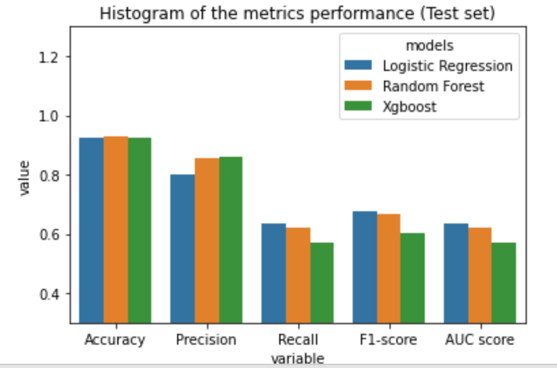
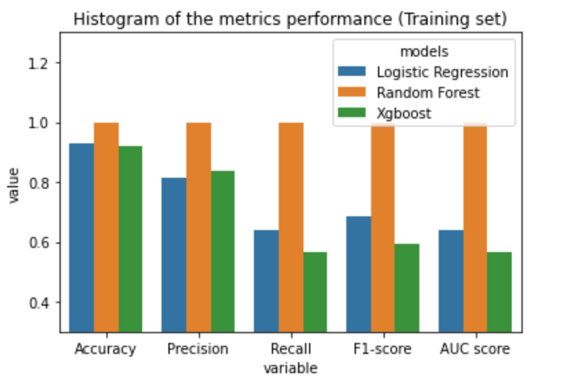
Part 1. Mortality Prediction in the ICU

(f).

By looking at the reports from (c)-(d) and histograms from (e):

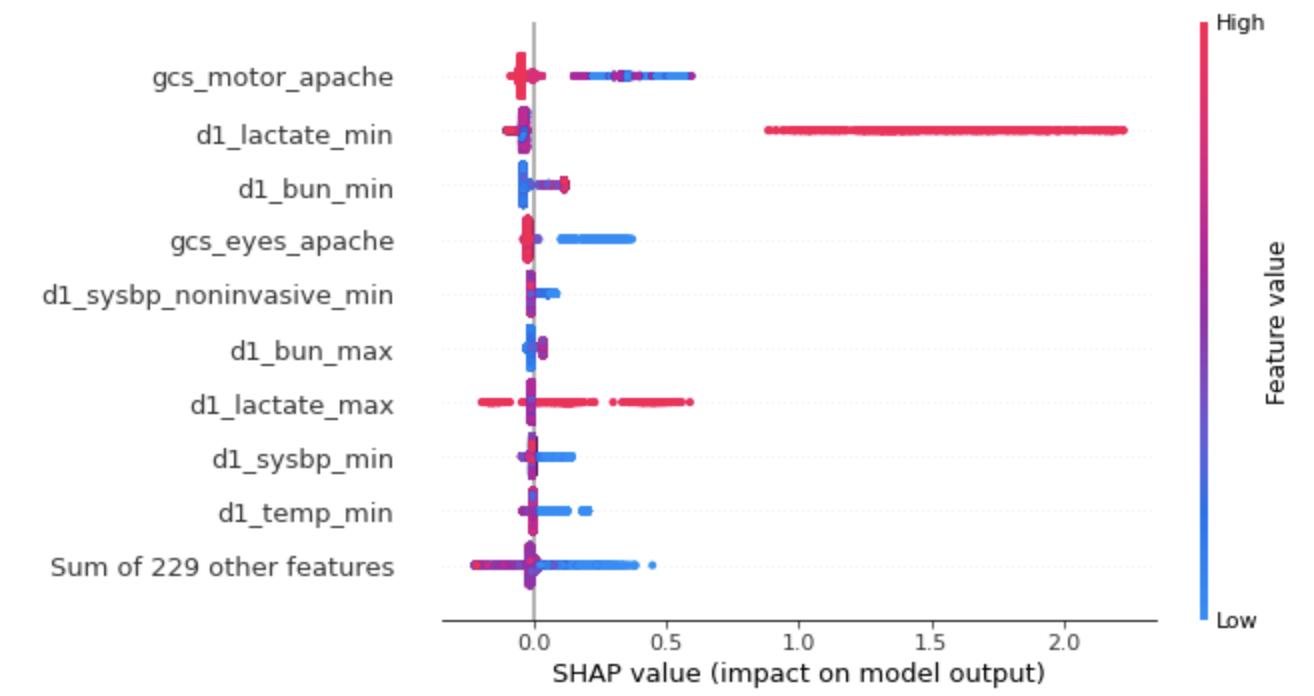


We can conclude that logistic regression model has the best performance among the test set.

