

Extreme Gradient Boosting Method for Early Detection of Sepsis

Tingting Zhao, Zhuoyang Xu, Zheng Cai

East China University of Science and Technology
Shanghai, China

Objective: The goal of this Challenge is the early detection of sepsis using physiological data. Predict sepsis 6 hours before the clinical prediction of sepsis is the optimal detection and too early or too late predictions of sepsis patients are not encouraged.

Methods: In this article, we use Extreme Gradient Boosting (XGBoost) to automatically identify a patient's risk of sepsis and make a positive or negative prediction of sepsis for every time interval. We think XGBoost is a good choice for this problem that's because: first, XGBoost allow us to define our own optimization objectives and evaluation criteria to rewards classifiers for early predictions of sepsis and penalizes them for late/missed predictions and for predictions of sepsis in non-sepsis patients; second, XGBoost can handle missing values cause there are many missing values in the sepsis data; the last, the boosting capabilities that this algorithm infuses in a predictive model guarantee its performance. Oversampling method were used to solve the unbalanced sepsis and non-sepsis patients.

Results: In our phase 1 of PhysioNet/CinC Challenge 2019, the cross validated score 0.42 was obtained. We plan to control our overall score in the future using more reasonable parameters and better optimization objectives as well as a pre-processing the input data.

Discussion: The ongoing work will attempt to increase the performance of the proposed approaches by considering the correlations between a sequence of measurements over time for per subject and we will extract or generate more significant features.