We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

An Explainable Machine Learning Model for Early Prediction of Sepsis Using ICU Data

Naimahmed Nesaragi and Shivnarayan Patidar

Abstract

Early identification of individuals with sepsis is very useful in assisting clinical triage and decision-making, resulting in early intervention and improved outcomes. This study aims to develop an explainable machine learning model with the clinical interpretability to predict sepsis onset before 6 hours and validate with improved prediction risk power for every time interval since admission to the ICU. The retrospective observational cohort study is carried out using PhysioNet Challenge 2019 ICU data from three distinct hospital systems, viz. A, B, and C. Data from A and B were shared publicly for training and validation while sequestered data from all three cohorts were used for scoring. However, this study is limited only to publicly available training data. Training data contains 15,52,210 patient records of 40,336 ICU patients with up to 40 clinical variables (sourced for each hour of their ICU stay) divided into two datasets, based on hospital systems A and B. The clinical feature exploration and interpretation for early prediction of sepsis is achieved using the proposed framework, viz. the explainable Machine Learning model for Early Prediction of Sepsis (xMLEPS). A total of 85 features comprising the given 40 clinical variables augmented with 10 derived physiological features and 35 time-lag difference features are fed to xMLEPS for the said prediction task of sepsis onset. A ten-fold cross-validation scheme is employed wherein an optimal prediction risk threshold is searched for each of the 10 LightGBM models. These optimum threshold values are later used by the corresponding models to refine the predictive power in terms of utility score for the prediction of labels in each fold. The entire framework is designed via Bayesian optimization and trained with the resultant feature set of 85 features, yielding an average normalized utility score of 0.4214 and area under receiver operating characteristic curve of 0.8591 on publicly available training data. This study establish a practical and explainable sepsis onset prediction model for ICU data using applied ML approach, mainly gradient boosting. The study highlights the clinical significance of physiological inter-relations among the given and proposed clinical signs via feature importance and SHapley Additive exPlanations (SHAP) plots for visualized interpretation.

Keywords: sepsis, early prediction, machine learning, explainable AI, electronic health records, clinical informatics, critical care, model-based diagnosis

1. Introduction

Sepsis is an enigmatic clinical condition that occurs when the patient's body reacts adversely to infection and as a consequence develops organ dysfunction.

Sepsis can practically affect all organ systems however, the organs involved and the degree of dysfunction varies distinctly among patients and can even lead to death in most cases [1, 2]. In the early stages of the disease, the treatment of sepsis seems to be relatively easy with the availability of broad-spectrum antibiotics [3]. While in the later stages of the disease, diagnosis of sepsis becomes much easier but extremely hard to treat. Therefore, early diagnosis of sepsis is the need of the hour for better clinical management [4].

Current manual assessment of sepsis using screening tools, like the Sequential Organ Failure Assessment (SOFA) score for ICU-patients, are complex in terms of measured clinical signs and even lack adequate sensitivity [5, 6]. On the other hand, AI and machine learning-based automated clinical decision support systems that use easily accessible clinical data have reflected a significant improvement in agreement with these treatment protocols in ICUs by guiding physicians through predefined work-flows [7–11]. In the current era wherein we have abundant availability of electronic medical records (EMRs) has brought more feasibility to such automated realizations [12]. However, almost every machine learning (ML)-based AI model and automated decision support system lack proper explainability because of their uninterpretable black-box nature [13, 14]. This is where Explainable Artificial Intelligence (XAI) comes in rescue to address some of these restrictions imposed by a Black-box AI system by adding explainability. And thus assist clinicians in the interpretation of their diagnosis, and recommend future actions to be taken thereby improving the quality of predictions [15–17]. The development of such an explainable ML framework for sepsis onset prediction is an important and active area of investigation.

This work presents a novel clinical application of developing an explainable ML framework for sepsis onset prediction among ICU patients based on the physiological medical knowledge of given clinical signs, obtained via extensive analysis, and using popular gradient boosting ML techniques. The framework's design includes an optimal explainable gradient boosting architecture for clinical decision making that investigates questions of generalizability and interpretability of the proposed system.

2. Methods

An overview of the proposed methodology from raw data to explainable decision framework is shown in **Figure 1**.

2.1 Dataset and study population

The publically available training set consists of data from two cohorts [18]. Cohort A has 790,215 records of 20,336 patients. Cohort B has 761,995 records of 20,000 patients. Particularly, data for every patient record contains 40 clinical covariates i.e. 8 vital signs, 26 laboratory values, and 6 demographic values. The labeling of the patient data was done adhering to Sepsis-3 clinical criteria. **Table 1** presents the details of various clinical covariates used under study together with their missing information in percentage [18, 19].

2.2 Feature extraction

Feature extraction takes place on the imputed version of given clinical data that generates features sample-wise on an hourly time grid. Two types of features ware generated namely:

Physiological features: In literature, inter-relations among the clinical values have been proven to enhance the capability of anomaly detection tasks [7, 20]. By reviewing various studies that justify the clinical significance of well-established physiological inter-relations among the given clinical signs 10 such physiological relations are derived from the given covariates: Three Shock Indices firstly the well defined Shock Index (SIndex) using Systolic BP and the other two are its modified versions proposed in this study for Diastolic BP (DPBSIndex) [21] and Mean