For models predicting the mortality rate in the test set, although the logistic regression model has the lowest precision among all 3 models, it have the highest recall, F1-score and AUC score. Thus, we can conclude that logistic regression model has the best performance.

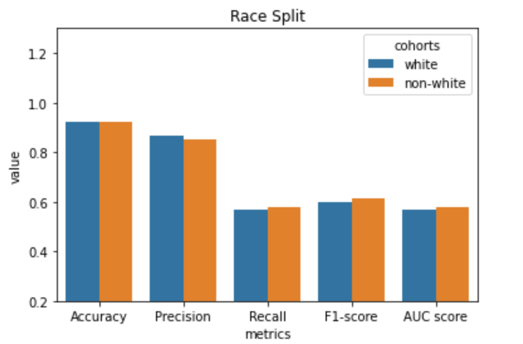
(g).

To get a high-level overview of which features contribute the most to our models’ pre- dictions, we made a beeswarm plot:

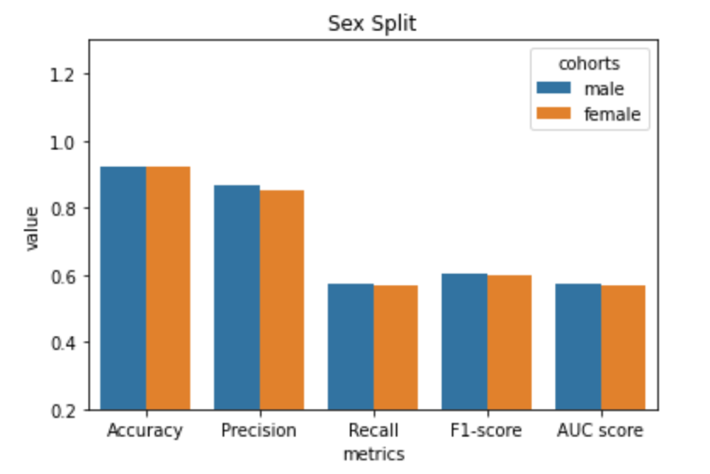


From the beeswarm plot, feature “gcs\_motor\_apache” contributes the most to the model’s predictions. This feature stands for the motor component of the Glasgow Coma Scale measured during the first 24 hours which results in the highest APACHE III score. It is a reasonable feature that the model can rely on: generally, lower GCS scores are correlated with higher risk of death. In this beewarm SHAP plot, a high “gcs\_motor\_apache” (motor component of the Glasgow Coma Scale measured during the first 24 hours which results in the highest APACHE III score) lowers the predicted death probability.

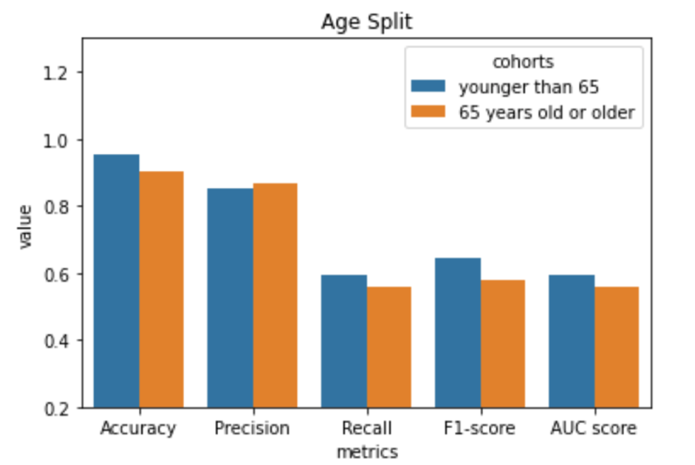
(j).



For the white/non-white split, the xgboost model perform better in the non-white split, for which it has higher Recall, F1-score and AUC score.



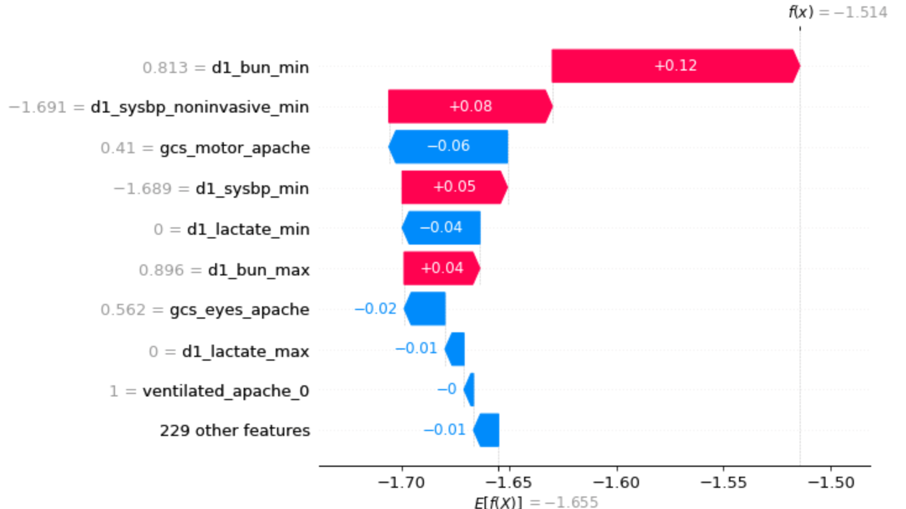
For the age split, the xgboost model perform better in the younger than 65 split, for which it has higher Accuracy, Recall, F1-score and AUC score.



