Name: Andrea Fox

Class: DSC630-T301 Predictive Analytics

Date: November 14, 2020

Assignment: 11.3 Assignment – Predictive Analytics Case Study

**Overview**

The Sisters of Mercy, now known as just Mercy is a health system that spans four states in the Midwest of the United States. The organization was originally founded in Dublin, Ireland with the very first House of Mercy opening in 1827 (Corbett, n.d.). This case study is about Mercy Health Systems and how they created a way to “predict severe sepsis and septic shock based on patient vital signs observed over time – detected 71% cases with an acceptable false positive rate.” (Siegel, 2016) Sepsis has been around for thousands of years and is considered a “dysregulated systemic inflammatory and immune response to microbial invasion that produces organ injury for which mortality rates are declining to 15-25%. Septic shock is “sepsis with hyperlactatemia and concurrent hypotension requiring vasopressor therapy, with in-hospital mortality rates approaching 30-50%. (Hotchkiss, et al., 2016) If you are unfamiliar with hospital care sepsis can lead to septic shock and in many cases death. This is something hospitals want to avoid as it is mostly a inpatient related infection.

**Data Understanding**

The business case for this project is simple, predict the risk of sepsis and septic shock by using patient data. Mercy started their project by using a large set of data using patient electronic health records (Hemsoth, 2012). This data is primarily unstructured if you open a patient record, but if accessing from the backend (making assumptions this is similar to how Cerner stores patient records) there are several structured tables that can be pulled in to make a holistic dataset. If I was performing this project, I would most likely create a checklist of sorts, that would lay out the different parts of the patient EHR that would be necessary to pull.

**Data Preparation**

Knowing that this dataset is going to be quite large I would work through my checklist to create a dataset that includes all the different variables that could help predict sepsis. For example, septic shock causes a huge downward swing in blood pressure, so making sure I include vital signs would be important. I would then start doing cleaning and some basic visuals that would help me identify if there are any outliers within the data. If I identify outliers I would then need to analyze if these are true outliers and something I need to incorporate, or possible errors that need to be filtered out. I would need to make sure I standardize everything as well since there will be several date and time stamps that may vary. Another piece of information I would need to remember when doing preparation and analysis is that most health data including height and weight are recorded using the metric system instead of what Americans are most familiar with. This could lead to inaccurate predictions later on if I do not standardize.

**Modeling**

For modeling of the data, I would most likely follow a similar process that we used in our term project and split the data then try it on a few different models. Based on experience I would say decision tree or random forest may potentially be a good option to use. In this case, sepsis and septic shock is also something CMS (Centers for Medicare & Medicaid Services) tracks and if a hospital system has a large percentage of patients contracting or dying from that may affect payment. Similar to my group’s term project I would also want to look specifically into false negatives, which would be cases where my model predicted the patient would not contract sepsis based on vitals but actually is the opposite of what happens.

**Revisit Business Understanding**

After looking at some initial models and before making a final decision on what model to deploy, I would need to revisit the business objectives to ensure that the data I am seeing matches to what the project is about. In this case study, Mercy had two major goals they wanted to achieve:

1. Compliance Analysis: “To evaluate and measure effectiveness of hospital compliance with SSC guidelines for addressing sever sepsis and S4 by analyzing patient EHR records.”
2. Early Detection Analysis: “To mine the EHR records for potential clinical indicators that could lead to early detection of S4.” (Hemsoth, 2012)

I would make sure that both of these goals are being addressed in the model before I move forward with deployment.

**Deployment**

For deployment I would have originally started planning at the very beginning of the project before extracting data since Abbott says this process can take time and there is not enough information out there for guidance (Abbott, 2014). The steps that I would need to plan for includes:

* Planning deployment
* Monitoring and maintaining models
* Creating a final report
* Reviewing the project and generating a list of lessons learned (Abbott, 2014)

Each of these are vital things that need to be considered for deployment. Once a model is deployed it will need to be monitored to make sure it is not depreciating too rapidly. If this starts to happen the model may need to be updated or completely rebuilt depending on how bad things are. Monitoring will also include putting together a report to prove that the model is being effective in predicting sepsis. This leads to a retrospective or going over how things went in data extraction and analysis, deployment of model, and the current state of the model. This will help predictive analysts improve in their skillsets and job.

# References

Abbott, D. (2014). *Applied Predictive Analytics.* Indianapolis: Wiley.

Corbett, N. (n.d.). *Mercy Quick Facts*. Retrieved from Mercy: https://www.mercy.net/newsroom/mercy-quick-facts/

Hemsoth, N. (2012, April 25). *Taking the Shock out of Sepsis with Big Data Analytics*. Retrieved from datanami: https://www.datanami.com/2012/04/25/taking\_the\_shock\_out\_of\_sepsis\_with\_big\_data\_analytics/

Hotchkiss, R. S., Moldawer, L. L., Opal, S. M., Reinhart, K., Turnbull, I. R., & Vincent, J.-L. (2016). Sepsis and Septic Shock. *PMC US National Library of Medicine*.

Siegel, E. (2016). *Predictive Analytics The Power to Predict Who Will Click, Buy, Lie, or Die.* Hoboken: Wiley.