2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A Machine Learning Understanding of Sepsis

Manish Shetty[†], Soumya Mary Alex[‡], Merlin Moni[‡], Fabia Edathadathil[‡], Preetha Prasanna[‡], Veena Menon[‡], Vidya P. Menon[‡], Prashanth Athri[§], Gowri Srinivasa[†⋆], *Senior Member, IEEE*

***Abstract*— Sepsis is a serious cause of morbidity and mortality and yet its pathophysiology remains elusive. Recently, medical and technological advances have helped redefine the criteria for sepsis incidence, which is otherwise poorly understood. With the recording of clinical parameters and outcomes of patients, enabling technologies, such as machine learning, open avenues for early prognostic systems for sepsis. In this work, we propose a two-phase approach towards prognostic scoring by predicting two outcomes in sepsis patients - *Sepsis Severity* and *Comorbidity Severity*. We train and evaluate multiple machine learning models on a dataset of 80 parameters collected from 800 patients at Amrita Institute of Medical Sciences, Kerala, India. We present an analysis of these results and harmonize consistencies and/or contradictions between elements of human knowledge and that of the model, using local interpretable model-agnostic explanations and other methods.**

I. Introduction

Sepsis - a condition caused most frequently by a systemic bacterial infection, but also by viral, fungal, and microbial infections, is amongst the leading causes of death in the world [1]. Mortality remains elevated post discharge, particularly in infants in low income settings [2]. The World Health Organization estimated 49 million cases and 11 million sepsis related deaths occurred in the year 2017, accounting for 20% of annual global deaths [1]. In the context of such alarming metrics, multiple studies have worked towards the early diagnosis and treatment of this condition [3]–[5]. More recent studies have also examined the effectiveness of existing treatment methods [6]. Effective treatment plans and preventive measures for patients at risk of developing sepsis is dependent on early and accurate diagnosis of the condition.

The definition of "sepsis" has evolved over the years with the expansion in knowledge of sepsis pathophysiology and medical prognosis [7], [8]. The first consensus definition of sepsis [9], developed in 1991, defines sepsis based on the occurrence of two or more Systemic Inflammatory Response Syndrome (SIRS) criteria in response to an infection. Based on severity, it also classifies sepsis sub-types as *"Sepsis"*, *"Severe Sepsis"*, and *"Septic Shock"*. Here, *"Severe Sepsis"* is the condition of sepsis with an associated organ dysfunction, and *"Septic Shock"* further includes refractory hypotension [10].

While the classification based on severity is useful, in the following years much has been discussed about the effectiveness of the SIRS criteria. Recently, recognizing the need to re-examine current definitions, the European Society of Intensive Care Medicine (ESICM) and Society of Critical Care Medicine (SCCM) convened a task force of 19 specialists. The current use of 2 or more SIRS criteria to identify sepsis was unanimously considered unhelpful. They defined sepsis (Sepsis-3) [11] as - *"a life-threatening organ dysfunction caused by a dysregulated host response to infection"*. They state further, that organ dysfunction can be represented by an increase in the Sequential [Sepsis-related] Organ Failure Assessment (SOFA) score of 2 points or more.

With advances in machine learning, multiple efforts have been made to tackle the task of predicting sepsis. Prior work includes exploration of survival models [12], Hidden Markov Models [13], and Recurrent Neural Networks [14]. Such models rely on the assumption that a patient's clinical parameters is a time series of variables and lack the interpretability necessary in the medical field. Other simpler approaches have also been explored, such as a modified Weilbull-Cox model for prediction of sepsis [5] and logistic regression models to predict mortality risk of sepsis patients [4].

In this work, we aim to develop and evaluate various simple machine learning methods for early prognosis of sepsis severity. This can enable downstream activities such as treatment planning and preventive measures to control the severity of the diagnosed condition. Inspired by various studies using admission-time data to diagnose patient health status and trends [15] and the potential of machine learning models to support decision making [16], we use data collected and provided by the Amrita Institute of Medical Sciences (AIMS) hospital in our study. Please refer to Section II for details on the data. For validation, we make use of the latest definition of "sepsis" as well as the severity classification and frame the primary problem as the prediction of ***"Sepsis Severity"*** using the "Surviving Sepsis Campaign" guidelines [17].

