Using daily inpatient Serum Creatinine to predict Chronic Kidney Disease in children

true

March 19, 2019

Abstract

To monitor kidney function and detect Acute Kidney Injury (AKI), pediatric inpatients typically have their serum creatinine measured daily. In this paper, I use these daily lab values to predict which patients will develop chronic kidney disease (CKD) after they leave the hospital. I pulled historical patient from the electronic health record system at Lucile Packard Children's Hospital and used these data to develop a supervised learning algorithm that estimates the risk of post-hospitalization CKD. I am working with the hospital to implement a risk prediction algorithm in the electronic health record system at the hospital. The algorithm will be used to identify patients who are good candidates for enrollment in a new clinic for managing CKD risk after hospitalization.

Introduction

Acute kidney injury (AKI) is a common comorbidity for pediatric inpatients. A growing body of evidence, summarized in (Macedo and Mehta 2016), suggests that some patients who experience AKI during their hospitalization go on to develop chronic kidney disease (CKD) over the course of the next year. Approximately 25% of pediatric intensive care patients meet the clinical criteria for AKI (Kaddourah et al. 2017), making it impossible for nephrologists to provide follow-up renal monitoring and care to every AKI survivor. The risk factors that determine which AKI survivors are more likely to develop CKD are largely unknown.

Dr. Scott Sutherland, a practicing nephrologist at Lucile Packard Children's Hospital (LPCH), would like to pilot a clinic where pediatric patients at risk of subsequent CKD development can recieve necessary care. I am working with Dr. Sutherland, Professor David Schienker from MS&E, and informatics stakeholders at LPCH to develop an algorithm that can use data on renal function from recently discharged patients to predict which patients are at elevated risk of developing CKD over the next 3 - 12 months.

To provide some relevant clinical background, creatinine is a waste product that is filtered out of the bloodstream by the kidneys. Therefore, serum creatinine (SCr), a measure of how much creatinine is in a patient's blood, is a critical and commonplace bioindicator for renal function. Most pediatric inpatients have their SCr measured on a daily basis. Acute Kidney Injury (AKI), one of the most common commodities for pediatric inpatients, is diagnosed based on an acute increase in serum creatinine over a seven-day period. A simple calculation involving a patient's height and serum creatinine is used to calculate the estimated glomerular filtration rate (eGFR), and a sustained eGFR is used to diagnose chronic kidney disease (CKD).

The goal of this project is to assign a CKD risk score to every patient meeting certain inclusion criteria at the time they are discharged from Lucile Packard Children's Hospital. I have pulled laboratory values and demographic information on about 4 thousand pediatric inpatients. Prior to this quarter, I have developed machine learning models (for example, random forest and regularized logistic regression) to predict CKD development using this dataset. However, these methods are limited because they cannot handle variable-length inputs and therefore the serum creatinine lab values must be pre-processed. The aim of this course project is to attempt to improve on the performance of these earlier models using deep learning methods.

Related work

I found no prior studies using serum creatinine to predict post-hospitalization CKD, but a few studies have used serum creatinine to predict in-hospital AKI. One such study [Weisenthal 2017] found that multilayer perceptron processing of the sum of patients' sCr led to a strong predictor of AKI in re-hospitalized patients. Other recent studies have attempted to predict AKI in the ICU (Malhotra et al. 2017), in pediatric stem cell transplant patients (Augustynowicz et al. 2019), and for liver failure patients (Maiwall et al. 2017).

Dataset and Features

I pulled data from the LPCH electronic health record for all hospitalizations abetween May 4, 2014 and Oct 16, 2017. Hospitalizations were excluded if the patient was diagnosed with CKD, end-stage renal disease, or experienced a kidney transfusion at this or a prior stay; if the patient was less than 3 months or more than 18 years of age at admission; if the patient was hospitalized for less than 24 hours; if their serum creatinine was not measured at least once during their stay; or if more than one value was reported for serum creatinine at the same data and time. Patients were considered positive for CKD stage 3 or higher after their discharge if their eGFR was below the clinical cutoff for at least two SCr measurements between 90 and 365 days after the discharge date. If their SCr was not measured at least twice after their discharge then their CKD status was considered unknown and they were therefore excluded from the CKD prediction dataset. A total of 4,179 patient stays were included in the labeled CKD prediction dataset. An additional 7,874 patient stays meet all other inclusion criteria but had unknown CKD status after discharge.

I wanted to ensure my train, dev, and test sets were balanced in terms of the label (CKD-positive or CKD-negative) as well as having patients with short and long stays in each. To accomplish this, I defined a temporary variable 'partition class' that assigned each patient to one of eight classess according to the following table. I then used StratifiedShuffleSplit from the sklearn.model_selection module to develop a train, develop, and test set that was balanced across these eight classes. My dataset had 3128 examples in the train set, 834 examples in the dev set, and 209 examples in the test set.

Table 1: Class assignment for stratified sampling

Outcome_Label	Count_SCr_measurements	Class	Count_examples
CKD-	1-2	A-	1056
CKD+	1-2	A+	67
CKD-	3	B-	586
CKD+	3	B+	41
CKD-	4	C-	455
CKD+	4	C+	33
CKD-	>4	D-	1727
CKD+	>4	D+	206

In the initial models described in this report, serum creatinine is the only feature being used to predict CKD. I do have additional features on each patient that I will attempt to use with serum creatinine in future work. Those features are: age, height, weight, sex, discharge disposition, race, and whether an AKI diagnosis was coded for the patient. For some examples, I also have serum creatinine measurements from before the hospital stay began.

