A Low Dimensional Algorithm for Detection of Sepsis From Electronic Medical Record Data

Aruna Deogire

A.C. Patil College of Engineering, Mumbai, India

Abstract

Sepsis is an emerging potential health issue, needs to be address on priority basis. Considering the importance, Physionet /Computing in Cardiology announced, the early detection of sepsis as a challenge of the year 2019. Detecting sepsis at earlier stage can save life and cut down financial expenses for the patient. The work presented here is a result of participation in the challenge.

Sepsis is detected by analysing the clinical parameters of the patient. Aim of this work is to study the existing methods of sepsis detection, select the parameters which can optimize the performance and to propose a new method with improved performance. With this aim, existing techniques sepsis-2 and sepsis-3 were studied and a modified approach was proposed to detect sepsis with set of six parameters.

The participation in the challenge was with a team name SHODH. The final utility score on full data set for the work submitted was 0.013 with overall ranking of 68. The work was tested on three data sets, A, B and C and out of which, set C was hidden to the participant. The AUROC (0.586), AUPRC (0.016), Accuracy (0.984) and F measure (0.048) was obtained for the hidden dataset C.

1. Introduction

Sepsis is defined as life threatening organ dysfunction caused by dysregulated host response to infection [1]. In spite of advances in sepsis detection technology and management tools, sepsis is one of the major cause of morbility and mortiality in critically ill patients and one of most expensive healthcare problem in United States [2-4]. Globally, 30 millions of people affected by sepsis every year, causing 6 million deaths annually [5]. Another important aspect of sepsis is, early detection as each hour of delay may increase 4-8% mortality [6, 7]. This will also help to decrease financial burden on sepsis patients.

Considering the importance, the Physionet/ Computing in Cardiology (CinC) had organized a challenge to detect the sepsis at early stage and invited automated solutions to address the problem [8].

The work presented in this paper is a result of participation in the challenge. In this work, a new procedure, with a set of six parameters, redefined weightage and criteria is proposed in this work. The method, is a quick low dimension algorithm for binary sepsis detection.

2. Method

2.1 Database

Sepsis data was reffered from Physionet website, which was made available freely for the challenge 2019 [8]. The data is of ICU patients, from three different hospitals. The first two were provided as training set A & B respectively for testing algorithms and the third was hidden for the participants, used by organizers to score the entries. SET A consists of 20,336 and SET B 20,000 records each. Every record has 40 parameters, out of which first 34 are clinical and remaining demographic. The clinical parameters are vital and laboratory measured values while demographic parameters consist of information like age, gender ICU length of stay information. The parameters in each record are available at hourly basis.

2.2. Proposed Method

2.2.1 Clinical Parameter Selection

A timely detection of sepsis is challenging and complex procedure. The clinical parameters associated with routine measurement and organ functioning plays an important role in the sepsis detection. There are guidelines to assess the impact of the parameters like Sepsis-3. The Sepsis-3 are latest guidelines while Sepsis-2 is an earlier version. Though the guidelines of sepsis identification were updated, number of controversies have been reported regarding utilization of the guidelines and sepsis treatment [9]. The two guidelines mainly differs in clinical parameters used, weightage defined to the selected parameters and threshold criteria applied to differentiates

between sepsis and non-sepsis patients [9, 10]. Though, the evolution has been happened in understanding the phenomenon and revisions are made accordingly, there is still an absence of streamlined sepsis management protocols and standards which can be considered as a golden rule for the disease detection. And therefore a difference in outcome of sepsis detection has been reported with the guidelines. In this work the existing guidelines were studied and a modified approach is proposed to detect sepsis. The method uses the following six parameters as shown in Table 1.

The respiration rate, SBP and MAP are vital and Creatinine, Bilirubin and Platelets are laboratory measured parameters available as per dataset. The vital parameter values are mostly available but laboratory parameters are not necessarily available, for every hour, in the given dataset

Table 1. Parameters used in the Proposed Algorithm

Parameter	Abbreviation
Respiration rate (breaths per minute)	Resp
Systolic BP (mm Hg)	SBP
Mean arterial pressure (mm Hg)	MAP
Creatinine (mg/dL)	Cr
Bilirubin total (mg/dL)	Bi
Platelets (count*10^3/μL)	Pt

2.2.2 Threshold Criteria

Following criteria is applied to the above parameters in the existing algorithm. Units of the parameters are as defined in Table 1.

> ➢ Respiration rate > 23
> ➢ Systolic Blood Pressure >135
> ➢ Mean Atrial Pressure < 67
> ➢ Creatinine >2
> ➢ Bilirubin total >2

> ➢ Platelets >=150 & <20
> > • >= 150 (0) , <150 (1), <100 (2), <50 (3) , <20 (4)

The above criteria threshold was obtained by experimenting with different threshold values and finalising the values which gave optimum response for sepsis detection.