Comorbidities are coexistent diseases to a disease of interest or an index disease, which may directly affect the prognosis of the disease of interest, or indirectly influence the choice of treatment [18]–[20]. Chronic comorbid medical conditions are present in 54–65% of all sepsis patients [21], [22] and strongly influence outcomes. Typically, the severity of patients' comorbidities are evaluated by summary measures that attribute fixed weights to various conditions, and

[†]M. Shetty and G. Srinivasa are with the PES Center for Pattern Recognition, Department of Computer Science and Engineering, PES University, Bengaluru, India.

[‡]S. M. Alex, M. Moni, E. T. Fabia, P. Prasanna, V. Menon, and V. P. Menon are with Amrita Institute of Medical Sciences, Kochi, India

[§]P. Athri, is with the Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India

[⋆]G. Srinivasa is the corresponding author: gsrinivasa@pes.edu

[1]https://www.who.int/news-room/fact-sheets/detail/sepsis

then sum the weights of those conditions that are present in a patient. We propose that such measures can be viewed as a transformation over recorded clinical parameters that define presence or absence of these comorbid conditions. Thus, we additionally aim to predict ***"Comorbidity Severity"*** as a bucket/range of one such prognostic metric - Charlson Comorbidity Index (CCI) [23]. Refer to Section II for more details on CCI bucketing.

Lastly, we evaluate the ability of our model to learn logical reasons to make critical decisions by interpreting the importance and weights assigned to features by the model. Additionally, we showcase our model's interpretability via local surrogate model explanations that can aid medical experts to validate and ensure effective treatment plans. Using these explanations, we prove that our model is in conformance with the latest clinical criteria defined to identify sepsis (Section V). In this work we make the following main contributions:

1) We propose a novel 2 phase approach towards a prognostic scoring system by predicting two complementary outcomes in sepsis patients - *Sepsis Severity* and *Comorbidity Severity*.
2) We evaluate multiple machine learning models on the data of 800 patients using 5 fold cross validation.
3) We interpret and explain multiple single patient predictions and map machine understanding to published domain knowledge that is currently used in the field.
4) Lastly, we open-source the code, model, and pre-processed data–https://interpretsepsis/repository.

II. Data Used in this Study

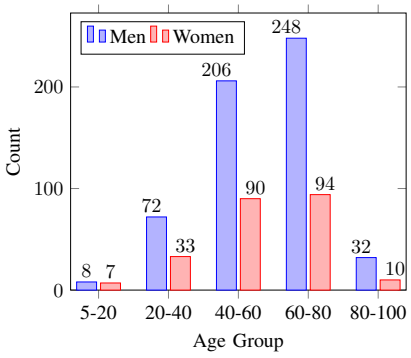

Fig. 1. Distribution of Patients

We make use of data of 800 patients, collected and provided by Amrita Institute of Medical Sciences (AIMS). The study has been approved by AIMS and complies with the Helsinki Declaration of 1975, as revised in 2000. It is ensured that all subjects in the data set have been thoroughly de-identified. This study is retrospective and has not influenced the course of treatment of the patients in anyway. Each patient record consists of around 80 features that can be categorized as–***On Admission Parameters*** and ***Clinical Parameters***. *On Admission Parameters* include features such as Age, Gender, Name, etc., that were collected on the patient's admission to the hospital. *Clinical Parameters* are features such as Heart Rate, Temperature, PH, etc., that were collected during treatment, i.e., the patient's stay in the Intensive Care Unit (ICU). Features also include specific sepsis and organ failure related parameters, such as *SOFA* (Sequential Organ Failure Assessment) scores [24]. SOFA scores were collected at 2 time-steps - on admission (*ASOFA*) and after a 72 hours stay in the ICU (*NSOFA*). For all practical purposes, a patient's data here is a flattened list of features that describes their overall health in the ICU, rather than a time series of recorded variables. This is particularly useful, as we can use instantaneous data, rather than historic, for near real-time prognosis of the severity of sepsis in any patient. Figure 1 shows the distribution of patient data with respect to their reported gender and age. In Figure 1, the upper limit for 'Age Group' is inclusive and lower is not.