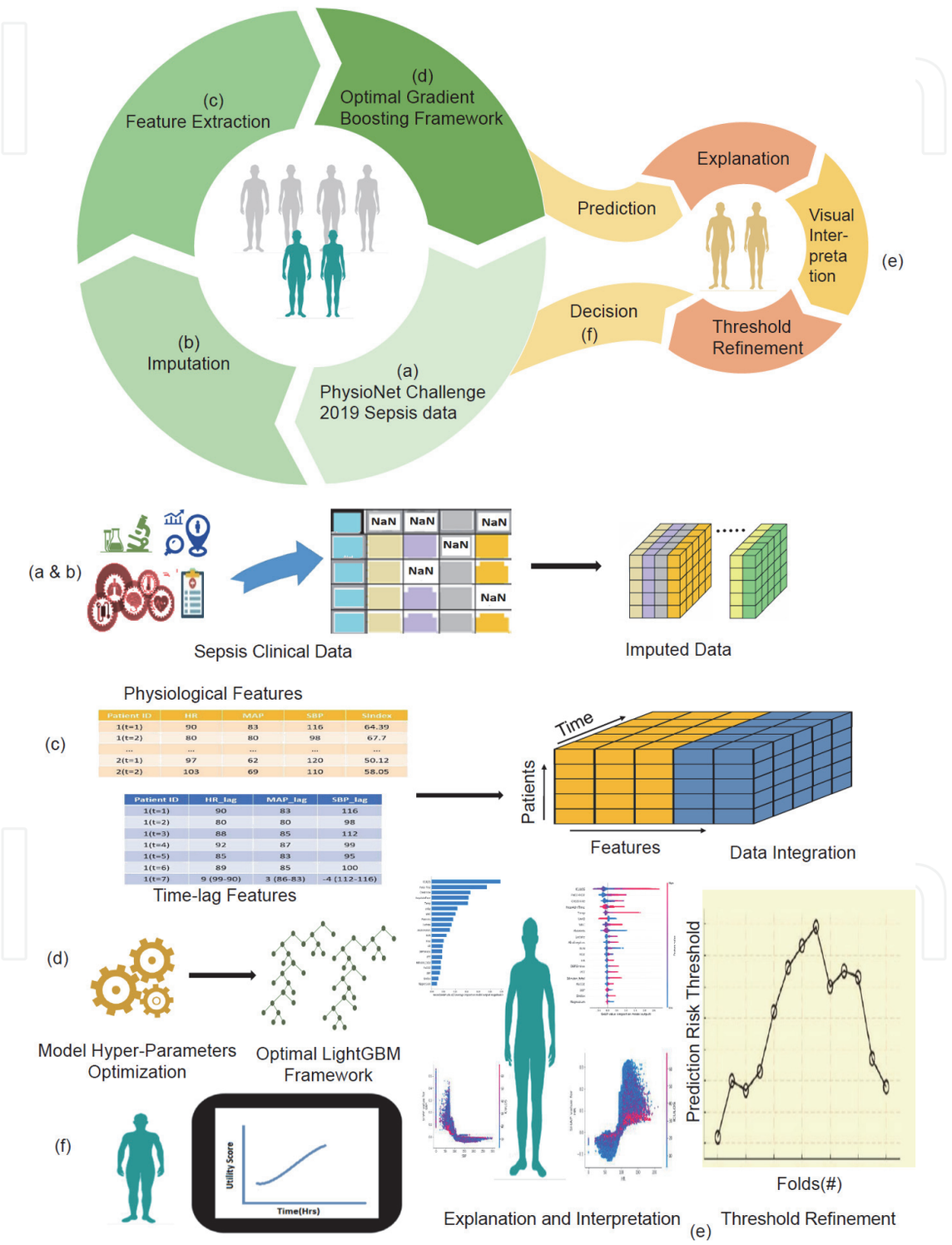

Figure 1.
Graphical overview: From given raw clinical data to explainable decision framework. (a & b) clinical data from two ICU cohorts is imputed. (c) Physiological inter-relations and time lag differences are computed as features. (d) an optimal sepsis onset prediction architecture is developed using LightGBM models via bayesian optimization. (e) the predictions are rendered to explanations and potentially their predictive power is increased by refining the threshold that drove the prediction at every time-point. (f) Final decision.

Sr. no.	Covariates	Missing values (%)	Units
1	Heart rate	9.8	beats/min
2	O_2Sat	13	%
3	Temperature	66	^{o}C
4	SBP	14.5	mm Hg
5	MAP	12.45	mm Hg
6	DBP	31.34	mm Hg
7	Resp	15.35	breaths/min
8	$EtCO_2$	96.28	mm Hg
9	Excess bicarbonate	95.57	mmol/L
10	Bicarbonate	94.81	mmol/L
11	FiO_2	91.66	%
12	pH	93.06	—
13	$PaCO_2$	94.44	mm Hg
14	SaO_2	96.54	%
15	Asparatate transaminase	98.37	IU/L
16	BUN	93.13	mg/dL
17	Alkaline phosphatase	98.39	IU/L
18	Calcium	94.11	mg/dL
19	Chloride	94.46	mmol/L
20	Creatinine	93.90	mg/dL
21	Direct bilirubin	99.8	mg/dL
22	Gl	82.89	mg/dL
23	Lactic acid	97.32	mg/dL
24	Magnesium	93.68	mmol/dL
25	Phosphate	93.98	mg/dL
26	Potassium	90.68	mmol/L
27	Total bilirubin	98.50	mg/dL
28	Troponin I	99.04	ng/mL
29	Hematocrit	91.14	%
30	Hemoglobin	92.61	g/dL
31	PTT	94.05	s
32	WBC	93.59	count/L
33	Fibrinogen concentration	99.34	mg/dL
34	Platelet count	94.05	count/mL
35	Age	—	yr
36	Gender	—	Male (1) or Female (0)
37	Unit 1	39.42	true (1) or false (0)
38	Unit 2	39.42	true (1) or false (0)
39	HospAdmTime	—	hours
40	LOS	—	hours

Sr. no.	Covariates	Missing values (%)	Units
41	SepsisLabel	—	septic (1) nonseptic (0).

Table 1.
Details of the various clinical variables used under study along with missing values information in percentage.

Sl. no	Abbreviation	Description	Formula
1	SIndex	Shock Index (SIndex) is the proportion of heart rate (HR) being divided by systolic blood pressure (SBP), normalized by age.	(HR/SBP)*Age
2	DBPSIndex	Diastolic Shock Index is the proportion HR being divided by systolic blood pressure (DBP), normalized by age.	(HR/DBP) *Age
3	MAPSIndex	It is defined as the proportion of HR being divided by Mean Arterial Pressure (MAP), normalized by age.	(HR/MAP) *Age
4	BUNCr	It is the ratio of Blood Urea Nitrogen(BUN) to Creatinine	BUN/ Creatinine
5	BILTCr	It is the ratio of Direct Bilirubin (Bilirubin_total) to Creatinine	Bilirubin_total/ Creatinine
6	SaO2 -FiO2	It is the ratio of oxygen saturation of arterial blood in percentage (SaO2) to the fraction of inspired oxygen (FiO2).	SaO2/FiO2
7	PaO2 -FiO2	It is defined as proportion of the partial pressure of oxygen PaO2 divided by the fraction of inspired oxygen (FiO2).	PaO2/FiO2
8	Pla_Age	It is the ratio of platelets to age	Platelets/Age
9	PP	Pulse Pressure (PP) is the difference between SBP and DBP	SBP-DBP
10	CO	Cardiac Output is the product of pulse pressure (PP) and HR.	PP * HR

Table 2.
Detailed definitions of the physiological features.

