

# A Deep Learning Program to Predict Acute Kidney Injury

Xiaoqiang Li<sup>a,1</sup>

<sup>a</sup>Heidelberg University, Germany

**Abstract.** Acute kidney injury is a dangerous and sometime fatal clinical situation, which can cause irreversible damage. If we can predict it earlier and make appropriate prevention before its outbreak, kidney injury could be avoided. One challenge of early recognition of AKI is that the most e-alerts have focused on creatinine-based algorithms, but the elevation of serum creatinine lags behind renal injury. We use recurrent neural network (RNN) to make data mining on laboratory results of MIMIC-III Database. At first, we transfer the case data into Pandas DataFrame of series framed for supervised learning. Then we can use RNN predicts the next serum creatinine values (SCr) based on the last laboratory test results after emergency admissions. We train the RNN on whole dataset (i.e. multi-cases prediction) with LSTM. As the result shown, this prototype can predict criteria (SCr) of AKI with a RMSE (Root Mean Square Error) of 0.017mg/dL.

**Keywords.** acute kidney injury, prediction, artificial intelligence, neural network

## 1. Introduction

AKI (Acute kidney injury) is one of a number of conditions that affect kidney structure and function. AKI is defined as any of the following (Not Graded) [1]:

1. Increase in SCr (Serum creatinine) by  $\geq 0.3$  mg/dl ( $\geq 26.5$   $\mu$ mol/l) within 48 hours; or
2. Increase in SCr to  $\geq 1.5$  times baseline, which is known or presumed to have occurred within the prior 7 days; or
3. Urine volume  $< 0.5$  ml/kg/h for 6 hours.

A recent clinical practice assessment concluded there was an unacceptable delay in recognizing AKI in 43% of those that developed the condition after admission [4]. One challenge of early recognition of AKI is that the most e-alerts have focused on creatinine-based algorithms, but the current rule-based alerts cannot improve survival, because the elevation of serum creatinine lags behind renal injury [7].

---

<sup>1</sup> Corresponding Author, Xiaoqiang Li, MSc. Medicine Informatics, Heidelberg University, Germany. LIMS Software Engineer, Carl-Sonnenschein-Str 66 65936 Frankfurt Germany; E-mail: livvddcc@gmail.com.

## 2. Method

### 2.1. Mathematical methodology of this study

This study focuses on creating a program prototype to predict acute kidney injury by using the deep learning algorithm. Deep learning architectures such as deep neural networks, deep belief networks, recurrent neural networks (RNN) and convolutional neural networks have been applied to many fields, where they have produced results comparable to and in some cases surpassing human expert performance [2].

Step 1: We transform Bayes' theorem to a format, which looks like Sigmoid Function:

$$\begin{aligned}
 P(C_1|x) &= \frac{P(C_1, x)}{P(x)} \\
 &= \frac{P(x|C_1)P(C_1)}{p(x|C_1)P(C_1) + P(x|C_2)P(C_2)} \\
 &= \frac{1}{1 + \frac{P(x|C_2)P(C_2)}{p(x|C_1)P(C_1)}} \\
 &= \frac{1}{1 + e^{-a}} \quad [sigmoid]
 \end{aligned}$$

Where:  $P(C|x)$  is conditional probability, “a” represents function “a(x)”:

$$a(x) = \ln \frac{P(x|C_1)P(C_1)}{P(x|C_2)P(C_2)}$$

Step 2: We use Gaussian distribution to calculate  $P(x|C)$ :

$$P(x|C_1) \sim N(x|\mu_1, \Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x - \mu_1)^T \Sigma^{-1}(x - \mu_1)\right\}$$

$$P(x|C_2) \sim N(x|\mu_2, \Sigma) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x - \mu_2)^T \Sigma^{-1}(x - \mu_2)\right\}$$

$$\ln P(x|C_1) = -\frac{D}{2} \ln(2\pi) - \frac{1}{2} \ln|\Sigma| - \frac{1}{2}(x - \mu_1)^T \Sigma^{-1}(x - \mu_1)$$

$$\ln P(x|C_2) = -\frac{D}{2} \ln(2\pi) - \frac{1}{2} \ln|\Sigma| - \frac{1}{2}(x - \mu_2)^T \Sigma^{-1}(x - \mu_2)$$

