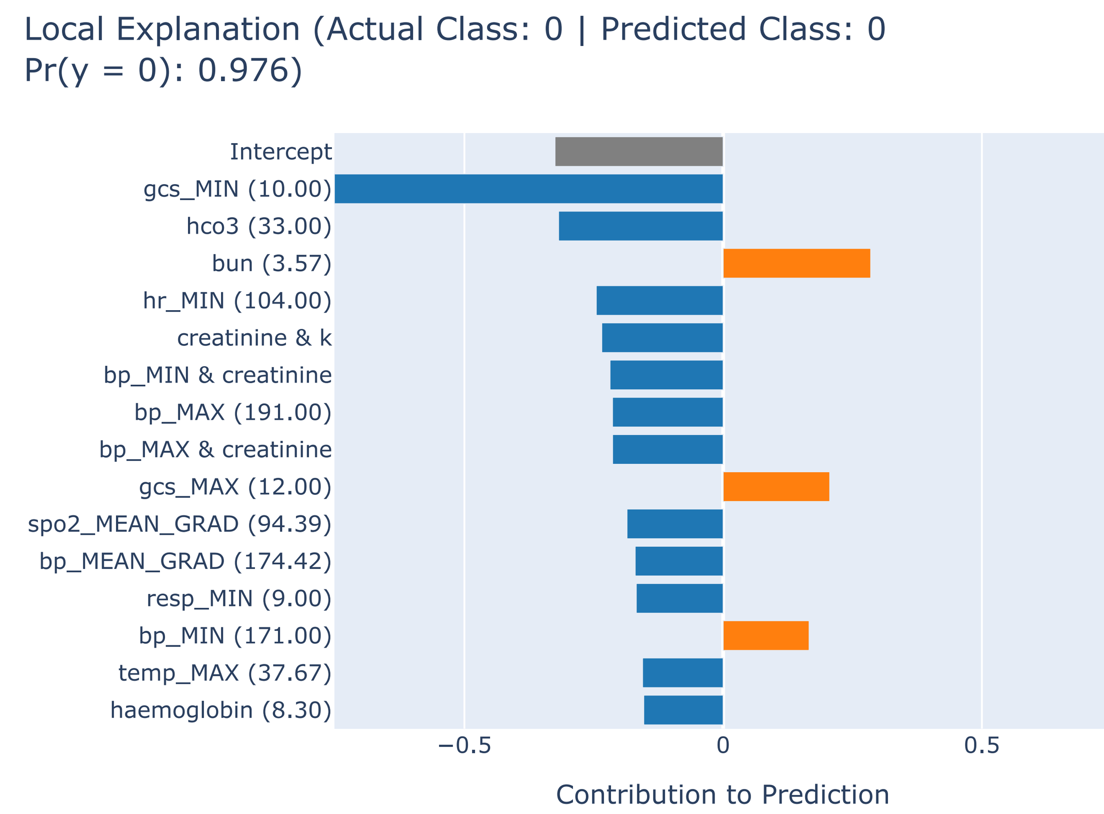
Future predictions

- By resampling patients from earlier stages in their ICU admissions they can be classified as having a positive or negative outcome if they were to be discharged in the future, rather than just at the current time.
- To avoid time of day effects, and to match once-per-day scheduled discharge decisions, separate EBM models are trained to predict outcomes for discharge 24, 48, 72 and 96 hours into the future.
- This proved viable, with diminishing accuracy the further into the future predictions are made.
- The combination of these models acts as a tool that can be used to predict when patients are ready for discharge up to 4 days into the future, with a granularity of 24 hours.
- In theoretical deployment, ICU doctors can query the tool to use the latest collected data to indicate how close patients are to being ready for discharge.
- This could enable better prioritisation of patients for assessment when it comes to the discharge decision time each day, potentially improving patient flow.



