

Real-time diagnosis of sepsis in intensive care using Logistic regression and Cox proportional hazards model

Fernando Andreotti, Anna Antoniou, Stojan Jovanovic, Rabia Khan, Andras Szabo, Joe Zhu

Sensyne Health plc, Oxford, UK

Aims: Sepsis is a wide-spread severe problem faced in intensive care units (ICUs). Patients with undetected infections develop a complex reaction to fight off pathogens, which often cause organ damage and failure, leading to a life-threatening condition and death in many cases. The difficulty lies in early-detection as the infection does not present obvious symptoms. In this work we aim at assessing machine learning (ML) methods for early prediction of sepsis using electronic patient records (EPRs) provided by the Physionet Computing in Cardiology Challenge 2019.

Methods: In this work we focus on comparing clinically used sepsis scores with the yet less wide-spread ML methods. For this purpose we propose a logistic regression model and the Cox proportional hazards model (Henry et al. 2015 Sci Trans Med). The latter method takes into account the time dependency into the prediction, which enables us to make live monitoring of patient health a reality.

Results: After imputing missing data using forward propagation, we perform a 10-fold cross validation of patient records keeping the same number of septic and non-septic patients. Our initial model performances yield a mean F1 of 0.04, 0.00 and 0.05 for the challenge's sample entry, logistic regression and Cox regression models, respectively on the cross-validation. The sample entry obtained an utility score of 0.21 on the challenge test set, remaining entries are still pending.

Conclusions: During the official phase, we plan on improving imputation of missing values and comparing real-time clinical scores with predictions of the Cox model, logistic regression, and recurrent neural networks. Additionally, we aim at further benchmarking our methods against Physionet's MIMIC-III and eICU data sets.