So we can transform the function “a” to a format like  $(w^T x + b)$ , take “a” back to step 1, we get the whole Sigmoid Function now:

$$\begin{aligned}
 a(x) &= \ln P(x|C_1) - \ln P(x|C_2) + \ln \frac{P(C_1)}{P(C_2)} \\
 &= (\mu_1 - \mu_2)^T \Sigma^{-1} x - \frac{1}{2} \mu_1^T \Sigma^{-1} \mu_1 + \frac{1}{2} \mu_2^T \Sigma^{-1} \mu_2 + \ln \frac{P(C_1)}{P(C_2)} \\
 &= w^T x + w_0
 \end{aligned}$$

Where:

$$w = \Sigma^{-1}(\mu_1 - \mu_2)$$

$$w_0 = \frac{1}{2}\mu_2^T \Sigma^{-1} \mu_2 - \frac{1}{2}\mu_1^T \Sigma^{-1} \mu_1 + \ln \frac{P(C_1)}{P(C_2)}$$

Step 3: if we only have C1 and C2, then  $P(C_2|x)$ :

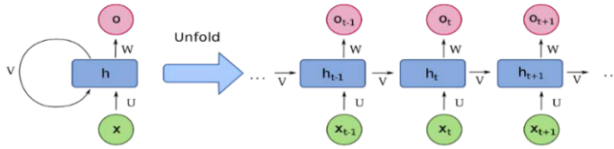
$$P(C_2|x) = 1 - P(C_1|x)$$

If we have C1, C2, ... Cn, then  $P(C_k|x)$ :

$$\begin{aligned} P(C_k|x) &= \frac{P(x|C_k)P(C_k)}{\sum_j P(x|C_j)P(C_j)} \\ &= \frac{\exp(a_k)}{\sum_j \exp(a_j)} \end{aligned}$$

Where  $k \in n$ .

After step 1 to 3, we get the Softmax function based on Bayes' theorem. At last, we build our recurrent neural network for prediction. On each ICU labor test (time series), the input is fed forward and a learning rule is applied. The back-connections save a copy of the previous values of the hidden units in the context units. Thus, it can learn the relation of time series data and make a prediction based on the history. That is why our RNN can predict the serum creatinine elevation, which normally lags behind the renal injury.



**Figure 1.** Structure of the recurrent neural network

Step 4: We use our neural network, which is based on Bayes' theorem, to predict real number (e.g., serum creatinine, range 0.00 – 2.00), but Bayesian network uses the probability value (0.00 – 1.00) as input and output. Therefore, we rescale the serum creatinine value to fit the input of Bayesian network and rescale the result (0.00 – 1.00) back to serum creatinine range (0.00 – 2.00) for prediction.

## 2.2. Data-Set

MIMIC-III (Medical Information Mart for Intensive Care III) is a database comprising de-identified health-related data associated with over forty thousand patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012 [3]. The patient data used in our study contains approximately 60000 admissions of patients including information such as patient demographics, vital signs, laboratory test results. In this prototype, we consider laboratory test results e.g., SCr (Serum creatinine), bicarbonate, blood urea nitrogen (BUN), chloride, international normalized ratio (INR), white blood count (WBC) as features, because these indicators are related to AKI

according to the past studies [4]. In this study the patient data do not have to be labeled manually.

2.3. Programming

Mining all the information from raw Data and dimensionality reduction are two main targets of data processing for neural network. Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning [5]. It cannot only process single data points, but also entire sequences of data. LSTM networks are well suited to making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. We use the ICU case data as multi-Case multivariate time series. At first, we transfer the case data into Pandas DataFrame of series framed for supervised learning. As example of this prototype, this program predicts the next serum creatinine values (SCr), which is the most important indicator of AKI. We use 3 time steps( $n=3$ ) for 1 prediction, i.e., we use the laboratory results of 3 time steps of one case to predict the next SCr value of next time step. For multi-cases predictions, the data-structure is as follow:

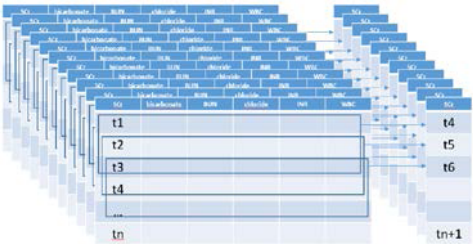


Figure 2. SCr. Prediction of multi-cases

2.4. Training of neural network

Keras API is used to train the dataset. Keras is a deep learning API written in Python, running on top of the machine learning platform TensorFlow [6], which is installed on CUDA GPU. GRU is also used and compared with LSTM to get better results and performance. 60% dataset is used for training and 40% for evaluation of training.

