Convolutional neural networks based model to provide early prediction of sepsis from Clinical Data

Mohammed H Baydoun[(1)], Lise E Safatly[(2)], Hassan M Ghaziri[(1)], Ali El-Hajj[(2)]

[(1)] Beirut Research and Innovation Center, Beirut, Lebanon
[(2)] American University of Beirut, Beirut, Lebanon

This work deals with the Physionet 2019 challenge which targets the early prediction of sepsis from clinical data obtained when patients are in the ICU. The challenge is deals with a life threatening condition that leads to an increase in mortality rates between 4-8%.

The provided data is comprised of 40 clinical variables obtained during the staying period of the patient in the ICU which usually exceeds 8 hours into days with many of the variables unrecorded. The preliminary approach relies on normalizing each of the variables between followed by creating 40 binary variables to indicate the missing values. This means that the data for each patient can be considered as (number of hours) x 80. This is resized into an image of size 64 x 80 which provides the input to the convolutional neural network (CNN). The output here is a vector of size 64 with a value between 0 and 1 to indicate the sepsis probability. This is resized to the correct number of hours. The network utilizes 4 convolutional layers with each followed by an average pooling layer along the height only. Afterwards, the network has two dense layers and the output. The network is trained to optimize the f-measure. The outputs are then binarized and shifted to optimize the main metric.

The proposed algorithm was not checked on the testing data due to not receiving scores. We provide the results for the validation performed by using the two provided training sets A and B as training and testing respectively in both directions. The scores are provided for some of metrics where the main score for the challenge is the utility score and is currently at 0.58 (0.042). The results are optimized for the utility score noting that other thresholds lead to better scores with other metrics.

Initial validation results

Utilized Score	Score Value (std)
Accuracy	0.93 (0.017)
F-measure	0.29 (0.089)
Utility score	**0.58 (0.042)**