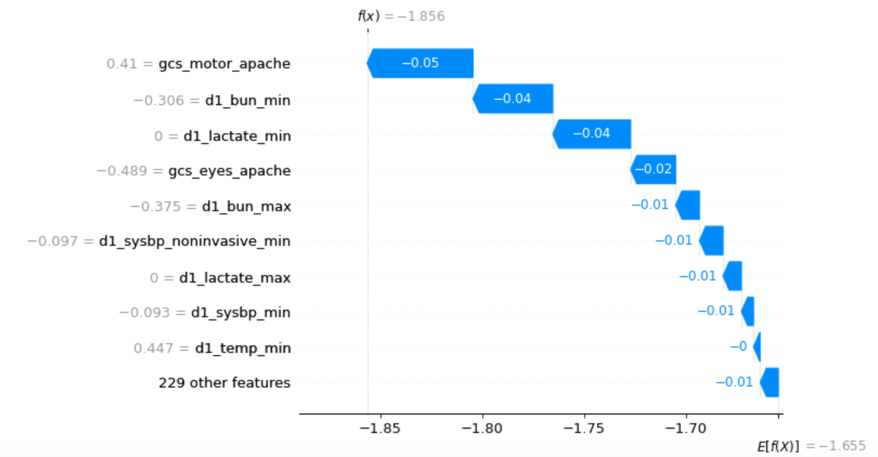
(k).

- Race

white patients:

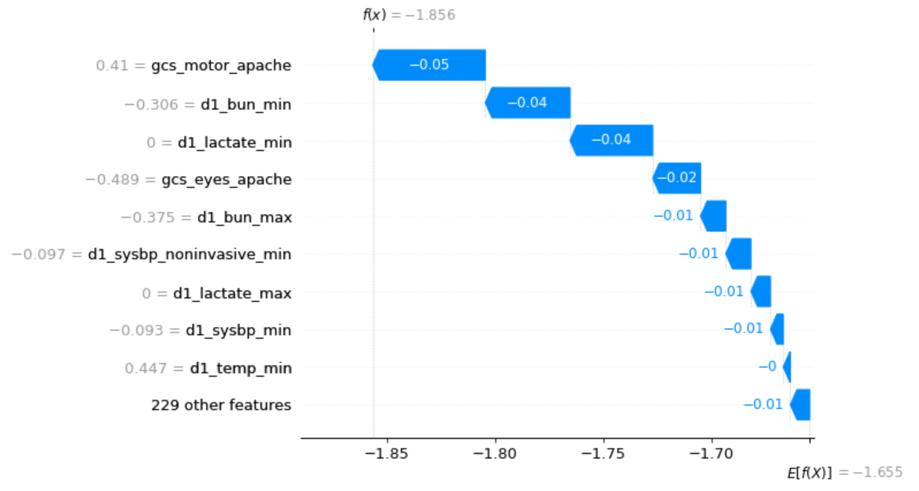


non-white patients:

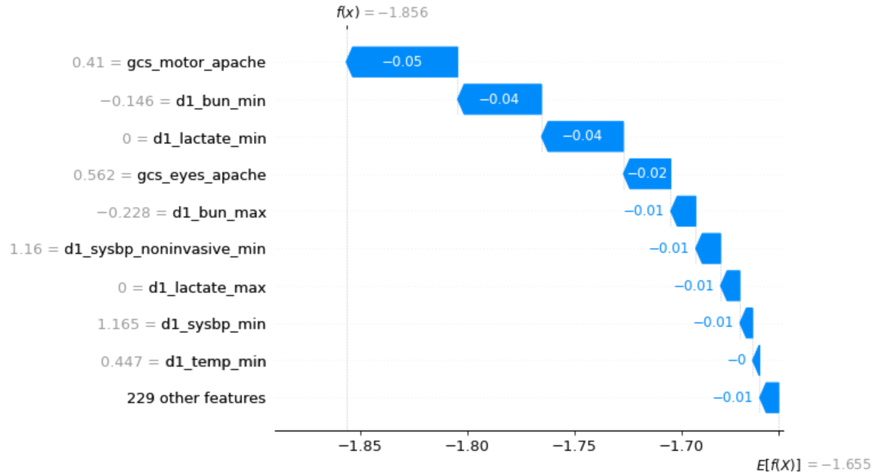


- Sex

male patients:

