Improved Characterization of Patient Phenotypes for Advanced Heart Failure

A Capstone Report
presented to the faculty of the
School of Engineering and Applied Science
University of Virginia

by

T Read Schlomer

May 6, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

T Read Schlomer

Capstone advisor: Josephine Lamp, Department of Computer Science

The first task on which I worked was taking data from other studies involving congestive heart failure (CHF) and matching their features to the ones used to train the algorithm devised in Lamp (2019). Providing more training data has the potential to make the machine learning algorithm more accurate. How to match features wasn't always obvious. There were features in each set that in one may track an amount and the corresponding feature in the other was binary (yes/no). Frequently medications were recorded this way with one study recording a dose and the other whether the medication was being taken by the patient. Some studies would combine certain conditions or equipment into a single feature. For example, one study may place atrial fibrillation and atrial flutter into one feature and another a patient as having either a pacemaker or an Implantable Cardioverter Defibrillator into one feature. In these instances, availed with my experience as a Respiratory Therapist, what I could glean from study documentation and guidance from Josie Lamp, I was able to match the features in a way that maintained the value of the data as a training tool.

Technically, this involved manipulating data frames with general Python and the pandas library. The data was taken from an unwieldy csv format and each useful feature was placed into a data frame for further evaluation.

Secondly, I completed work to facilitate a comparison between CHF risk scores and the scores produced by the MVDD algorithm. Josie provided a list of trials which produced similar CHF risk scores. After careful evaluation it became apparent that comparison of some of these scores to the algorithm output would be of dubious value. Some of the scores relied heavily on features for which the algorithm wasn't programmed to consider as a feature. For instance, the Larissa score produced by Xanthopoulos et al. (2018) made use of red cell distribution width,

which accounted for 1 point on a four point scale. Attempting to compare this score and the MVDD algorithm output on data sets that didn't include this crucial feature didn't seem useful.

After we narrowed down which scores could be compared meaningfully, I wrote some python code, again manipulating pandas data frames to calculate the CHF risk scores. Using documentation on the scores I converted the text to python code. A data frame is passed to a score function which calculates the score for each patient, represented with a de-identified id, and stores both in a key-value pair in a dictionary which is returned to the calling function.

What I accomplished was limited in scope. There is much further work to be completed. Some of the risk scores have online calculators. A savvy individual could package up the features, submit them to the server and retrieve the score to be stored for comparison. I was unable to do this in my time on this project. I didn't compare the scores. This requires some alignment of the scores as they have different scales and the ranges for categories of severity of heart failure differ. The algorithm could be modified to accept multiple point of care measurements to account for the change in condition a patient undergoes from hospital admission to discharge. This project has the potential to aid clinicians in alleviating the suffering of those with CHF either directly or as a building block for further work. I'm grateful that the opportunity to contribute was extended to me.

References

Lamp, J.; Y. Wu, S. Lamp, L Feng, S Mazimba (2019). Intelligently Characterizing Patient Hemodynamic Phenotypes for Advanced Heart Failure in the ESCAPE Trial Using Learned Multi Valued Decision Diagrams.

Xanthopoulos, A., Tryposkiadis, K., Giamouzis, G., Konstantinou, D., Giannakoulas, G., Karvounis, H., Kattan, M, Skourlarigis, J., Parissis, J., Starling, R., Triposkiadis, F. (Mar. 2018). Larissa Heart Failure Risk Score: a proposed simple score for risk stratification in chronic heart failure. *European Journal of Heart Failure. 20(3)* 614-616 doi:10.1002/ejhf.1132