Early Prediction of Sepsis: Using State-of-the-Art Machine Learning Techniques on Vital Sign Inputs

Manmay Nakhashi[1*], Anoop Toffy[1*], Achuth P V[1*], Lingaselvan Palanichamy[1], Vikas C M[1]

[1]Tricog Health India Pvt LTD, Bangalore, India

Abstract

Electronic Health Records (EHR) gives a lot of information regarding a patient's progress in health, who is admitted in an Intensive Care Unit. Sepsis is a critical condition suffered by a patient whom, if not treated in a timely manner can cause casualties. Machine Learning algorithms have evolved to utilize Electronic Health Records to help doctors detect the onset of sepsis. In this work we present Random Forest based ensemble machine learning technique to work on patient's data, also called vital sign inputs, from Intensive Care Unit. The proposed technique performs well on data which contain a major chunk as missing values due to the sparsity of measurement taken in an Intensive Care Unit. We used a combined classifier and early predictor approach to accomplish the task. The classifier do the job of classification when early prediction is not possible due to lack of data. While early predictor predicts the onset of sepsis based on the patient's information it received from previous recordings of vital sign inputs. A utility metric score is used to evaluate the early prediction. The utility function rewards early predictions and penalizes late predictions as well as false alarms. We were able to achieve a good performance based on the so called utility metric.

Key Words: critical care; organ failure; sepsis; machine learning; random forest

1. Introduction

Sepsis is defined as life-threatening organ dysfunction caused by a dysregulated host response to infection [1]. It is a major public health concern, accounting for more than $20 billion (5.2%) of total US hospital costs in 2011 [2]. It is said sepsis is a major concern for the population in the coming decade, that if not treated in a timely manner can affect the mortality rate and critical illness worldwide. A patient who survive sepsis can have serious health disorders and cognitive disabilities leading to healthcare and social implications [3]. The early treatment of sepsis hence play an important role in minimizing the side effects caused by the sepsis condition to a patient. Various definitions are developed over the years for sepsis. From Bone et al [5] in 1991, International Sepsis Definitions Conference developed initial definitions that focused sepsis as host's systemic inflammatory response syndrome (SIRS) to infection. Later scoring methods were introduced for quantifying the organ dysfunction leading to the introduction of Sequential Organ Failure Assessment (SOFA) [6] scores. Higher the SOFA score leads to increased mortality rate for the patient. In 1991, an international consensus panel described instances in which sepsis is complicated by acute organ dysfunction, codified it as 'severe sepsis' or 'septic shock' [7]. A SOFA score ≥ 2 reflects a case of sepsis for the general hospital population suspected with sepsis symptoms. Although another metric qSOFA [1] is introduced to identify sepsis with high sensitivity it never replaced the SOFA score.

Many worked in the area of early detection of sepsis to accurately predict the onset of sepsis from the vital sign inputs taken from patient from the time of admission. Studies conducted on electronic health records using machine learning algorithms to accurately predict the onset of sepsis shows that early detection of sepsis is possible using vital sign inputs [8, 9, 10]. In the work by joseph et at [8] uses patient clinical time series with multi-output Gaussian processes, maintaining uncertainty about the physiological state of a patient while also imputing missing values. The study shows it improved the early clinical detection of sepsis but provided room

* All authors marked, contributed equally for the work

for improvement with more in depth analysis of electronic health records. In another work by Horng et al [9], apart from vital sign and demographic data they use free text data from emergency department to identify the infection. The work improved the ROC curve (AUC) for vitals model. Lauritsen et al [4] proposed an early detection method which uses deep learning models on clinical time series data. It overcomes the shortcomings of machine learning using a deep learning approach on a diverse multicenter data set. The proposed method uses a combination of a convolutional neural network and a long short-term memory network.

The paper is organized into the following section. A short description of the content of the work is given in the abstract. Section 1, gives an introduction of sepsis by its definitions, it also gives different measure to identify the presence of sepsis. Section 2 describes the dataset used for getting the results. Section 3, gives a detailed overview of the methodology used in this paper to early detect sepsis. In Section 4, we cover the results and accompanying discussion. Finally, conclude the paper identifying the possible future directions.

2. Dataset and Preprocessing

We used dataset of 40336 patients Electronic Health Records for building our models. The dataset used is very sparse and contains a lot of missing values. It covers features related to physiology and demographics of the patient admitted to Intensive Care Unit. A detailed list of features and its units is given in Table 1.

Table 1. List of features available in dataset.

