Sepsis Prediction in Intensive Care Unit Using Multivariate Clinical Data

Morteza Zabihi[1], Serkan Kiranyaz[2], Moncef Gabbouj[1]

[1] Tampere University, Tampere, Finland
[2] Qatar University, Doha, Qatar

Objective: Sepsis is an organ response to infection and can cause shock, organ failure, and patient mortality if it not diagnosed promptly. According to the WHO fact sheets, it is estimated that sepsis affects more than 30 million people worldwide every year, and potentially is the main cause of 6 million death. As such, the development of a robust and accurate model toward sepsis predication can lead to improving the treatment outcome. The aim of this study is to develop a method to predict the onset of sepsis occurrence using clinical data.

Methods: In total, 102 features are used including the given covariates and handcrafted features. The handcrafted features are extracted from 12 time-varying covariates using a moving window with a length of 5 hours. The covariates used for feature extraction are the heart rate, oxygen saturation, temperature, systolic blood pressure, mean arterial pressure, respiration rate, oxygen saturation from arterial, creatinine, bilirubin direct, lactate, total bilirubin, and platelets. The features are the median, maximum, minimum, variance of the sign changes of the numerical gradients, and the exponential moving average of the 12 measurements. The features are then fed into an ensemble of 10 extreme gradient boosting (XGBoost) classifier and the average of the predictions are used to drive the final decision. In the training phase, the training set is separately balanced using the random undersampling technique for each of the XGBoost to tackle the unbalance and censoring problems.

Results: In the unofficial phase, the proposed method is evaluated over the PhysioNet/CinC 2019 Challenge training dataset. So far, using 10-fold cross-validation scheme, we have achieved an average area under the ROC curves of 0.863 and the utility function (provided by the challenge) of 0.406, respectively. As a future work, we plan to enhance our approach by extracting more informative features. (*Username in Physionet website: morteza.zabihi@gmail.com*)