Arterial Pressure (MAPSIndex) [22] followed by ratios BUN/Creatinine (BUNCr) [7], Bilirubintotal/Creatinine (BILTcr), SaO2/FiO2 [23], PaO2/FiO2 [24], Platelets/Age (PlaAge), the difference between SBP and DBP called Pulse Pressure (PP) [25], and lastly Cardiac Output (CO) [26]. **Table 2** gives a detailed description of Physiological features.

Time-Lag difference features: A set of 35 time-lag features are computed with 6 hours of time-lag difference among vital signs and lab values from the given 40 clinical variables excluding the last 5 demographic values.

Finally, the obtained 45 features are combined with the given 40 clinical signs, thereby increasing the final feature count to 85 features. The resultant feature set is then fed to train the proposed xMLEPS framework.

2.3 Implementation of xMLEPS

Together with Bayesian optimization and the refinement of prediction risk threshold an optimal disease onset detection method before six hours for sepsis called xMLEPS is developed. As shown in **Figure 1** the given clinical sepsis data has large amount of missing information (approximately 20%). So at the onset of the algorithm computation, filling of these missing values is carried out as as a pre-processing step. The data imputation to fill in the missing values is done by employing forward fill imputation on the given EHR clinical data. In the real-time scenario, the current missing values encountered are to be filled with previous

available measurements. Thus only the previous clinical values of given EHR data are fetched for data imputation of current observation.

In this study, imputation is carried out into two rounds: first local imputation, for each individual record, and then global imputation for all the combined records together. In the case of local imputation, the trailing missing values in a row for a particular clinical covariate (or feature vector) are forward filled with the nearest past non-missing value in that row locally for the given record. Ipso facto, if the record encounters 'NaN' values, in the beginning, i.e. for the first alone measurement at t = 0, they are retained as it is initially and then later replaced with 'global mean' for that covariate row obtained by combining all records [19].

During model development, a ten-fold cross-validation scheme is employed wherein 10 LightGBM classifiers with the same complexity of model hyper-parameters obtained during Bayesian optimization are developed for the corresponding fold. The total feature set used to develop these models comprises of 85 features as described in *sub-Section 2.2*. Generally, hyper-parameter optimization aims at looking for the best hyper-parameter values to minimize the objective loss function. The hyper-parameter settings maximizing the custom-defined challenge metric- utility score on the subset of training data during the Bayesian optimization phase are later used to build models. These built models generate the predictions on the hold out 10% of validation data in each fold. The training process of the model in each fold stops when the utility score of the validation set does not show further improvements over 32 consecutive iterations, i.e. early stopping to best iteration is achieved to reset the model and thereby to avoid over-fitting.

The initial predictions generated by each optimal model on the corresponding validation data of each fold undergo refinement of the prediction risk threshold to enhance the utility score. The search space for the prediction risk thresholds lies in the range of 0 to 1 and is varied in steps of 0.05. Thus the threshold search space has 20 values. So the initial predictions of validation data of each fold are compared with each of these 20 values. After comparison, the threshold value that gives the maximum utility score for the set of predictions of that fold is said to be optimal. Such 10 optimum threshold values are later used by the corresponding models to refine the predictive power in terms of utility score for generated labels in each fold.

This LightGBM based gradient boosting framework serves with a specific processing method for sparse data which is important in our classification task with class imbalance problem [27]. For the interpretability of the proposed framework, the LightGBM uses its feature importance attribute to quantify each variable, and the explainability component is addressed by employing SHAP summary and dependency plots wherein the distribution of the variable importance is illustrated [28, 29].

3. Results

The proposed framework performs the prediction from the given patient-records to determine the risk of development of sepsis onset in the next 6 hours. This is achieved using a continuous-valued utility score as defined by challenge organizers for each prediction [18]. The utility function rewards or penalizes classifiers for their predictions within 12 hours before and 3 hours after sepsis onset time and was normalized as described in [18]. Using a ten-fold cross-validation scheme 10 LightGBM models are designed based on patient-wise stratified ten-folds each containing unique 10% of the entire training set. The hyper-parameters of the above models that minimize cross-validation loss are obtained by using automatic hyper-parameter optimization utility 'bayesopt' in Python [30, 31]. The underlying

objective function formulated for the optimization is intended to maximize the AUROC. The given software utility finds optimal parameters automatically using Bayesian optimization. At the outset, the optimized models includes: 60 '*num_leaves*', 120 '*min _data_in_leaf*', '*max _depth*' of 2, '*learning_rate*' of 0.01, '*scale_pos_weight*' of 20, '*min _samples_split*' of 4.

Table 3 gives a summary of the results by the proposed framework on the entire training data in a ten-fold cross-validation scheme. Results also include performances of inter-cohort and baseline studies. To ensure that the models trained in the proposed study learn dependencies not only between the patient-records but also among the cohorts, we considered inter-cohort training and testing scheme. i.e. model trained with the data of cohort A was scored on cohort B data and vice versa. This certainly avoids the doubt of the over-fitting, thus increasing the robustness of the framework. Inter-cohort scores for A and B were 0.3191 and 0.3284 respectively.

3.1 Comparison of xMLEPS with baseline

Further, to emphasize the clinical relevance of the derived features under this proposed method, a comparative analysis of results is done by carrying out three baseline studies as shown in **Figure 2**.

