1. Data pre-processing

把outlier(大於Q3+1.5\*IQR或 小於Q1-1.5\*IQR)設為mean，把missing data設為median的preprocessing處理。

2. Details of model setup

The model I use is random forest. And random\_state = 42.

3. Tables or plots for performance evaluation metrics

一張含有 文字, 字型, 螢幕擷取畫面, 白色 的圖片

自動產生的描述

AUROC在4. 中。

4. ROC curve, confusion matrix, calibration plot

ROC curve

一張含有 文字, 圖表, 繪圖, 行 的圖片

自動產生的描述

confusion matrix

一張含有 文字, 字型, 設計 的圖片

自動產生的描述

calibration plot

一張含有 文字, 行, 圖表, 繪圖 的圖片

自動產生的描述

5. Feature importance

一張含有 文字, 螢幕擷取畫面, 字型, 數字 的圖片

自動產生的描述

The figure shows that urineoutput is the most important feature.

6. Fairness analysis (report AUC scores in different race groups)

在hw3.csv中病人的種族有很多種，一張含有 文字, 螢幕擷取畫面, 字型, 文件 的圖片

自動產生的描述，我把他們分為5種，分別是Black、White、Asian、Hispanic、Others，再去看他們的AUC score。

一張含有 文字, 字型, 螢幕擷取畫面, 代數 的圖片

自動產生的描述