Machine Learning Algorithms for Prediction of Acute Kidney Injury in Neurocritical Ill Patients with Acute Brain Injury

Stephanie Wang[1*†], Jan-Yue Lin[1†], Xun Zhe Yan[1†], Ming-Chun Yeh[1]

[1*]National Tsing Hua University, Computer Science, No. 101, Section 2, Kuang-Fu Road, 30013, Hsinchu, Taiwan.

*Corresponding author(s). E-mail(s): steph107062176@gapp.nthu.edu.tw;
Contributing authors: leolin0118@gmail.com; l159753807@gmail.com; emilyyeh0621@gmail.com;
†These authors contributed equally to this work.

Abstract

This study introduces two recurrent neural network (RNN) architectures for predicting Acute Kidney Injury (AKI) in ICU patients using urine output data. The first model, employing a single-layer Long Short-Term Memory (LSTM), utilizes the "Longest Length of Stay" method for input dimension standardization, padded with average urine output values. However, this approach predominantly learns from padding data, leading to skewed accuracy. To address this, we developed a "24Hr" model using a five-layer stacked LSTM (LSTM5) with dropout layers to prevent overfitting. This model adopts a sliding window technique for extracting 24-hour continuous data, predicting AKI status 6 hours later. Despite a modest accuracy, the 24Hr model achieves the highest recall in our tests, demonstrating enhanced capability in handling uneven data distributions typical in critical care settings. This research underscores the importance of model design in predictive accuracy and recall, particularly in imbalanced medical datasets.

Keywords: Recurrent neural network, Long short-term memory, Acute kidney injury, urine output, Machine Learning

1 Introduction

Globally, stroke remains a significant health concern, with an alarming 12.2 million new cases each year, translating to one stroke every three seconds. Notably, the prevalence of stroke has risen by 50% over the past 17 years, with current data indicating that 1 in 4 individuals will experience a stroke in their lifetime. A striking trend has emerged in recent years: strokes are increasingly affecting the younger population. In 2019, 63% of stroke incidents occurred in individuals under 70 years of age, challenging the traditional perception of stroke as an ailment predominantly of the elderly.

The aftermath of stroke affects approximately 101 million people worldwide, leading to a range of debilitating conditions including impaired speech, physical limitations, and partial paralysis. This demographic shift and the growing incidence underline the urgent need for effective predictive and preventive strategies for stroke.

One notable complication following stroke is the development of Acute Kidney Injury (AKI). Our analysis reveals that 27% of stroke patients suffer from AKI, which significantly impacts long-term outcomes. The 10-year mortality rate for stroke patients with AKI stands at 75.9%, compared to 57.7% for those without AKI, a disparity highlighted by a log-rank test of 45.0 ($P = 0.001$). This evidence positions AKI as a potent, independent predictor of both long-term mortality and subsequent cardiovascular events. The relationship between renal function impairment and cardiovascular diseases (CVD), including stroke, is well-documented. This association persists across various stages of renal dysfunction, from predialysis conditions—particularly when compounded by CVD and anemia—to dialysis-dependent patients. Furthermore, preexisting renal dysfunction in acute coronary syndrome or stroke patients correlates with higher risks of mortality and cardiovascular comorbidities. The majority of existing studies have primarily focused on two aspects: first, identifying the factors linking Chronic Kidney Disease (CKD) with CVD, and second, understanding the role of CKD as an independent prognostic factor for both short- and long-term mortality and new cardiovascular incidents.

Given the increasing incidence of AKI in the past five years and its significant relation to stroke outcomes, our prospective study is designed to examine the incidence factors for AKI in patients who have experienced an acute stroke. We aim to provide insights that could be pivotal in the prevention of AKI, thereby contributing to improved stroke management and patient outcomes.

2 Results

In the course of this investigation, the dataset was judiciously partitioned into subsets designated for training, validation, and testing, adhering to a 7:2:1 ratio, respectively. The focal point of our methodology was the 24-hour model architecture, which incorporates LSTM5—a multi-layered LSTM network with five layers—as the primary mechanism for predicting the incidence of Acute Kidney Injury (AKI).

While the accuracy of our model did not surpass that of established studies, the model demonstrated a superior recall rate of 86%, as explicated in Table 1. This high recall rate is indicative of the model's robust capacity to accurately identify AKI instances within datasets characterized by an imbalanced distribution of cases.

As demonstrated in Table 2, we conducted a comparative analysis between the two model architectures discussed in the Methods section. The 24-hour model, which circumvents the use of padding, achieved an accuracy of 86.35%, modestly eclipsing the 84.29% accuracy achieved by the Longest Length of Stay approach. The slight yet significant edge in performance attained by the 24-hour model attests to its capability in handling unmodified data inputs.

Further examination of the impact of different hidden layer configurations, particularly contrasting a multi-layered LSTM against a single-layered approach, revealed a propensity for the former to discern more complex data patterns. This is substantiated by the improved accuracy rates associated with the multi-layer LSTM5 architecture, detailed in Table 3. The advancement in accuracy with the adoption of a deeper neural network underscores the merits of applying more complex machine learning architectures in the field of medical data analysis.