2.2.3 Weights Assignment

To get a quantitative measure sepsis, the parameters are assigned certain weight in any sepsis detection procedures. The Sepsis-3 guidelines uses 0-4 point scale for its Sequential Organ Failure Assessment (SOFA) and accordingly ranges are distributed [1]. In present work the weights are defined on 0 to 1 point scale for all parameters except Platelets, which is asses on 0 - 4 scale because of its wide range variation. Table 2 shows the weights assigned to each of the parameters.

The proposed method assigns weight of 0.5 maximum to the Respiration and Systolic Blood Pressure parameters as they are routine measurements and may not directly linked to organ dysfunction measure. Higher weightage is assigned to organ specific critical parameters. Selection and finalization of the parameters and the ranges was done by experimenting and comparing results of different combination of parameters and ranges.

Table 2. Weights Assigned to Each Parameter.

Parameter	Weight
Respiration Rate (Resp)	0 - 0.5
Systolic BP (SBP)	0 - 0.5
Mean arterial pressure (MAP)	0 – 1
Creatinine (Cr)	0 – 1
Bilirubin total(Bi)	0 – 1
Platelets (Pt)	0 - 4

2.2.4 Sepsis Detection

Figure 1 shows the flow of the sepsis detection algorithm. For a given record data, the parameter values were read. The predefined criteria was applied and weights calculated for every parameter.
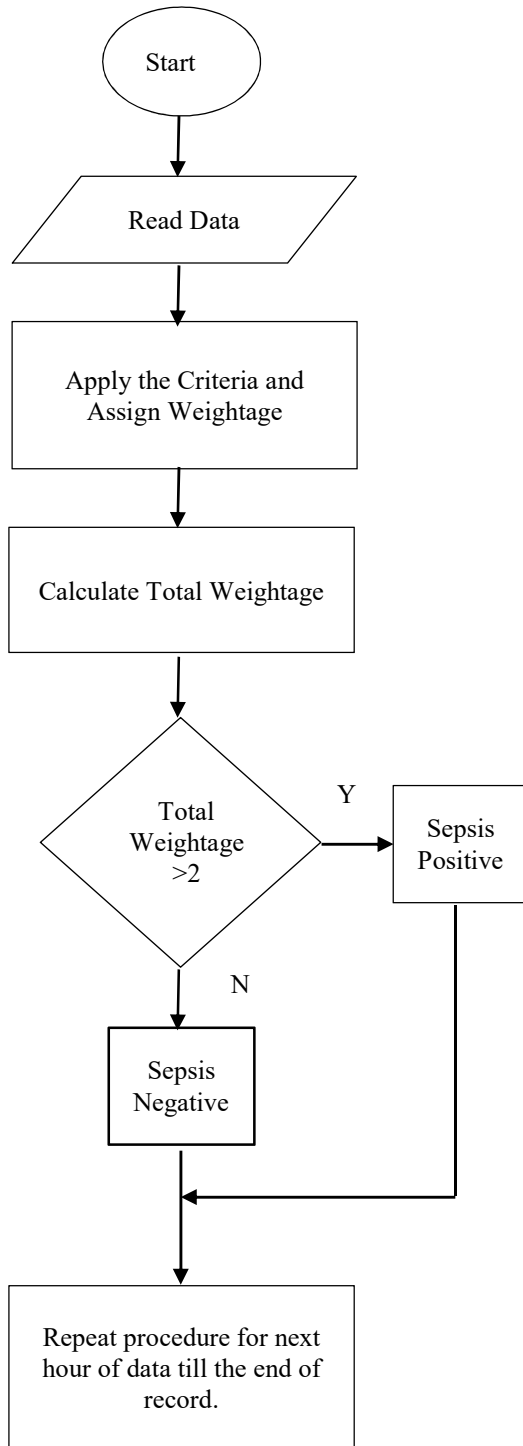
Figure 1. Sepsis Detection Algorithm

Total Weight = $\sum$ Weights of Individual Parameter

Sepsis is marked as positive (1) if,

Total Weight is > 2

Otherwise sepsis is marked as negative (0) in sepsis detection column.

The procedure was repeated for each hour data till the end of the record. And if any one instances was positive the patient was considered as positive case of sepsis.