As a part of comparative analysis three well-tuned baseline studies are performed: Firstly, the proposed method with feature set of 85 features is tested without optimal threshold refinement (default threshold value of 0.5 with no skill is used) in a 10-fold cross-validation scheme. In the second and third methods, the given 40 clinical variables only are directly fed to LightGBM models with and without refinement of optimal threshold respectively in a 10-fold cross-validation scheme. **Table 3** presents the results of these three baseline studies accordingly. As expected the proposed method xMLEPS outperforms these three studies. The third study carried out without derived features and optimal threshold refinement shows

Fold	AUROC	F1	Utility	Threshold
1	0.8456	0.1420	0.4234	0.35
2	0.8436	0.1501	0.4029	0.35
3	0.8605	0.2208	0.4452	0.40
4	0.8610	0.1737	0.4331	0.35
5	0.8568	0.1507	0.4069	0.35
6	0.8607	0.1290	0.4253	0.25
7	0.8628	0.1475	0.4302	0.25
8	0.8649	0.1294	0.4205	0.20
9	0.8648	0.1299	0.3973	0.20
10	0.8704	0.1285	0.4282	0.20
Average (Std)	0.8591 (0.0085)	0.1502 (0.0286)	0.4214 (0.0148)	—
Baseline 1	0.8560	0.1517	0.3870	—
Baseline 2	0.8502	0.1376	0.3509	—
Baseline 3	0.8124	0.1197	0.3198	—
xMLEPS	Set A (Training) and Set B (Test)		0.3191	
xMLEPS	Set B (Training) and Set A (Test)		0.3284	

Table 3.
Results summary of the proposed framework.

worst performance. Even for the first baseline study, results are significantly lower by 3% in terms of the utility score as compared to the proposed method.

3.2 Explanation and visualization of feature importance

The cumulative feature importance of the first top 50 features is shown in **Figure 3**. Here the LightGBM feature importance attribute is used for the gradient boosting framework developed. The approach used is to count the number of times a feature gets involved to split the dataset across all trees. The failure of such an approach is that it accounts for different impacts due to different splits. The next best approach is to attribute the gain achieved with the reduction in average training loss when using a feature for splitting. This "Gain" measure used for feature importance recovers the correct mutual information between feature inputs and label outputs [32]. The limitation of this approach is that it gets easily biased when greedy trees are built in the finite ensembles. So other methods are designed to compensate for the bias in feature selection using gain approach [33, 34].

SHAP summary plot with the 20 most important clinical features that cause sepsis onset identified by the xMLEPS framework is shown in **Figure 4(a)**. Here the approach used for the feature importance is to sort all the relevance scores across the entire population in decreasing order of mean relevance as computed for local, but considering only those individuals who were positive for sepsis. The mean relevance is displayed as blue horizontal bars in **Figure 4(a)**. While local explanations summary is shown in **Figure 4(b)**, wherein all the individual data points are displaced by mean relevance for sepsis and are colored by feature values. As shown from **Figure 4(b)** we can draw that the increase in the length of stay (ICULOS) and higher value of clinical ratio's like PaO2/FiO2, Shock indices: DBPSIndex and SIndex, etc. leads to the development of sepsis, whereas lower Platelets, DBP and Magnesium levels cause sepsis. These findings are found to be consistent with previous studies on it [7, 21, 35, 36].

Further, the impact of each feature and the interactions among them for sepsis development can also be illustrated using SHAP dependency plots. As an example, in **Figure 4(c)** the dependency plot showing the interaction of Heart rate with ICULOS is depicted. As seen the xMLEPS model seems to associate high heart rate values in the range 120–180 with increased ICULOS and hence causing sepsis.

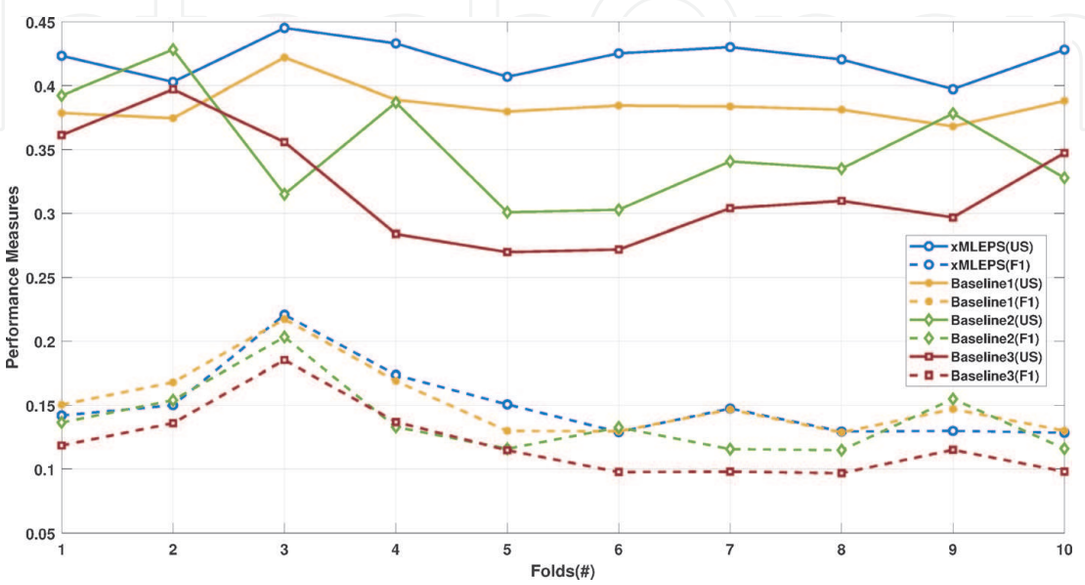

Figure 2.
Comparison of results by xMLEPS with the three base-line studies. US: Utility score, F1: F1 score.