Vital Signs	
HR	Heart rate (beats per minute)
O2Sat	Pulse oximetry (%)
Temp	Temperature (Deg C)
SBP	Systolic BP (mm Hg)
MAP	Mean arterial pressure (mm Hg)
DBP	Diastolic BP (mm Hg)
Resp	Respiration rate (breaths per minute)
EtCO2	End tidal carbon dioxide (mm Hg)

Laboratory values	
BaseExcess	Measure of excess bicarbonate (mmol/L)
HCO3	Bicarbonate (mmol/L)
FiO2	Fraction of inspired oxygen (%)
pH	N/A
PaCO2	Partial pressure of carbon dioxide from arterial blood (mm Hg)
SaO2	Oxygen saturation from arterial blood (%)
AST	Aspartate transaminase (IU/L)
Alkalinephos	Alkaline phosphatase (IU/L)
Calcium	(mg/dL)
Chloride	(mmol/L)
Creatinine	(mg/dL)
Bilirubin_direct	Bilirubin direct (mg/dL)
Glucose	Serum glucose (mg/dL)
Lactate	Lactic acid (mg/dL)
Magnesium	(mmol/dL)
Phosphate	(mg/dL)
Potassium	(mmol/L)
Bilirubin_total	Total bilirubin (mg/dL)
TroponinI	Troponin I (ng/mL)
Hct	Hematocrit (%)
Hgb	Hemoglobin (g/dL)
PTT	partial thromboplastin time (seconds)
WBC	Leukocyte count (count*10^3/µL)
Fibrinogen	(mg/dL)
Platelets	(count*10^3/µL)
BUN	Blood urea nitrogen (mg/dL)

Apart from vital signs and laboratory values the dataset also give demographic values such as age, gender, Hospital admission time and ICU length of stay. Since the measurements are done on a need to basis most of the values of the vitals and others were filled with NA's. We used mean values from across the dataset to fill the NA's since their presence can skew the predictions..

3. Methodology

The data was first scanned for outliers since a lot of garbage values were present in a few records. Box plots were used to clean up the outliers and remove the ambiguous records from the data set that can effect the prediction. The thus cleaned dataset is then removed of missing values using mean values across the columns from the overall dataset. Since only 5% of the dataset consists of sepsis cases the data is very sparse.

The method developed uses a two classifier based approach in which one classifier serves do the job of classification when the number of recording of the admitted patient is less that a window size (w). The other classifier is invoked when the w^{th} measurement is done for the patient in ICU. This classifier, so called predictor does the job of early prediction from the given window size, w. The functionality of the system is given in figure 1.

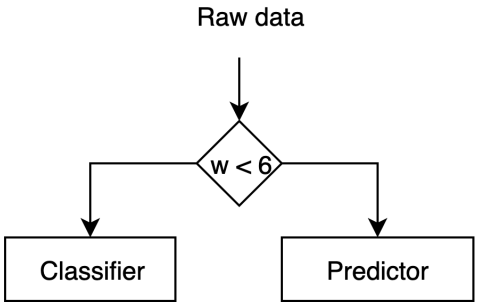

Figure 1: Approach to early detection of sepsis

3.1. Random Forest Classifier and Predictor

We used machine learning ensemble technique random forest for conducting our studies. The random forest is trained on approximately 90% of the available data. The other 10% is used a blind test set for testing the classifier and arriving at a cross validation score to make sure that the classifier generalize the given dataset rather than overfit it.

A window size, (w) of six (6) is used for early prediction of sepsis. Initially we use classification if the window size is less than six, if the results are showing onset of sepsis we use zero filling to create a window of data and verify the same using early predictor as shown in the block diagram in Figure 2.

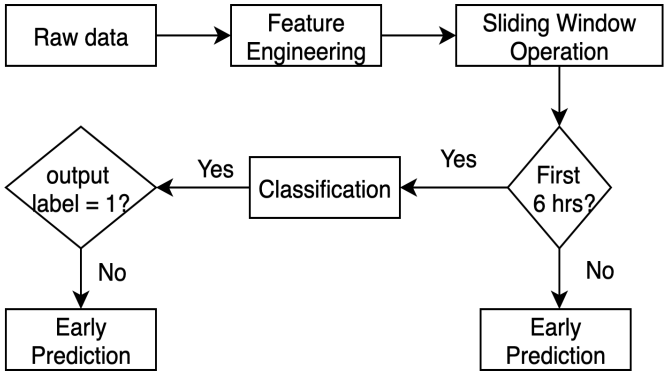

Figure 2: Overall architecture

4. Results and Discussions

The approach yielded decent utility score in a held out dataset and the blind dataset taken from an entirely different hospitals Electronic Health Records.

4.1. Utility function

We first define a score U(s,t) for each prediction, i.e., for each patient, s and each time interval, t:

$$U(s,t) = \begin{cases} U_{\mathrm{TP}}(s,t), & \text{positive prediction at time } t \text{ for sepsis patient } s, \\ U_{\mathrm{FN}}(s,t), & \text{positive prediction at time } t \text{ for non-sepsis patient } s, \\ U_{\mathrm{FP}}(s,t), & \text{negative prediction at time } t \text{ for sepsis patient } s, \\ U_{\mathrm{TN}}(s,t), & \text{negative prediction at time } t \text{ for non-sepsis patient } s. \end{cases}$$

This utility function rewards or penalizes classifiers using their predictions on each patient:

For patients that eventually have sepsis (i.e., all entries of 1) the utitlity function reward classifiers that predict sepsis between 12 hours before and 3 hours after sepsis, where the maximum reward is a parameter (1.0). The function penalize classifiers that do not predict sepsis or predict sepsis more than 12 hours before sepsis, where the maximum penalty for very early detection is a parameter (0.05) and the maximum penalty for late detection is also a parameter (-2.0). For patients that do not eventually have sepsis (i.e., all entries of 0), function penalize classifiers that predict sepsis, where the maximum penalty for false alarms is a parameter (0.05; equal to the very early detection penalty). We neither reward nor penalize those that do not predict sepsis.