Methods

I first trained a fully-connected neural network with 100 hidden units in a single hidden layer to predict the output using the three last serum creatinine measurements from the patient stay. To do so, I had to use reduced training and dev sets, which I'll call train_reduced_1 and dev_reduced_1, that excluded examples with fewer than 3 lab values (class A- and A+ from table 1). This decreased my training set from 3128 examples to 2286, and my dev set from 834 examples to 609.

Next, I wanted to see whether using 4 lab values instead of 3 was improved the ability to predict CKD. To do so, I had to further reduce the training and dev sets (train_reduced_2 and dev_reduced_2) to only include stayw with at least 4 lab values (class C+, C-, D+, and D-). Train_reduced_2 and Dev_reduced_2 have only 1861 and 481 examples, respectively. Each model was trained using cross-entropy loss using adam optimizer with a learning rate of 0.0001. Models were trained for 2,000 iterations.

Results

In all models, the probability of CKD assigned to every patient in the train and validation set was less than 0.5, which is unsurprising given the class imballance. For this reason, accuracy is a poor measure of performance. For this reason, I instead used AUC. Table two shows the AUCs for the three models trained.

Table 2: Class assignment for stratified sampling

Description	Dataset	Train_AUC	Dev_AUC
3 inputs, 1 hidden layer, 100 units	>2 SCr	0.7876	0.7866
3 inputs, 1 hidden layer, 100 units	>3 SCr	0.7890	0.7779
4 inputs, 1 hidden layer, 100 units	>3 SCr	0.7911	0.7762

For the model that uses three lab values trained with examples with least three lab values, the AUC was 0.786 in training. AUC actually decreased to 0.7866 for the dev set, suggesting overfitting is not a problem. For the model using three lab values trained with examples with at least four lab values, the AUC surprisingly increased for the training set, but was lower for the dev set. This is to be expected since the model had fewer training examples to learn from. For the model trained using the four most recent lab values, the training AUC went up as expected. The dev AUC decreased to below that of the other two models, which suggests overfitting. Becuase the model is already fairly simple, I suspect regularization would help address the overfitting.

Conclusion and Future Work

In this project, I showed that post-hospitalization CKD can be predicted from inpatient serum creatinine lab values alone. Unfortunately, I did not have time to construct a sequential model for predicting CKD from variable-length vectors of serum creatinine. That will be my next step.

Assuming that the sequential model outperforms the non-sequential models, I will then experiment with adding in other covariates. I plan to experiments with models structured as in the following figure.

Additionally, in my data I have thousands of patient stay examples that are unlabeled, because the patient did not have the post-say laboratory measurements necessary to classify the patient as CKD positive or negative. I plan to explore the use of transfer learning to train my RNN on predicting the next lab value in the sequence for these patients. Hopefully this will improve performance for predicting CKD in the patients with labels.

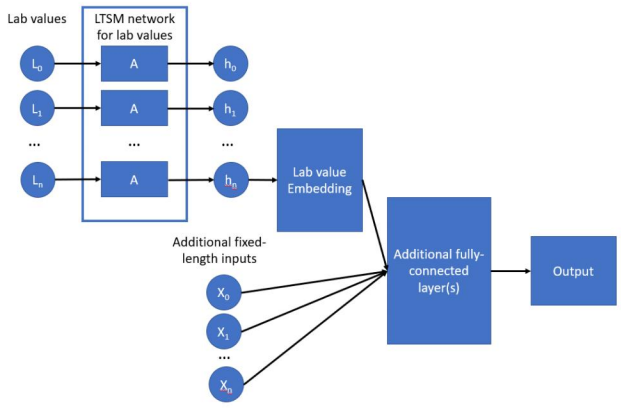

Figure 1: Diagram of model structure that I think would do well for this prediction task. I plan to experiemnt with models that have this structure.

Contributions

This work is part of my doctoral research. I am the only person directly analyzing data, but Professor David Scheinker, Management Science and Engineering, and Professor Scott Sutherland, Medicine, are advising the work.

References

Augustynowicz, Monika, Agnieszka Bargenda-Lange, Krzysztof Kałwak, Danuta Zwolińska, and Kinga Musiał. 2019. "Markers of acute kidney injury in children undergoing hematopoietic stem cell transplantation." *Advances in Clinical and Experimental Medicine* 28 (8): 0–0. doi:10.17219/acem/101573.

Kaddourah, Ahmad, Rajit K. Basu, Sean M. Bagshaw, and Stuart L. Goldstein. 2017. "Epidemiology of Acute Kidney Injury in Critically Ill Children and Young Adults." *New England Journal of Medicine* 376 (1): 11–20. doi:10.1056/NEJMoa1611391.

Macedo, Etienne, and Ravindra L. Mehta. 2016. "Renal recovery after acute kidney injury." *Contributions to Nephrology*. doi:10.1159/000443289.

Maiwall, Rakhi, Shiv Kumar Sarin, Suman Kumar, Priyanka Jain, Guresh Kumar, Ajeet Singh Bhadoria, Richard Moreau, et al. 2017. "Development of predisposition, injury, response, organ failure model for predicting acute kidney injury in acute on chronic liver failure." *Liver International* 37 (10): 1497–1507. doi:10.1111/liv.13443.

Malhotra, Rakesh, Kianoush B. Kashani, Etienne Macedo, Jihoon Kim, Josee Bouchard, Susan Wynn, Guangxi Li, Lucila Ohno-Machado, and Ravindra Mehta. 2017. "A risk prediction score for acute kidney injury in the intensive care unit." *Nephrology Dialysis Transplantation* 32 (5): 814–22. doi:10.1093/ndt/gfx026.