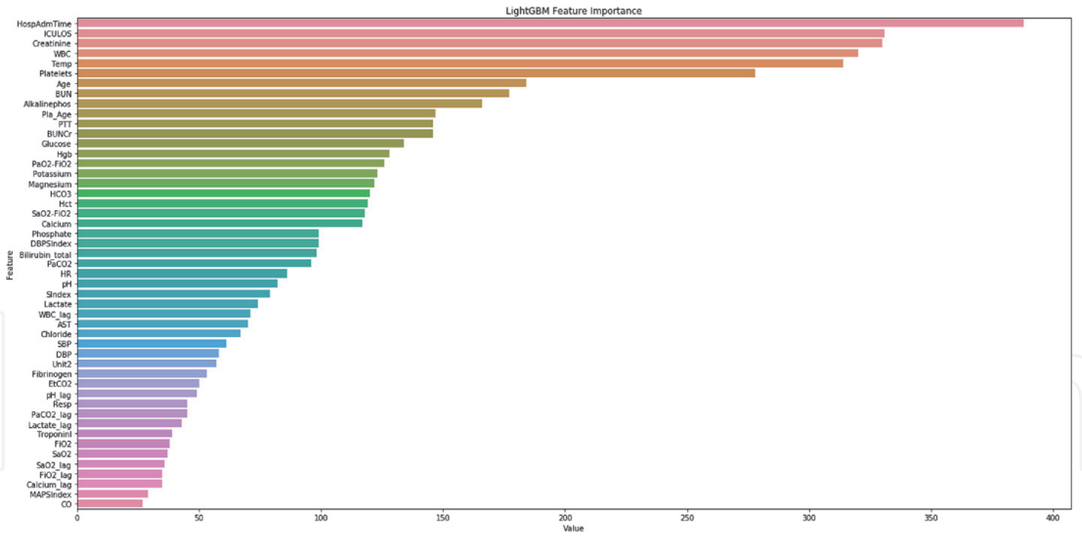

Figure 3.
Cumulative feature importance of the top 50 features using the feature importance attribute of LightGBM.

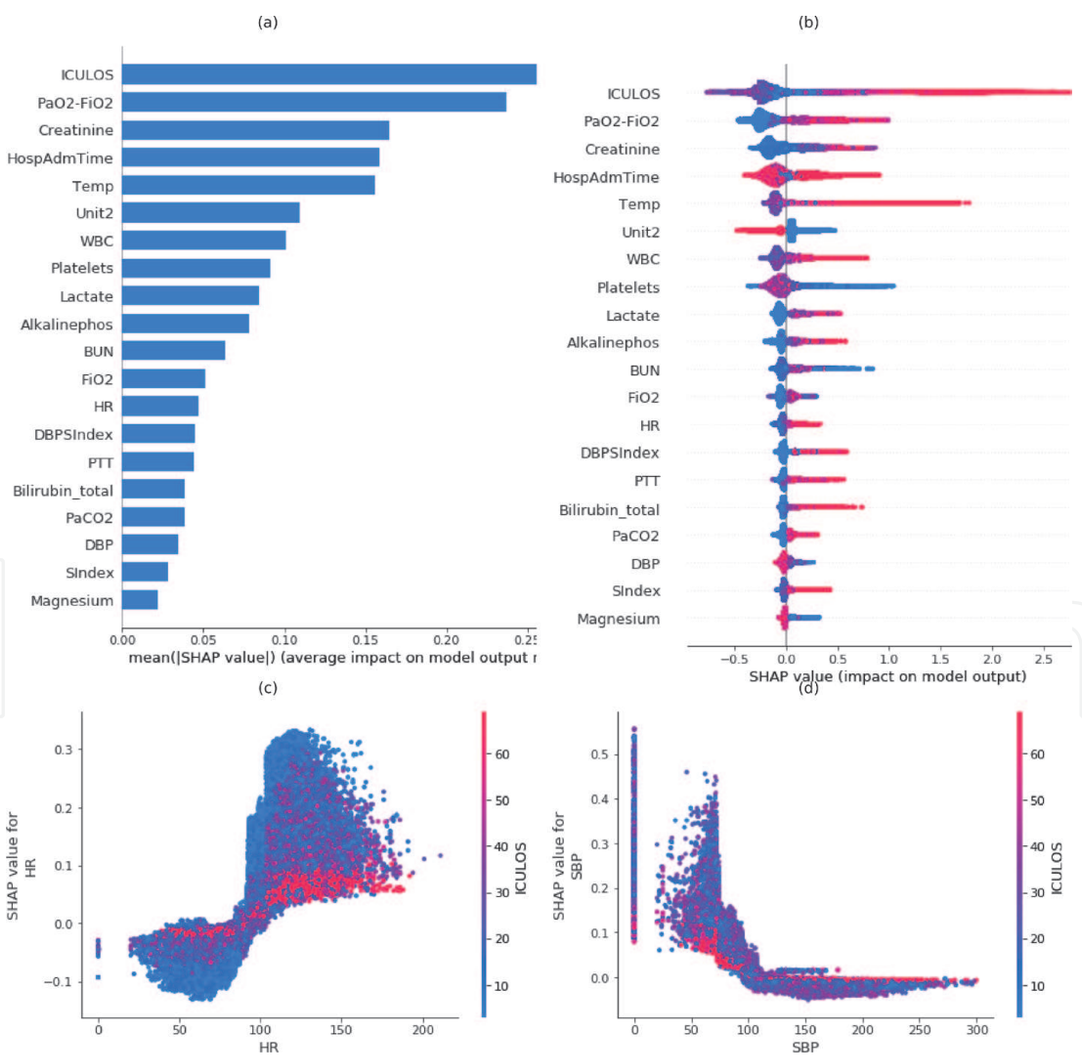

Figure 4.
Results from the SHAP explanation module showing the global feature importance together with local explanation summary. PaO_2: partial pressure of oxygen. FiO_2: fraction of inspired oxygen. HR: Heart Rate. DBP: Diastolic Blood Pressure. SBP: Systolic Blood Pressure. SIndex: Shock Index. DBPIndex: Diastolic Blood Pressure Shock Index. $PaCO_2$: Partial pressure of carbon dioxide. PTT: Partial thromboplastin time. WBC: Leukocyte count. BUN: Blood Urea Nitrogen.