3. Results

The algorithm was tested on the training data set and then submitted for scoring of the challenge. The organisers graded the algorithm by the utility function, which rewards algorithm for early detection and penalize for late or missed detection or predicting sepsis in non-sepsis patients [8]. The participation in the challenge was with a team name SHODH. The final utility score received on full data set for the work submitted was 0.013 with ranking 68. Table 3 summarizes the results obtained for the present work, for the three data sets, A, B and C. As stated earlier dataset A & B were training data set and was available for testing the algorithm while dataset C was hidden to the participants and used to score the entries.

Table 3. Result Summary

	Test Set A	Test Set B	Test Set C
Utility Score	0.012	0.013	0.017
AUROC	0.529	0.506	0.586
AUPRC	0.023	0.015	0.016
Accuracy	0.968	0.978	0.984
F-Measure	0.033	0.040	0.048

4.   Discussion

According to results the utility score of the present method was on the lower side compare to the entry received highest score. This is because with the algorithm could detect sepsis for each hour of data but the early detection part was not able to complete. Also the lower values of Receiving Operating Curve (ROC) is because of low sensitivity. Though in terms of accuracy results are good, the algorithm needs further modification to improve the overall performance.

Also variations in the results are observed from one dataset to another. Reason for this is the conditions of different sepsis patients varies according to many factors such as, individual health condition, body response to drugs, antibiotics administrated. Therefore strategy used for one patient not necessarily work on other patient equally. And therefore with a good set of guidelines and advanced algorithms a clinical expertise of doctors is also an important aspect [10].

Another challenging part which was limiting the performance of the algorithm was missing data. In present study sepsis was detected for each hour, but if the data for that instance was missing then it limits the capability of detection. In such situations either immediate previous data or average of earlier data can be refer. This part was not implemented therefore need to be complete further.

5.   Conclusion

In this study, an automated low dimensional sepsis detection algorithm was proposed with optimum clinical parameters and modified criteria. The algorithm shows a capability of correct detection in terms of accuracy. So it can be concluded that selection of parameters and its weightage plays an important role in identification of sepsis patients. At the same time, sepsis detection remains challenging topic because it is a complex phenomenon to understand and treat.

References

[1] Singer M, Deutschman CS, Seymour CW, et al. "The third international consensus definitions for sepsis and septic shock (sepsis-3)," JAMA, 315:801-10, February, 2016.

[2] Paoli, Carly J., Mark A. Reynolds, Meenal Sinha, Matthew Gitlin and Elliott D Crouser.,"Epidemiology and costs of sepsis in the united states—an analysis based on timing of diagnosis and severity level," Critical care medicine, November, 2018.

[3] Kaukonen K, Bailey M, Suzuki S, et al., "Mortality related to severe sepsis and septic shock among critically ill patients in Australia and New Zealand 2000–2012", JAMA, 311(13):1308–1316, April, 2014.

[4] Torio CM , Andrews RM," National inpatient hospital costs:the most expensive conditions by payer, 2011" Statistical Brief #160, August, 2013.

[5] Bishal Gyawali, Karan Ramakrishna, Amit S Dhamoon, "Sepsis: The evolution in definition, pathophysiology, and management", SAGE Open Medicine, Volume 7: 1– 13, May, 2019.

[6] Kumar Anand, Daniel D. Roberts, Kenneth E. Wood, Bruce Light, Joseph E. Parrillo, Satendra Sharma, Robert Suppes, Daniel Feinstein, Sergio Zanotti, Leo Taiberg, David P. Gurka, Aseem Kumar and Mary S. Cheang, "Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock," Critical care medicine, 34,6,1589-96, June, 2006.

[7] Seymour, Christopher W., Foster C. Gesten, Hallie C Prescott, Marcus E Friedrich, Theodore J. Iwashyna, Gary S. Phillips, Stanley Lemeshow, Tiffany Medlin Osborn, Kathleen M Terry and Mitchell M. Levy, "Time to treatment and mortality during mandated emergency care for sepsis," The New England journal of medicine, 376, 23 ,2235-2244, May, 2017.

[8] Reyna MA, Josef C, Jeter R, Shashikumar SP, Westover MB, Nemati S, Clifford GD, Sharma A," Early prediction of sepsis from clinical data: the PhysioNet/Computing in Cardiology Challenge 2019," CriticalCare Medicine, 2019; In press.

[9] M. Shankar-Hari, D. A. Harrison, G. D. Rubenfeld and K. Rowan," Epidemiology of sepsis and septic shock in critical care units: comparison between sepsis-2 and sepsis-3 populations using a national critical care database", British Journal of Anaesthesia, 119 (4): 626–36, October, 2017.

[10]Singer,   Mervyn.   "Sepsis: personalization   v protocolization?" Critical Care, July, 2019.

Address for correspondence:

Aruna Deogire
arunadeogire@gmail.com