We take a 2 phase approach towards a prognostic scoring system by predicting 2 outcomes (predictor variables) in sepsis patients-***Sepsis Severity*** and ***Comorbidity Severity***. In our study, *Sepsis Severity* can take 3 possible values:

1) `Sepsis` (34%)
2) `Severe Sepsis` (54.3%)
3) `Septic Shock` (11.7%)

Typically, *Comorbidity Severity* in patients is evaluated by a measure that attributes weights to various conditions present in a patient. The Charlson Comorbidity Index (CCI) [23] is the most widely used and verified comorbidity index. In our study, similar to Huang et al. [25], patients were assigned *Comorbidity Severity* by bucketing them into three groups based on CCI:

1) `Mild` (26.3%), if CCI < 3
2) `Moderate` (33.6%), if $3 \leq$ CCI ≤ 4
3) `Severe` (40.1%), if CCI ≥ 5

Note that, while predicting *Comorbidity Severity* we do not use the features–Charlson Comorbidity Index (CCI) and the known predictor variable *Sepsis Severity*, to avoid any bias.

III. Proposed Method

A. Data Pre-processing

Patient data collected at the bedside are prone to errors such as missing values, inconsistent spellings, etc. This is mainly due to manual filling of records as and when certain parameters are collected during treatment. All such instances were manually identified and fixed. Further, all binary valued parameters were converted to {True, False}. All categorical features were cleaned and converted to lowercase. Next, the cleaned data was pre-processed to be used as features for machine learning models. Categorical features were encoded [26] and any missing inputs were mapped to the value -1. Note, that we do not impute any categorical features.

On the other hand, numeric features were imputed using *Regression Imputation* [27]. Here, missing values can be imputed by fitting a curve to existing complete data, i.e., estimate the missing values by regression using other variables as parameters. A challenge with this approach is that the imputed data does not have an estimation of error. Thus, any imputed value can cause relationships to be over identified (reduce variance) and suggest greater precision. Thus, we

use a variation called *Stochastic Regression Imputation* [28] and add the average regression variance to the imputations.

TABLE I
COMPARISON OF 5 FOLD CROSS VALIDATION METRICS FOR ***Sepsis Severity*** CLASSIFICATION

Model	Label	Pre.	Rec.	F1	Acc. (± std)
AdaBoost	1	0.77	0.78	0.77	
	2	0.89	0.81	0.85	0.82 (± 0.037)
	3	0.48	0.67	0.56	
	Avg	0.80	0.78	0.79	
GradientBoosting	1	**0.82**	0.80	0.81	
	2	**0.93**	0.88	**0.91**	**0.92 (± 0.016)**
	3	0.52	0.67	**0.58**	
	Avg	0.85	0.83	0.84	
Linear SVM	1	0.66	0.54	0.60	
	2	**0.93**	0.42	0.58	0.64 (± 0.05)
	3	0.24	**0.90**	0.38	
	Avg	0.76	0.52	0.56	
Random Forest	1	**0.82**	**0.89**	**0.85**	
	2	0.86	**0.94**	0.90	**0.93 (± 0.02)**
	3	**0.89**	0.38	0.53	
	Avg	0.85	0.85	0.84	

TABLE II
COMPARISON OF 5 FOLD CROSS VALIDATION METRICS FOR ***Comorbidity Severity*** CLASSIFICATION

Model	Label	Pre.	Rec.	F1	Acc. (± std)
AdaBoost	1	0.87	**0.94**	**0.91**	
	2	**0.77**	0.49	0.60	0.81 (± 0.02)
	3	0.69	0.90	**0.78**	
	Avg	0.77	0.76	0.75	
GradientBoosting	1	0.88	0.91	0.90	
	2	0.72	**0.56**	**0.63**	**0.82 (± 0.05)**
	3	**0.71**	0.85	0.77	
	Avg	0.76	0.76	0.76	
Linear SVM	1	0.72	0.88	0.79	
	2	0.65	0.34	0.45	0.57 (± 0.05)
	3	0.67	0.87	0.76	
	Avg	0.68	0.68	0.66	
Random Forest	1	**0.91**	0.88	0.89	
	2	0.74	0.52	0.61	**0.85 (± 0.04)**
	3	0.69	**0.91**	**0.78**	
	Avg	0.77	0.76	0.74	