The values obtained are normalized so that the optimal classifier (highest possible score) receives a normalized score of 1 and that a completely inactive classifier (no positive predictions) receives a normalized score of 0.

The above utility score is measured on the classifier and predictor system on a held out data set that produced a utility score of 0.1128 with a give AUROC value of 79.35 %. The model is then tested against a completely blind dataset taken from new hospital setup. It yielded a normalized utility score of 0.249.

5. Conclusion

The suggested approach gave a reasonably good performance on the dataset provided. Even though data contains majority of values missing we were able to develop an approach which can be used for early detection of sepsis approximately six hours before its clinical onset.

6. Future directions

The approach used a basic machine learning ensemble technique with some data preprocessing. There are huge scope of improvement since the clinical time series can be modelled using deep learning models, which leverage the availability of data. Further research can be done on developing RNN based classifiers to accurately early predict the onset of sepsis from an approximately fixed window size. Later a study can be done by varying the window sizes to figure out the optimal window size to predict the onset of sepsis.

7. Conflict of interest statement

The authors Manmay Nakhashi, Anoop Toffy, Achuth P V, Lingaselvan Palanichamy and Vikas C M are employed at Tricog Health India Pvt. Ltd.

Acknowledgments

We acknowledge the steering committee for CinC 2019 for allowing us to present our work as a part of the conference. We acknowledge Dr. Charit Bhograj, Dr. Udayan Dasgupta and Dr. Zainul Charbiwala form Tricog for providing us with the necessary guidance and resources for completing this work. Funding: This work was supported by the Tricog Health India Pvt Ltd.

References

[1]Singer, M., Deutschman, C.S., Seymour, C.W., Shankar-Hari, M., Annane, D., Bauer, M., Bellomo, R., Bernard, G.R., Chiche, J.D., Coopersmith, C.M. and Hotchkiss, R.S., 2016. The third international consensus definitions for sepsis and septic shock (Sepsis-3). Jama, 315(8), pp.801-810.

[2] Torio, C. M., & Andrews, R. M. (2006). National inpatient hospital costs: the most expensive conditions by payer, 2011: statistical brief# 160.

[3] Iwashyna, T. J., Ely, E. W., Smith, D. M., & Langa, K. M. (2010). Long-term cognitive impairment and functional disability among survivors of severe sepsis. *Jama*, *304*(16), 1787-1794.

[4] Lauritsen, S.M., Kalør, M.E., Kongsgaard, E.L., Lauritsen, K.M., Jørgensen, M.J., Lange, J. and Thiesson, B., 2019. Early detection of sepsis utilizing deep learning on electronic health record event sequences. arXiv preprint arXiv:1906.02956.

[5] Levy, M. M., M. P. Fink, J. C. Marshall, E. Abraham, D. Angus, D. Cook, J. Cohen, S. M. Opal, J. L. Vincent, and G. Ramsay. "International Sepsis Definitions Conference: SCCM." ESICM/ACCP/ATS/SIS international sepsis definitions conference. Intensive Care Med 2001, 2003 (29): 530–538, 2003.

[6] Vincent, J-L., Rui Moreno, Jukka Takala, Sheila Willatts, Arnaldo De Mendonça, Hajo Bruining, C. K. Reinhart, PeterM Suter, and L. G. Thijs. "The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure." Intensive care medicine 22, no. 7 (1996): 707-710.

[7] Angus, D.C. and Van der Poll, T., 2013. Severe sepsis and septic shock. New England Journal of Medicine, 369(9), pp.840-851.

[8] Futoma, Joseph, Sanjay Hariharan, Mark Sendak, Nathan Brajer, Meredith Clement, Armando Bedoya, Cara O'Brien, and Katherine Heller. "An improved multi-output gaussian process rnn with real-time validation for early sepsis detection." arXiv preprint arXiv:1708.05894 (2017).

[9] Horng, S., Sontag, D. A., Halpern, Y., Jernite, Y., Shapiro, N. I., & Nathanson, L. A. (2017). Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. PloS one, 12(4), e0174708.

[10] Futoma, J., Hariharan, S., & Heller, K. (2017, August). Learning to detect sepsis with a multitask Gaussian process RNN classifier. In Proceedings of the 34th International Conference on Machine Learning-Volume 70 (pp. 1174-1182). JMLR. org.

Address for correspondence:

Tricog Health India Pvt LTD
94/H 9th Cross, RMV Extn,
Bangalore - 560080, Karnataka, India
anoop@tricog.com