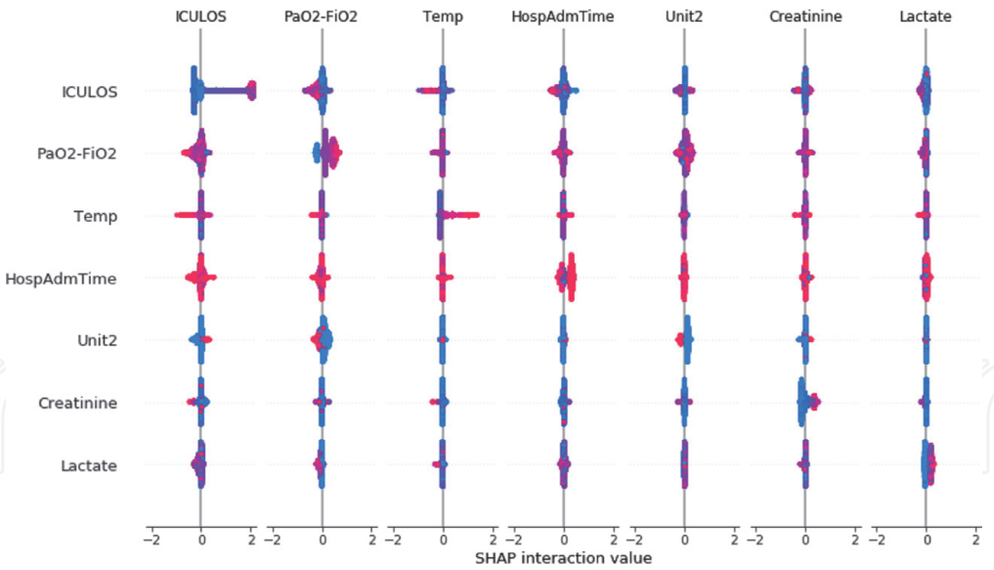

Figure 5.
Summary plot of a SHAP interaction value matrix.

Further **Figure 4(d)** shows lower values of SBP (approx. Below 90) is associated with increase ICULOS causing sepsis. A summary plot of a SHAP interaction value matrix is shown in **Figure 5** wherein the diagonal reflects the main effects, while across the diagonal show interaction effects. The explainable model will produce a high probability when it is confident about a decision, resulting in larger relevance scores due to the availability of more relevance for distributing backward. On the contrary, the model will output a lower probability when it is less confident about the patient to develop sepsis and as a result, yields lower relevant scores. This summary of scores distribution assists the clinicians with the hints to what to be expected from the designed model for clinical practice.

4. Discussion

This study justifies the clinical significance of the derived physiological interrelations among the clinical signs via feature importance and SHAP plots for visualized interpretation. Though SHAP values cannot be used as a generalized approach for early prediction of sepsis, they certainly help in generating relevant clinical hypotheses for desired events. The SHAP illustrations indeed assist in mitigating the concerns of the black-box issue associated with prediction models and might assist clinicians with a better understanding of the important features of the xMLEPS framework. The the proposed framework has the ability to establish the significance of the individual features contributing to enhance prediction of the utility score. Thus ensuring interpretablity of the framework to its clinical users. Furthermore, the proposed prediction framework, deploying clinical ICU data in the routine practice care can be potentially integrated into a computerized clinical decision support system instead of employing advanced molecular biomarkers.

The recent research literature relevant to early diagnosis of sepsis comes from the articles of various submission entries to PhysioNet Challenge 2019 [18]. This challenge aimed at the design and development of algorithms for early and automated prediction of sepsis onset with the optimal window definition of six hours before the actual clinical recognition of disease onset. The predictions of the machine learning algorithms were rewarded if they were able to detect true positives correctly up to 12 hours before disease onset and were slightly penalized if

Reference	Methodology	AUROC	Utility Score
Chang et al. [42]	Temporal Convolutional Networks (TCN)	—	0.4170
Li et al. [45]	A Time-phased model	—	0.4300
Morrill et al. [39]	A signature transform-based model	—	0.4340
Zabihi et al. [40]	XGBoost Ensemble models	0.8333	0.4280
Yang et al. [41]	Fusion-based XGBoost	0.8400	0.4300
Du et al. [46]	Gradient Boosting Scheme	0.8630	0.4090
Lee et al. [43]	Graph Convolutional Networks (GCN)	0.8170	0.3820
Lyra et al. [44]	Using Random forest classifier	0.8100	0.3760
Nesaragi and Patidar [38]	Ratio and Power-based features	0.8432	0.4013
Nesaragi et al. [19]	PMI-based Tensor factorization	0.8621	0.4519

Table 4.
Summary of the results obtained by our previous works and the submitted solutions to PhysioNet 2019 challenge under 5/10-fold cross-validation scheme using training data.

they were false positive. However predictions were strongly penalized if they were incorrect near disease onset. The reason for choosing the optimal prediction window to be six hours comes from the clinical fact that the ratio of observed median time to antimicrobial therapy is found to be 6 hours [37]. Furthermore, delay in each hour of treatment results in average decrease of survival rate of 7.6% [37].

The comparative analysis of the results obtained by the proposed method with our previous works [19, 38] and submission approaches [39–46] that reported the best results in the PhysioNet 2019 Challenge [18] is listed in **Table 4**. Most of these approaches utilized 5 or 10 fold cross-validation scheme and yielded utility scores in the range of 0.36–0.45.

This study supports the usage of the Utility score as an effective metric on ICU data for sepsis onset. However, experiments showed that even the F1 score gave reliable results aligning with utility score. i.e. the increase and decrease of F1 score follow accordingly with the Utility score. However, the bounds for utility score vary from −2 to 1 whereas the F1 Score has bounds from 0 to 1. The other conventional metrics namely AUROC, AUPRC, and Accuracy are insignificant to use with such a highly unbalanced dataset and are misleading for sepsis onset. Further, the fact that the interpretation of these results together with utility score is quite difficult cannot be ignored as mentioned by Roussel et al. [47].

The limitation of this study is, it constrains only to a two-center cohort design from the available training data, which might create doubt that the trained models may get over-fit towards the particular cohort data and it's patient-records. However, the analyzed ICU patient admissions originate from a diverse population covering the entire spectrum of ICU patients, and further, the validation in terms of inter-cohorts train-test approach along with optimum threshold refinement demonstrates the deployment of our framework in other ICUs.