B. Models

Considering the size of our dataset and the high dimensionality of features, we chose supervised machine learning models that are known to perform well with such conditions. Support vector machines (SVMs) [29] are effective in high dimensional spaces with multiple choices for kernels. After experiments, we found that a linear kernel SVM with a one-vs-rest scheme is the best option for our high dimensional dataset. Another popularly used model is Random Forests [30] which achieves high accuracies and generalizability by constructing a multitude of decision trees with random feature selection to reduce overfitting. Recently, other models such as AdaBoost [31] and Gradient Boosting [32] classifiers have also been widely used. Both the algorithms boost the performance of a simple base-learner by iteratively shifting the focus towards problematic observations that are challenging to predict. We train and evaluate our models using the traditional K-fold cross validation, with an 80:20 train:test split. Note, in this work, we aim to use simple machine learning approaches over complex algorithms, guided by the scale of data, overcoming overfitting, and also prioritizing interpretability (Refer Section V).

IV. EVALUATION

Tables I & II present precision, recall, F1 (harmonic mean of precision and recall), and accuracy for 5-fold cross validation of all models for *Sepsis Severity* and *Comorbodity Severity* prediction respectively. The row marked *Avg* for each model refers to the weighted average of each metric (column) over all label types. In Table I, the label types 1,2, and 3 map to "Sepsis", "Severe Sepsis", and "Septic Shock" respectively. In Table II, the label types 1,2, and 3 map to "Mild", "Moderate", and "Severe" respectively.

From Table I, Random Forest and Gradient Boosting are the top contending models for sepsis severity predicition. To further improve the performance of the classifiers, we explored ensembling techniques that use a voting mechanism over different combinations of these models: a combination of Random Forest and Gradient Boosting yielded the highest F1 scores–{Sepsis: 0.86, Severe Sepsis: 0.91, Septic Shock: 0.65}. This combination performs no worse than either classifier used in exclusion and performs better than any other combination of classifiers in Table I. Similarly, for comorbidity severity prediction (refer Table II), we see that Random Forest and Gradient Boosting are the best models and also that their ensemble combination performs the best. Lastly, from Table I, we see that all models are relatively less accurate in predicting "Septic Shock" (Class 3). To understand this further, we explored the general variance within each class by measuring the mean (μ) and the variance (σ^2) of individual feature variances. We find that datapoints labeled "Septic Shock" ($\mu = 0.08, \sigma^2 = 0.006$) are more dispersed compared to "Sepsis" ($\mu = 0.05, \sigma^2 = 0.004$) and "Severe Sepsis" ($\mu = 0.06, \sigma^2 = 0.005$). This explains the inherent difficulty for models to predict "Septic Shock".

V. INTERPRETING MODEL EXPLANATIONS

When machine learning models are used as black-boxes in research, scientific findings remain completely hidden if models only provide predictions without explanations. In this study, by explaining the model's understanding of a complex condition such as sepsis, we aim to harmonize consistencies and/or contradictions between elements of human knowledge and that of the model. Here, we interpret our best model for sepsis severity prediction–an ensemble voting classifier using Random Forest and Gradient Boosting, by analyzing the importance assigned to features and model internals (weights/rules learnt). Further, we attempt to map

the model's understanding to published domain knowledge and literature that is widely used in the field today.

TABLE III
TOP FEATURES AND PERMUTATION FEATURE IMPORTANCES

Feature	Importance	± std
Was Lactate Measured?	0.145	0.023
ASOFA	0.114	0.025
NSOFA	0.025	0.016
SOFA diff	0.025	0.0077
SBP	0.02145	0.014
PLR	0.021	0.0087

TABLE IV
IMPORTANCES OF SIRS CRITERIA FEATURES

Feature	Importance	± std
Temperature	-0.00083	0.005
Heart rate	0.006	0.004
Respiratory rate	0.009	0.005
White blood cell count	0.015	0.008