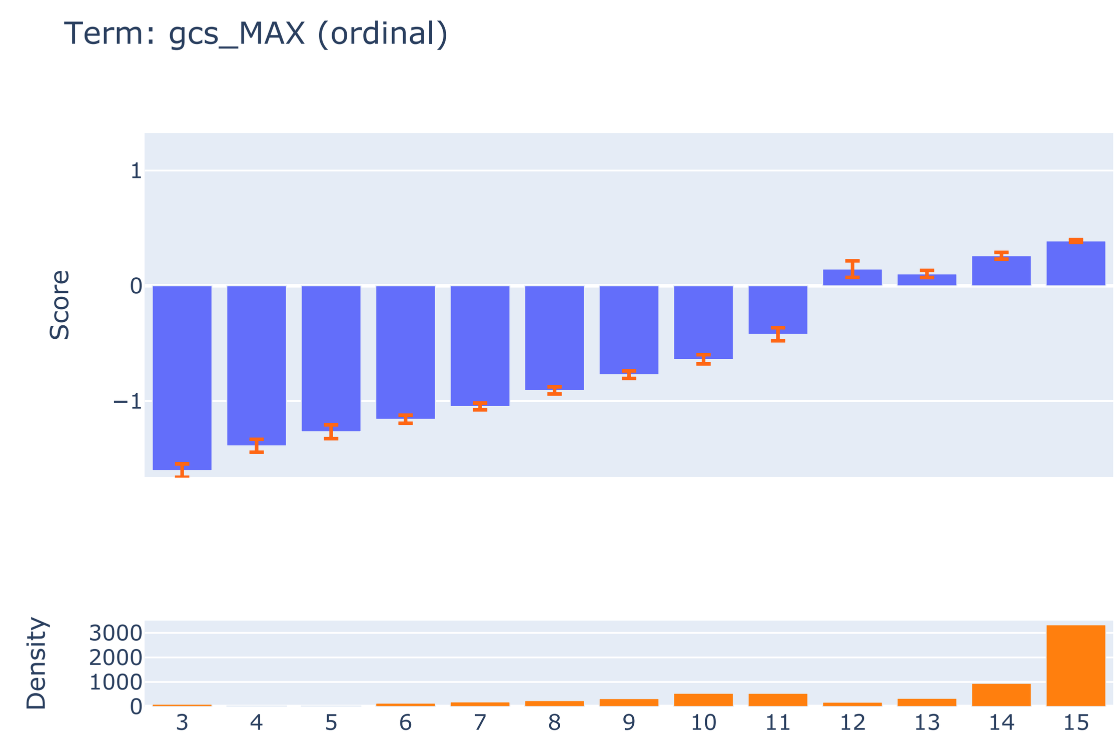
Interpreting predictions

For each individual classification the EBM models give a visualised breakdown of the contributions of each feature to their decision, as well as a confidence score. These are directly interpretable, and their reliability can be evaluated in conjunction with clinical expertise.



Editable models

The contribution of each feature to classifications can be visualised and easily edited. In this example the model erroneously learnt that a GCS score of 12 indicates better health condition than a GCS score of 13, perhaps caused by a lack of instances of this value in this particular dataset. This can be simply edited, which directly changes the response to this feature for all further classifications made with the model.



Conclusion

Achievements

- Proof of concept of future predictions of readiness for discharge at 24 hour intervals
- Predictive power captured in features based on temporal gradient of physiological measurements. However, overall performance of ICU discharge prediction methods were not improved by this.
- Development of Highly interpretable and editable model moving towards deployment.

Further Work

- Clinical evaluation of model feature functions to gain model trust as well as to remove erroneous responses that have been learnt due to limitations of the data.
- Multiple data-source validation to assess and improve generalisability between hospitals.

MIMIC: Medical Information Mart for Intensive Care (https://mimic.mit.edu/), GCS: Glasgow Coma Scale, LOS_AC: length of stay from admission to current time, hr: heart rate, bp: blood pressure, temp: temperature, k: potassium, bun: blood urea nitrogen