5. Conclusion

This study presents xMLEPS – an explainable machine learning framework for the early prediction of sepsis using clinical data in the ICU setting. These predictive explanations justify the clinical significance of physiological inter-relations among the given clinical signs via visualized interpretation. And thus assist the clinicians in

decision making for diagnosis and recommend future actions to be taken to improve the quality of predictions. This certainly ensures that the data-driven automated ML models have the potential to make the paradigm shift from conventional detection and treatment to an automated early prediction that prevents the failure of the organ system due to sepsis.

Conflict of interest

The authors declare no conflict of interest.

Author details

Naimahmed Nesaragi† and Shivnarayan Patidar*†
National Institute of Technology, Goa, India

*Address all correspondence to: shivnarayan.patidar@nitgoa.ac.in

† These authors contributed equally.

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] M. Singer, C. S. Deutschman, C. W. Seymour, M. Shankar-Hari, D. Annane, M. Bauer, R. Bellomo, G. R. Bernard, J.-D. Chiche, C. M. Coopersmith, R. S. Hotchkiss, M. M. Levy, J. C. Marshall, G. S. Martin, S. M. Opal, G. D. Rubenfeld, T. van der Poll, J.-L. Vincent, D. C. Angus, The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3), JAMA 315 (2016) 801.

[2] L. Giesen, M. Singer, What Is Sepsis?, in: W. J. Wiersinga, C. W. Seymour (Eds.), Handbook of Sepsis, Springer International Publishing, Cham, 2018, pp. 3–14.

[3] H. Nishie, Guidelines for management of severe sepsis and septic shock, Okayama Igakkai Zasshi (Journal of Okayama Medical Association) 125 (2013) 153–157.

[4] C. W. Seymour, J. N. Kennedy, S. Wang, C.-C. H. Chang, C. F. Elliott, Z. Xu, S. Berry, G. Clermont, G. Cooper, H. Gomez, D. T. Huang, J. A. Kellum, Q. Mi, S. M. Opal, V. Talisa, T. van der Poll, S. Visweswaran, Y. Vodovotz, J. C. Weiss, D. M. Yealy, S. Yende, D. C. Angus, Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis, JAMA 321 (2019) 2003.

[5] R. Goulden, M.-C. Hoyle, J. Monis, D. Railton, V. Riley, P. Martin, R. Martina, E. Nsutebu, qSOFA, SIRS and NEWS for predicting inhospital mortality and ICU admission in emergency admissions treated as sepsis, Emergency Medicine Journal 35 (2018) 345–349.

[6] V. Anand, Z. Zhang, S. S. Kadri, M. Klompas, C. Rhee, Epidemiology of Quick Sequential Organ Failure Assessment Criteria in Undifferentiated Patients and Association With Suspected Infection and Sepsis, Chest 156 (2019) 289–297.

[7] K. E. Henry, D. N. Hager, P. J. Pronovost, S. Saria, A targeted real-time early warning score (TREWScore) for septic shock, Science Translational Medicine 7 (2015) 299ra122–299ra122.

[8] J. Calvert, T. Desautels, U. Chettipally, C. Barton, J. Hoffman, M. Jay, Q. Mao, H. Mohamadlou, R. Das, High-performance detection and early prediction of septic shock for alcohol-use disorder patients, Annals of Medicine and Surgery 8 (2016 a) 50–55.

[9] J. S. Calvert, D. A. Price, U. K. Chettipally, C. W. Barton, M. D. Feldman, J. L. Hoffman, M. Jay, R. Das, A computational approach to early sepsis detection, Computers in Biology and Medicine 74 (2016 b) 69–73.

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

[11] T. Desautels, J. Calvert, J. Hoffman, M. Jay, Y. Kerem, L. Shieh, D. Shimabukuro, U. Chettipally, M. D. Feldman, C. Barton, D. J. Wales, R. Das, Prediction of Sepsis in the Intensive Care Unit With Minimal Electronic Health Record Data: A Machine Learning Approach, JMIR Medical Informatics 4 (2016) e28.

[12] K. Kawamoto, C. A. Houlihan, E. A. Balas, D. F. Lobach, Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success, BMJ 330 (2005) 765.

[13] F. Cabitza, R. Rasoini, G. F. Gensini, Unintended Consequences of Machine Learning in Medicine (2017) 2.

[14] D. Petkovic, L. Kobzik, C. Re, Machine learning and deep analytics for

biocomputing: call for better explainability, in: Biocomputing 2018, WORLD SCIENTIFIC, Kohala Coast, Hawaii, USA, 2018, pp. 623–627.

[15] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, D. Pedreschi, A Survey of Methods for Explaining Black Box Models, ACM Computing Surveys 51 (2019) 1–42.

[16] written on behalf of AME Big-Data Clinical Trial Collaborative Group, Z. Zhang, M. W. Beck, D. A. Winkler, B. Huang, W. Sibanda, H. Goyal, Opening the black box of neural networks: methods for interpreting neural network models in clinical applications, Annals of Translational Medicine 6 (2018) 216–216.

[17] S. Tan, M. Soloviev, G. Hooker, M. T. Wells, Tree Space Prototypes: Another Look at Making Tree Ensembles Interpretable, arXiv: 1611.07115 [cs, stat] (2020). ArXiv: 1611.07115.

[18] M. A. Reyna, C. S. Josef, R. Jeter, S. P. Shashikumar, M. B. Westover, S. Nemati, G. D. Clifford, A. Sharma, Early Prediction of Sepsis From Clinical Data: The PhysioNet/Computing in Cardiology Challenge 2019, Critical care medicine 48 (2020) 210–217. Publisher: Lippincott Williams & Wilkins.

[19] N. Nesaragi, S. Patidar, V. Aggarwal, Tensor learning of pointwise mutual information from ehr data for early prediction of sepsis, Computers in biology and medicine (2021) 104430.

[20] R. J. Delahanty, J. Alvarez, L. M. Flynn, R. L. Sherwin, S. S. Jones, Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis, Annals of Emergency Medicine 73 (2019) 334–344.