3. Result

Serum creatinine is reported as milligrams of creatinine to a deciliter of blood (mg/dL),  $n$ : timesteps for 1 prediction,  $inv\_yhat$ : forecast,  $inv\_y$ : actual, RMSE: Root Mean Square Error. We got following results:

$n=1$ :  $inv\_yhat$ : [1.33949034 1.44861898 1.37046677],  $inv\_y$ : [0.8 0.8 1.3], Test RMSE: 0.489  
 $n=2$ :  $inv\_yhat$ : [1.32214486 1.41522023],  $inv\_y$ : [0.8 1.3], Test RMSE: 0.378  
 $n=3$ :  $inv\_yhat$ : [1.28334119],  $inv\_y$ : [1.3], Test RMSE: 0.017

If set  $n=4$ , we got “out of range” error, so we use  $n=3$  as our result. The reason of “out of range error” is that some of our patients have only 4 labor results during their ICU stays. If  $n = 4$ , no prediction can be made. We should not delete such patients’ data

because they can give contribution for prediction if we set  $n < 4$ . We train the RNN on whole dataset (i.e., multi-cases prediction). As the result shown, this prototype can predict criteria (SCR) of AKI with a RMSE (Root Mean Square Error) of 0.017mg/dL.

#### 4. Discussion

The prediction of AKI is very challenging; therefore, this study is only trying to make a prototype of a prediction approach, following points can be considered for further improvement of the prediction's precision:

- Prediction should be grouped by demographics to avoid Simpson's paradox. E.g., using data of same age/sex group for prediction is better than mixed data.
- Influence of medications and comorbidities should be considered. Vital signs measured at the bedside such as ECG, SpO<sub>2</sub>, and respiration rate could also be helpful for AKI development prediction. However, because our activation function of neural network is based on Bayesian theorem, the conditional independence of features should be further analyzed and managed before we can use them to minimize overfitting.

#### 5. Conclusion

Because of the complexity of the clinical data, it is difficult to precisely analyze the clinical features manually by using statistical methods such as linear regression. Our purpose of feature engineering is keeping the original information of data and dimension reduction at same time, our algorithm can achieve these goals automatically, it can predict some important laboratory test results of acute kidney injury, such as serum creatinine, with a RMSE (Root Mean Square Error) of 0.017mg/dL. Therefore, our program can make early alert before AKI occurs.

#### References

- [1] KDIGO Clinical Practice Guideline for Acute Kidney Injury. *Kidney International Supplements* (2012) 2, 2; doi:10.1038/kisup.2012.2 <https://kdigo.org/>
- [2] Andrew Ng, Jiquan Ngiam, Chuan Yu Foo, Yifan Mai, Caroline Suen, Adam Coates, Andrew Maas <http://ufldl.stanford.edu/tutorial/supervised/MultiLayerNeuralNetworks/>
- [3] MIMIC-III, Johnson AEW, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, and Mark RG. *Scientific Data* (2016). DOI: 10.1038/sdata.2016.35. Available at: <http://www.nature.com/articles/sdata201635>
- [4] Predictive Modeling of the Risk of Acute Kidney Injury in Critical Care: A Systematic Investigation of The Class Imbalance Problem. Zhenxing Xu, Yujuan Feng, Yun Li, Anand Srivastava, Prakash Adekkanattu, Jessica S. Ancker, Guoqian Jiang, Richard C. Kiefer, Kathleen Lee, Jennifer A. Pacheco
- [5] Sepp Hochreiter; Jürgen Schmidhuber (1997). "Long short-term memory". *Neural Computation*. 9 (8): 1735–1780. doi:10.1162/neco.1997.9.8.1735. PMID 9377276. S2CID 1915014.
- [6] Tomašev, N., Glorot, X., Rae, J.W. et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 572, 116–119 (2019). <https://doi.org/10.1038/s41586-019-1390-1>
- [7] Lachance P, Villeneuve PM, Rewa OG, Wilson FP, Selby NM, Featherstone RM, and Bagshaw SM, "Association between e-alert implementation for detection of acute kidney injury and outcomes: a systematic review," *Nephrology Dialysis Transplantation*, vol. 32, no 2, pp. 265–272, 2017.