Extracting the importance of features in Tree-based models using mean decrease in impurity measures (Gini, Entropy, etc.) can give high importance to features that may not be predictive on unseen data when the model is overfitting. Instead, we use Permutation Feature Importance [30], [33], which measures the increase in prediction error after permuting a feature's values. A feature is "important" if shuffling its values increases the model error, because the model relied on the feature for the prediction (and vice versa). From Table III, we observe that SOFA score (ASOFA & NSOFA) and the change in SOFA score (SOFA diff) are among the most important features for the model. Additionally, Systolic Blood Pressure (SBP), a direct indicator of hypotension for Septic Shock patients, and Platelet-to-Lymphocyte Ratio (PLR) [34], were also top features for the model. The above observations were found to be completely consistent with the latest decisions and recommendations from the *Third International Consensus Definitions for Sepsis and Septic Shock* (Sepsis-3) [11]. The Consensus also unanimously considered the current use of 2 or more SIRS criteria [9] to be unhelpful in determining sepsis incidence. They state the changes in white blood cell count, temperature, and heart rate reflect inflammation, but not necessarily a dysregulated, life-threatening response. They also mention the poor discriminant and convergent validity of these criteria in determining sepsis [35], [36]. In congruity, from Table IV, we observe that SIRS criteria features are relatively less important for the model's predictions.

To further explain the model's understanding and logical reasoning, we explore model-agnostic local explanations. Local interpretable model-agnostic explanations (LIME) [37] generates a new dataset consisting of perturbed samples

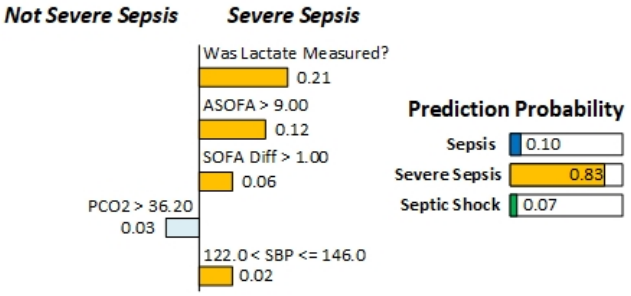

Fig. 2. Explanations for Severe Sepsis Prediction

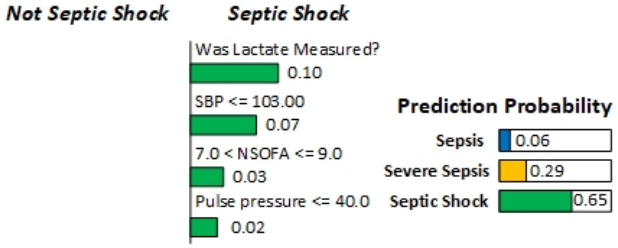

Fig. 3. Explanations for Septic Shock Prediction

for an individual data point (local) and the corresponding predictions of any black box model. On this new dataset, LIME trains an interpretable surrogate model (such as decision tree or lasso), which is weighted by the proximity of the sampled instances to the instance of interest. We then extract and interpret the weights of the surrogate model to explain the black box model. Figures 2 and 3, that show explanations for predictions, contain a series of features and their weights that directed the model towards a particular prediction. Figure 2 shows the explanations for the model predicting (correctly) a patient's severity as *Severe Sepsis*. We observe that explanations like `SOFA Diff > 1.00` are in conformance with consensus defined rules [11], i.e., an increase in SOFA of 2 points or more. We also see that Systolic Blood Pressure (SBP) is in the range $[122, 146]$, indicating an absence of hypotension (SBP < 90 mm Hg), thus lesser chances of septic shock. On the other hand, Figure 3 shows explanations for the model predicting (correctly) a patient's severity as *Septic Shock*. Here, we observe that the explanation–`SBP <= 103` indicates the model's ability to learn the presence of hypotension in septic shock patients. Singer et al. [11] also state that septic shock patients can be clinically identified by a vasopressor requirement to maintain a mean arterial pressure of 65 mm Hg or greater. This maps to the model's understanding that an `Arterial Pulse pressure <= 40` indicates septic shock. A similar analysis of a model's assignment of weights to features can be made to gain insight to comorbidity severity as well.

VI. CONCLUSION

With the analysis and model explanations in the foregoing Section, we conclude that simple models can learn logical rules to determine the severity of sepsis as well as the severity of accompanying comorbodities. These explanations not

only support conformance of model understanding with the knowledge of medical experts, but also provide interpretable assistance to doctors that can aid in early development of effective treatment plans. In the spirit of **reproducible research**, we have made available the preprocessed data, code, models and an interpretability notebook used to produce the results in this paper– https://interpretsepsis/repository.

