

HSD5230 High Performance Computing

Homework 13

*Miao Cai**

2019-04-13

Read both the main paper and the appendix then answer the questions below.

Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. Crit Care Med. 2018;46(4):547-553. doi:10.1097/CCM.0000000000002936

- 1. Provide a brief summary (1 paragraph) of this paper.**
 - (a) What was the hypothesis or goal of the paper?**
 - (b) What was the outcome of the study?**

The goal of this paper is to develop and validate an Artificial Intelligence Sepsis Expert algorithm using electronic medical records and high frequency physiologic data for early prediction of sepsis. The outcome variable was whether sepsis occurred to the patient or not.

- 2. What statistic did the authors use to summarize the model performance (what classification metric)? Name 2 other potential statistics they could have chosen. What are the advantages/disadvantages of the one they chose compared with the other two you came up with?**

The authors used the area under receiver operating characteristic (AUROC), as well as specificity (1-false alarm rate) and accuracy at a fixed 85% sensitivity level to summarize model performance.

The other two potential statistics can be:

- mean square error
- loglikelihood

The ones that I choose are more general statistics, not statistics specific to survival analysis models. The statistics that the authors used were specific to the modified Weibull-Cox proportional hazard model in this study.

- 3. Did the study authors use a validation or testing dataset? Describe their approach to validation and testing. Do you think that is a good approach? Why or why not?**

The authors used a testing dataset. They randomly split the data into 80% training and 20% testing datasets, and calculate AUROC statistic for both training and testing sets. This may not be a good approach since one-time random split may be subject to sampling error. A better approach is cross-validation.

- 4. Would you describe their modeling approach as high or low capacity? Why? Support your answer.**

*Department of Epidemiology and Biostatistics, College for Public Health and Social Justice, Saint Louis University. Email miao.cai@slu.edu

I would say this is a high-capacity model since they have 65 features in this dataset, which is relatively high compared to a typical 10-variable model, although they have quite a few patients ($n = 27,537$).

5. What optimization algorithm did the authors use for model convergence?

They used a mini-batch stochastic gradient descent approach to maximize the log likelihood of the data.

6. Did the authors regularize the model? What sort of regularization did they use?

Yes, they used L1 - L2 regularization to minimize overfitting and optimize generalizability of the learned model.

7. Review the appendix graph summarizing the testing and training error. Do you think this model overfits the training data? Why or why not?

No. The training set and testing set have a close agreement, which indicates good generalizability and no overfitting issue.

8. How did the authors calculate variable importance? Why is this important to do?

They used AUROC, specificity, and accuracy to evaluate the relative importance of each variable. It is important since the authors want to know what variables can lead to better prediction.

9. Why is this study/approach novel compared to other prior sepsis prediction models? The authors discuss this in the discussion.

- Their algorithm is the first one to use data collected at different resolutions (low-resolution electronic medical records data and high-resolution BR and HR data).
- Their algorithm is the first one to demonstrate acceptable performance of sepsis prediction over incrementally longer time windows.

10. The authors have a nice discussion of prediction versus interdiction in the context of defining model value. Summarize that discussion in your own words. Do you think this model is proven to be of high value based on this report? How do we know if a tool like this can affect patient outcomes?

The algorithm is of high value since it can be used to identify the sources of sepsis, adapted and generalized to other hospitals, make real-time predictions for patients who have the risk of sepsis. These machine learning tools can be useful and affect patient outcomes if it has high prediction accuracy and specificity so that it can be used to guide clinical practice.