[21] G. A. Ospina-Tascón, J.-L. Teboul, G. Hernandez, I. Alvarez, A. I. Sánchez-Ortiz, L. E. Calderón-Tapia, R. Manzano-Nunez, E. Quiñones, H. J. Madriñan-Navia, J. E. Ruiz, J. L. Aldana, J. Bakker, Diastolic shock index and clinical outcomes in patients with septic shock, Annals of Intensive Care 10 (2020) 41.

[22] Y.-c. Liu, Modified shock index and mortality rate of emergency patients, World Journal of Emergency Medicine 3 (2012) 114.

[23] A. Linder, R. Arnold, M. Zindovic, I. Zindovic, A. Lange-Jendeberg, M. Paulsson, P. Nyberg, B. Christensson, P. Åkesson, Heparin-binding protein improves prediction of severe sepsis in the emergency department, Critical Care 17 (2013) P3.

[24] Y. Wang, H. Yang, L. Qiao, Z. Tan, J. Jin, J. Yang, L. Zhang, B. M. Fang, X. Xu, The predictive value of PaO2/FIO2 and additional parameters for in-hospital mortality in patients with acute pulmonary embolism: an 8-year prospective observational single-center cohort study, BMC Pulmonary Medicine 19 (2019) 242.

[25] A. Y. A. Amer, J. Vranken, F. Wouters, D. Mesotten, P. Vandervoort, V. Storms, S. Luca, B. Vanrumste, J.-M. Aerts, Feature Engineering for ICU Mortality Prediction Based on Hourly to Bi-Hourly Measurements, Applied Sciences 9 (2019) 3525.

[26] J.-L. Vincent, Understanding cardiac output, Critical Care 12 (2008) 174.

[27] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, Lightgbm: A highly efficient gradient boosting decision tree, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17, Curran Associates Inc., Red Hook, NY, USA, 2017, p. 3149–3157.

[28] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R.

Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, Explainable AI for Trees: From Local Explanations to Global Understanding, arXiv:1905.04610 [cs, stat] (2019). ArXiv: 1905.04610.

[29] S. Lundberg, S.-I. Lee, A Unified Approach to Interpreting Model Predictions, arXiv:1705.07874 [cs, stat] (2017). ArXiv: 1705.07874.

[30] E. Brochu, V. M. Cora, N. de Freitas, A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning, arXiv:1012.2599 [cs] (2010). ArXiv: 1012.2599.

[31] J. Snoek, H. Larochelle, R. P. Adams, Practical bayesian optimization of machine learning algorithms, in: Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2, NIPS'12, Curran Associates Inc., Red Hook, NY, USA, 2012, p. 2951–2959.

[32] G. Louppe, Understanding Random Forests: From Theory to Practice, arXiv: 1407.7502 [stat] (2015). ArXiv: 1407.7502.

[33] S. Chebrolu, A. Abraham, J. P. Thomas, Feature deduction and ensemble design of intrusion detection systems, Computers & Security 24 (2005) 295–307.

[34] V. A. Huynh-Thu, A. Irrthum, L. Wehenkel, P. Geurts, Inferring Regulatory Networks from Expression Data Using Tree-Based Methods, PLoS ONE 5 (2010) e12776.

[35] X. Li, Y. Kang, X. Jia, J. Wang, G. Xie, Tasp: A time-phased model for sepsis prediction, in: 2019 Computing in Cardiology (CinC), pp. Page 1–Page 4.

[36] D. W. Shimabukuro, C. W. Barton, M. D. Feldman, S. J. Mataraso, R. Das, Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial, BMJ Open Respiratory Research 4 (2017) e000234.

[37] A. Kumar, D. Roberts, K. E. Wood, B. Light, J. E. Parrillo, S. Sharma, R. Suppes, D. Feinstein, S. Zanotti, L. Taiberg, et al., Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock, Critical care medicine 34 (2006) 1589–1596.

[38] N. Nesaragi, S. Patidar, Early prediction of sepsis from clinical data using ratio and power-based features, Critical Care Medicine 48 (2020) e1343–e1349.

[39] J. H. Morrill, A. Kormilitzin, A. J. Nevado-Holgado, S. Swaminathan, S. D. Howison, T. J. Lyons, Utilization of the signature method to identify the early onset of sepsis from multivariate physiological time series in critical care monitoring, Critical Care Medicine 48 (2020) e976–e981.

[40] M. Zabihi, S. Kiranyaz, M. Gabbouj, Sepsis Prediction in Intensive Care Unit Using Ensemble of XGboost Models.

[41] M. Yang, X. Wang, H. Gao, Y. Li, X. Liu, J. Li, C. Liu, Early Prediction of Sepsis Using Multi-Feature Fusion Based XGBoost Learning and Bayesian Optimization.

[42] Y. Chang, J. Rubin, G. Boverman, S. Vij, A. Rahman, A. Natarajan, S. Parvaneh, A multi-task imputation and classification neural architecture for early prediction of sepsis from multivariate clinical time series, in: 2019 Computing in Cardiology (CinC), IEEE, pp. Page1–Page4.

[43] B. T. Lee, O.-Y. Kwon, H. Park, K.-J. Cho, J.-M. Kwon, Y. Lee, et al., Graph convolutional networks-based noisy

data imputation in electronic health record, Critical Care Medicine 48 (2020) e1106–e1111.

[44] S. Lyra, S. Leonhardt, C. Hoog Antink, Early Prediction of Sepsis Using Random Forest Classification for Imbalanced Clinical Data.

[45] X. Li, X. Xu, F. Xie, X. Xu, Y. Sun, X. Liu, X. Jia, Y. Kang, L. Xie, F. Wang, et al., A time-phased machine learning model for real-time prediction of sepsis in critical care, Critical Care Medicine 48 (2020) e884–e888.

[46] J. A. Du, N. Sadr, P. d. Chazal, Automated prediction of sepsis onset using gradient boosted decision trees, in: 2019 Computing in Cardiology (CinC), pp. Page 1–Page 4.

[47] B. Roussel, J. Behar, J. Oster, A recurrent neural network for the prediction of vital sign evolution and sepsis in icu, in: 2019 Computing in Cardiology (CinC), pp. Page 1–Page 4.

IntechOpen