REFERENCES

[1] K. E. Rudd, S. C. Johnson, K. M. Agesa, K. A. Shackelford, D. Tsoi, D. R. Kievlan, D. V. Colombara, K. S. Ikuta, N. Kissoon, S. Finfer *et al.*, "Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the global burden of disease study," *The Lancet*, vol. 395, no. 10219, pp. 200–211, 2020.

[2] B. Nemetchek, L. English, N. Kissoon, J. M. Ansermino, P. P. Moschovis, J. Kabakyenga, S. Fowler-Kerry, E. Kumbakumba, and M. O. Wiens, "Paediatric postdischarge mortality in developing countries: a systematic review," *BMJ open*, vol. 8, no. 12, p. e023445, 2018.

[3] R. O'Neill, J. Morales, and M. Jule, "Early goal-directed therapy (egdt) for severe sepsis/septic shock: which components of treatment are more difficult to implement in a community-based emergency department?" *The Journal of emergency medicine*, vol. 42, no. 5, pp. 503–510, 2012.

[4] G. Kong, K. Lin, and Y. Hu, "Using machine learning methods to predict in-hospital mortality of sepsis patients in the icu," *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, pp. 1–10, 2020.

[5] S. Nemati, A. Holder, F. Razmi, M. D. Stanley, G. D. Clifford, and T. G. Buchman, "An interpretable machine learning model for accurate prediction of sepsis in the icu," *Critical care medicine*, vol. 46, no. 4, p. 547, 2018.

[6] R. Ferrer, A. Artigas, D. Suarez, E. Palencia, M. M. Levy, A. Arenzana, X. L. Pérez, and J.-M. Sirvent, "Effectiveness of treatments for severe sepsis: a prospective, multicenter, observational study," *American journal of respiratory and critical care medicine*, vol. 180, no. 9, pp. 861–866, 2009.

[7] B. Gyawali, K. Ramakrishna, and A. S. Dhamoon, "Sepsis: The evolution in definition, pathophysiology, and management," *SAGE open medicine*, vol. 7, p. 2050312119835043, 2019.

[8] J.-L. Vincent, E. O. Martinez, and E. Silva, "Evolving concepts in sepsis definitions," *Critical Care Nursing Clinics*, vol. 23, no. 1, pp. 29–39, 2011.

[9] R. C. Bone, R. A. Balk, F. B. Cerra, R. P. Dellinger, A. M. Fein, W. A. Knaus, R. M. Schein, and W. J. Sibbald, "Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis," *Chest*, vol. 101, no. 6, pp. 1644–1655, 1992.

[10] D. Annane, E. Bellissant, and J.-M. Cavaillon, "Septic shock," *The Lancet*, vol. 365, no. 9453, pp. 63–78, 2005.

[11] 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 *et al.*, "The third international consensus definitions for sepsis and septic shock (sepsis-3)," *Jama*, vol. 315, no. 8, pp. 801–810, 2016.

[12] S. Ghosh, J. Li, L. Cao, and K. Ramamohanarao, "Septic shock prediction for icu patients via coupled hmm walking on sequential contrast patterns," *Journal of biomedical informatics*, vol. 66, pp. 19–31, 2017.

[13] K. E. Henry, D. N. Hager, P. J. Pronovost, and S. Saria, "A targeted real-time early warning score (trewscore) for septic shock," *Science translational medicine*, vol. 7, no. 299, pp. 299ra122–299ra122, 2015.

[14] H. J. Kam and H. Y. Kim, "Learning representations for the early detection of sepsis with deep neural networks," *Computers in biology and medicine*, vol. 89, pp. 248–255, 2017.

[15] S. B. Asiimwe, A. Abdallah, and R. Ssekitoleko, "A simple prognostic index based on admission vital signs data among patients with sepsis in a resource-limited setting," *Critical Care*, vol. 19, no. 1, pp. 1–8, 2015.

[16] L. M. Fleuren, T. L. Klausch, C. L. Zwager, L. J. Schoonmade, T. Guo, L. F. Roggeveen, E. L. Swart, A. R. Girbes, P. Thoral, A. Ercole *et al.*, "Machine learning for the prediction of sepsis: a systematic review and meta-analysis of diagnostic test accuracy," *Intensive care medicine*, vol. 46, no. 3, pp. 383–400, 2020.