Initially, within the Longest Length of Stay architecture, we utilized a padding value of -1 to address gaps in the time-series urine output data. This method, however, adversely affected the model's predictive accuracy, as detailed in Table 4.

Subsequently, we implemented a revised padding approach, utilizing the average urine output across all patients, thereby mitigating the negative impact of the initial padding strategy. The post-adjustment results, as reflected in Table 4, unequivocally indicated an enhancement in the model's performance.

The progression of our study also entailed an evolution in prediction targets. The initial aim was to predict the potential AKI stage at specific time points, with labels 0, 1, 2, and 3 representing the absence and varying stages of AKI. However, the complexity of the task proved to be too challenging for the 24-hour model's architecture, as illustrated in Table 5.

Therefore, we refined the predictive objective to a binary classification of AKI presence or absence, denoted by 0 and 1. This binary approach simplified the prediction task and, as a result, the model achieved noteworthy performance outcomes, as evidenced by the accuracy and recall rates previously in Table 5.

3 Equations

The Long Short-Term Memory (LSTM) cell is a sophisticated architectural unit within recurrent neural networks, specifically designed to regulate the flow of information. Through the intricate use of various gates and state updates, the LSTM cell adeptly maintains and modifies its internal state over time. This capability is crucial for capturing temporal dependencies and allows for the retention of information across lengthy sequential data streams. The following equations formalize the operations within an LSTM cell:

Input Gate (i_t): Responsible for modulating the influx of new information into the cell state, the input gate's activation is computed as follows:

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \tag{1}$$

where, σ represents the sigmoid function, U^i and W^i are the weight matrices for the input and the recurrent connection, respectively.

Forget Gate (f_t): This gate determines the information to be discarded from the cell state. Its activation is calculated using the current input and the previous output, weighted accordingly:

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \tag{2}$$

Cell State (C_t): The cell state is the memory part of the LSTM cell. It is updated by partially forgetting the previous state and adding new candidate values C_t, which are a function of the current input and the previous output, modulated by the input and forget gate activations.

Candidate Values:

$$\tilde{C}_t = \tanh(x_t U^g + h_{t-1} W^g) \tag{3}$$

Update:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4}$$

Output Gate (o_t): The output gate controls the output flow of cell state information to the rest of the network. Like the other gates, it uses the sigmoid function and has its own weights U^o and W^o.

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \tag{5}$$

Hidden State (h_t): The hidden state is the output of the LSTM cell that gets passed to the next time step. It is calculated by multiplying the output gate activation with the tanh of the updated cell state.

$$h_t = \tanh(C_t) \cdot o_t \tag{6}$$

4 Tables

Table 1: Comparison of Model Performance

Model Architecture	**Accu**	**Recall**
FNN [2]	—	82%
XGB [3]	87.9%	77.8%
Random Forest [3]	87.2%	79.4%
24hr[1]	86.35%	86%

[1]Model Architecture: 24hr. LSTM cell: LSTM5. Label: 0 for not having AKI, 1 for having AKI. Data split into train/validation/test: 0.7/0.2/0.1.

Table 2: The performance of Model Architectures

Model Architecture	Accu	Accu (padding eliminated)	Accu (last 5 hours)
LLS[1]	99.19%	84.29%	83.07%
24hr[2]	—	86.35%	—

[1]LLS: Longest Length of Stay. LSTM cell: LSTM5. Labels: 0, 1, 2, 3, -1.

[2]24hr: Sliding windows method.

Table 3: The performance of LSTM Architectures for Longest Length of Stay

LSTM cells	Accu	Accu (padding eliminated)	Accu (last 5 hours)
LSTM1[1]	99.03%	80.74%	79.82%
LSTM5[2]	99.10%	82.13%	81.43%

[1]LSTM1: stack 1 simple LSTM cell.

[2]LSTM5: stack 5 simple LSTM cells with DropOut() between.

Table 4: Accuracy of Models with Different Padding Methods

Padding method[1]	Accu	Accu (padding eliminated)	Accu (last 5 hours)
Pmean[1]	99.03%	80.74%	79.82%
P-1[2]	98.87%	76.08%	74.87%

[1]Pmean: Padding method where Umn = Mean of existing values.

[2]P-1: Padding method where Umn = -1.

Table 5: Accuracy of the 24hr Model with Different Label Methods

Label method	Accu
0, 1, 2, 3	70.79%
0, 1[1]	86.35%

[1]In the label method 0, 1: 0 represents having AKI, 1 represents not having AKI, and -1 is used for padding.

