Automatic Prediction of Sepsis using the Clinical Data

Nadi Sadr, John Anda Du, Philip de Chazal

University of Sydney, Sydney, Australia

In this study, we develop automatic algorithms that identify sepsis conditions. The training and testing data from the PhysioNet/Computing in Cardiology 2019 Challenge was used.

Our initial system processed the heart rate, blood pressure, oxygen saturation from pulse oximetry signal and respiratory rate information. We aim to use all of the available clinical information in the next phase and apply them to a deep learning algorithm to compare the results with our other approach of using machine learning methods. We extracted different features from the time series of clinical data including minimum, maximum and mean value as well standard deviation and other statistical features. We also aim to examine the performance results of expanding the feature set for each epoch using the information from the surrounding time series to boost the classification performance. Features were processed with a linear discriminant classifier. We aim to examine the results using different neural network algorithms including linear discriminant analysis (LDA), logistic regression (LR) and single hidden layer feed-forward neural networks (SHLN).

We applied 10-fold cross-validation on 20,000 data of the training set A and obtained an accuracy of 64.6% for predicting sepsis condition. We require to train our algorithm on the entire training set including training set B and evaluate the performance on the test set.

Future ideas include adding features based on all of the available clinical data as there are markers used clinically to identify sepsis according the commonly used protocols for sepsis diagnosis. Time-frequency, statistical and entropy features will be considered as well as amplitude-based variations in the haematology measurements, respiratory and blood pressure signals. Further, we will apply preprocessing to the time series to reduce the undesired components of the signals. We will apply feature selection using support vector machine (SVM) to extract the most effective information for sepsis diagnosis. Moreover, we will explore the performance of deep learning techniques to compare the performance with the performance results of other machine learning techniques.