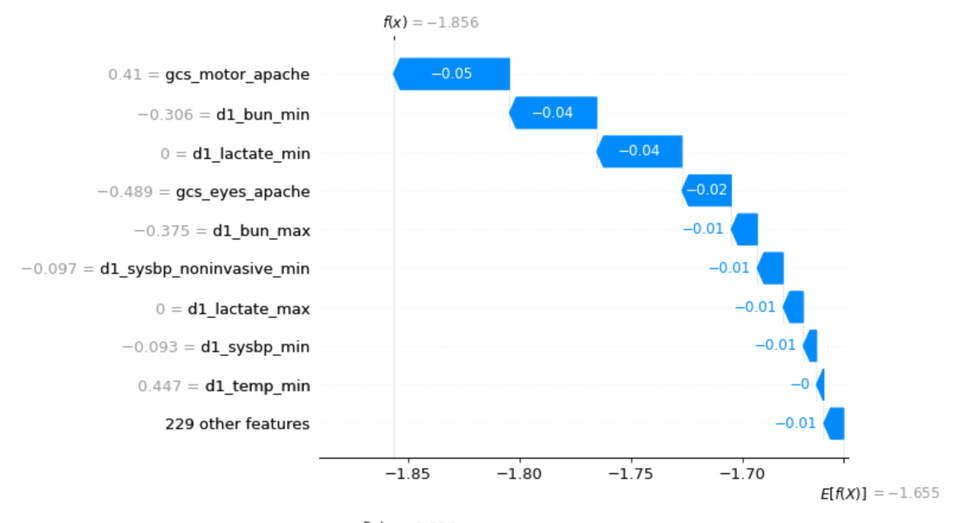


female patients:

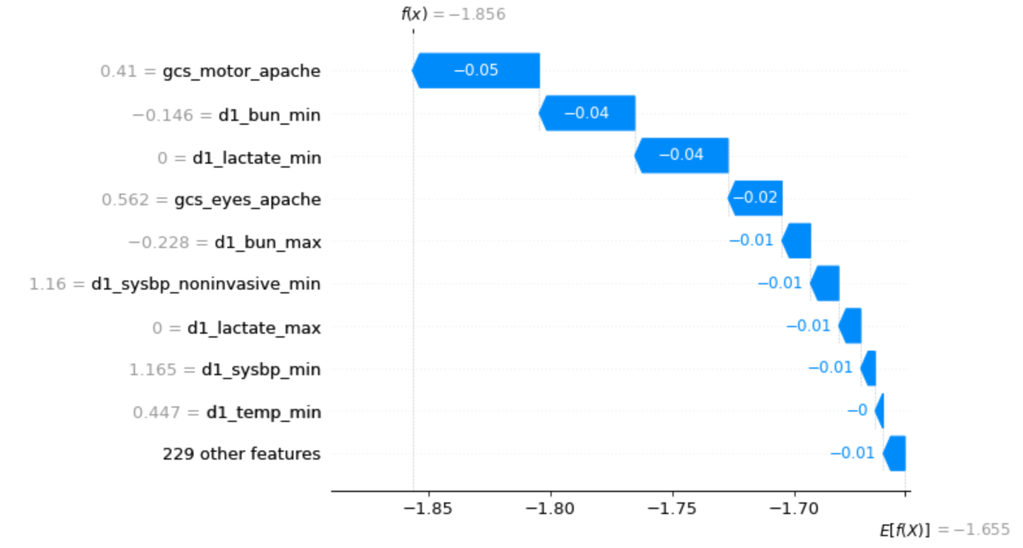


- Age

patients younger than 65:



patients 65 years old or older:



(l).

I’ve noticed discrepancies in the features used by the xgboost model to make predictions for the white/non-white split. Among the white patients,

If not, briefly explain why you think such discrepancies were not observed.

Part 2. Delving into Disparities

(a).

Seeing from the 6 pie plots, we do not see too much differences between the train and test cohorts. For the ethnicity, 78% of the patients are Caucasian, which is a huge imbalance. For the gender, 54% among the patients are male while 46% among the patients are female. For the age, younger patients are slighter than the elder, although it make sense for the normal age distribution in the population.

()