[17] A. Rhodes, L. E. Evans, W. Alhazzani, M. M. Levy, M. Antonelli, R. Ferrer, A. Kumar, J. E. Sevransky, C. L. Sprung, M. E. Nunnally *et al.*, "Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016," *Intensive care medicine*, vol. 43, no. 3, pp. 304–377, 2017.

[18] V. De Groot, H. Beckerman, G. J. Lankhorst, and L. M. Bouter, "How to measure comorbidity: a critical review of available methods," *Journal of clinical epidemiology*, vol. 56, no. 3, pp. 221–229, 2003.

[19] S. F. Hall, "A user's guide to selecting a comorbidity index for clinical research," *Journal of clinical epidemiology*, vol. 59, no. 8, pp. 849–855, 2006.

[20] R. Yancik, W. Ershler, W. Satariano, W. Hazzard, H. J. Cohen, and L. Ferrucci, "Report of the national institute on aging task force on comorbidity," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 62, no. 3, pp. 275–280, 2007.

[21] D. C. Angus, W. T. Linde-Zwirble, J. Lidicker, G. Clermont, J. Carcillo, and M. R. Pinsky, "Epidemiology of severe sepsis in the united states: analysis of incidence, outcome, and associated costs of care," *Read Online: Critical Care Medicine— Society of Critical Care Medicine*, vol. 29, no. 7, pp. 1303–1310, 2001.

[22] G. S. Martin, D. M. Mannino, S. Eaton, and M. Moss, "The epidemiology of sepsis in the united states from 1979 through 2000," *New England Journal of Medicine*, vol. 348, no. 16, pp. 1546–1554, 2003.

[23] M. Charlson, T. P. Szatrowski, J. Peterson, and J. Gold, "Validation of a combined comorbidity index," *Journal of clinical epidemiology*, vol. 47, no. 11, pp. 1245–1251, 1994.

[24] J.-L. Vincent, R. Moreno, J. Takala, S. Willatts, A. De Mendonça, H. Bruining, C. Reinhart, P. Suter, and L. G. Thijs, "The sofa (sepsis-related organ failure assessment) score to describe organ dysfunction/failure," 1996.

[25] Y.-q. Huang, R. Gou, Y.-s. Diao, Q.-h. Yin, W.-x. Fan, Y.-p. Liang, Y. Chen, M. Wu, L. Zang, L. Li *et al.*, "Charlson comorbidity index helps predict the risk of mortality for patients with type 2 diabetic nephropathy," *Journal of Zhejiang University Science B*, vol. 15, no. 1, pp. 58–66, 2014.

[26] A. Von Eye and C. C. Clogg, *Categorical variables in developmental research: Methods of analysis*. Elsevier, 1996.

[27] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[28] R. J. Little and D. B. Rubin, *Statistical analysis with missing data*. John Wiley & Sons, 2019, vol. 793.

[29] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[30] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[31] Y. Freund, R. E. Schapire *et al.*, "Experiments with a new boosting algorithm," in *icml*, vol. 96. Citeseer, 1996, pp. 148–156.

[32] J. H. Friedman, "Stochastic gradient boosting," *Computational statistics & data analysis*, vol. 38, no. 4, pp. 367–378, 2002.

[33] A. Fisher, C. Rudin, and F. Dominici, "Model class reliance: Variable importance measures for any machine learning model class, from the" rashomon" perspective," *arXiv preprint arXiv:1801.01489*, vol. 68, 2018.

[34] E. Can, Ş. Hamilcikan, and C. Can, "The value of neutrophil to lymphocyte ratio and platelet to lymphocyte ratio for detecting early-onset neonatal sepsis," *Journal of pediatric hematology/oncology*, vol. 40, no. 4, pp. e229–e232, 2018.

[35] M. M. Churpek, F. J. Zadravecz, C. Winslow, M. D. Howell, and D. P. Edelson, "Incidence and prognostic value of the systemic inflammatory response syndrome and organ dysfunctions in ward patients," *American journal of respiratory and critical care medicine*, vol. 192, no. 8, pp. 958–964, 2015.

[36] K.-M. Kaukonen, M. Bailey, D. Pilcher, D. J. Cooper, and R. Bellomo, "Systemic inflammatory response syndrome criteria in defining severe sepsis," *New England Journal of Medicine*, vol. 372, no. 17, pp. 1629–1638, 2015.

[37] M. T. Ribeiro, S. Singh, and C. Guestrin, "" why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.