5 Figures

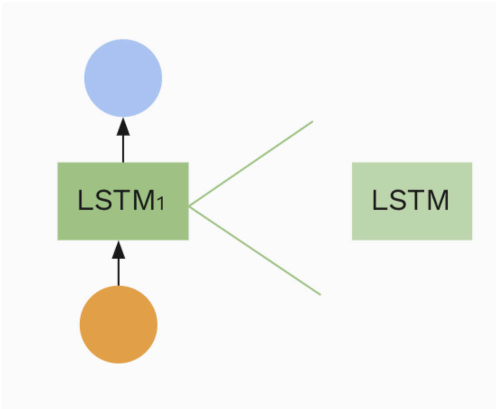

Fig. 1: LSTM1: A single LSTM cell

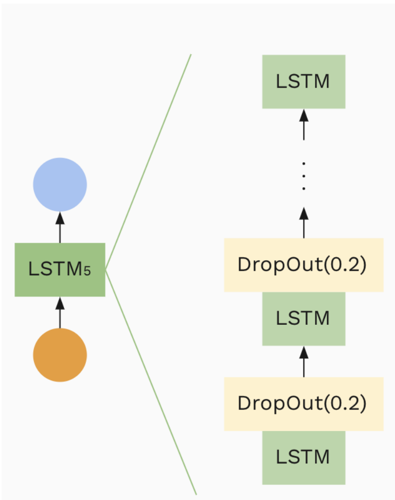

Fig. 2: LSTM5: Stack 5 LSTM cells, with DropOut() between each of them

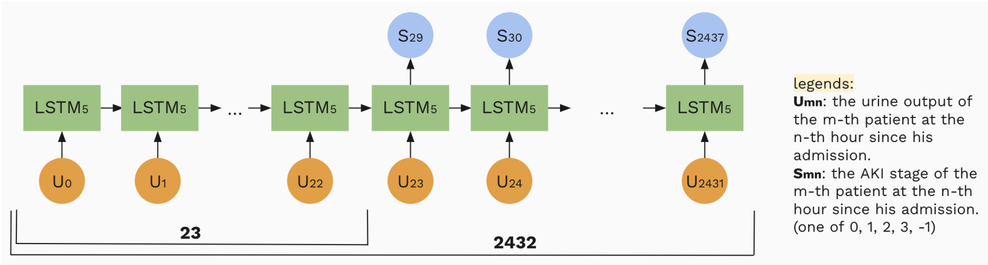

Fig. 3: Longest Length of Stay

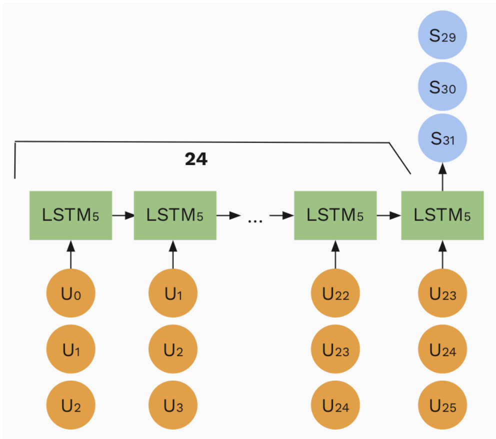

Fig. 4: 24Hrs Sliding Window

6 Methods

In this study, we implemented two distinct Recurrent Neural Network (RNN) architectures, each characterized by a unique feature extraction module. The primary architecture consists of a single-layer Long Short-Term Memory (LSTM) network, while the secondary model employs a more complex five-layer LSTM with integrated dropout layers to mitigate overfitting.

6.1 Data Prepossessing

For this study, we focused on the urine output data of patients aged 18 and older who were admitted to the ICU for more than 30 hours, as recorded in the MIMIC-IV dataset. The dataset comprised urine output measurements from 29,000 patients, of which 18,000 had developed Acute Kidney Injury (AKI) and the remaining 11,000 had not. This dataset also included instances of the same patients being admitted to the ICU multiple times.

A challenge encountered in the data was the irregularity of urine output collection intervals. Although urine output is expected to be recorded at one-hour intervals, the actual data collection exhibited inconsistencies, with some intervals extending beyond an hour. This irregularity necessitated the implementation of a padding method to address missing data. The padding process involved calculating the duration of missing data and then evenly distributing the recorded urine volume across the missing time slots. For instance, if a urine output of 150 is recorded at 16:00, with the previous measurement at 12:00, the missing four-hour interval is filled by allocating an average urine output of 37.5 (150 divided by 4) for each hour from 13:00 to 16:00. Additionally, the dataset's charttime variable required transformation to align with the study's analytical framework. Given the variability in each patient's admission time, it was essential to standardize the time sequence of their ICU stay. To facilitate the model's learning process, which involves predicting the future probability of AKI based on temporal changes in urine volume, we converted each patient's charttime into a sequential time series starting from 0, using one hour as the time unit. This approach allows for a consistent temporal framework, accommodating the differing lengths of ICU stays across patients.

6.2 Longest Length of Stay

Our first approach utilizes a single-layer LSTM to predict the likelihood of Acute Kidney Injury (AKI) based on urine output at individual time points for each patient. Recognizing the variability in ICU stay durations, we adopted the "Longest Length of Stay" method as shown in Figure 3 and Figure 1. This approach identifies the patient with the longest ICU stay and uses their stay duration as a benchmark for the input dimension of our model. This method facilitates the prediction of AKI status six hours subsequent to each recorded time point. For patients with shorter urine output records, we implemented an average padding technique, extending their data to match the input dimension. In these instances, the model output is represented by -1 to indicate padding at the respective time points. However, this method predominantly trains the model on padded data, leading to a skewed accuracy favoring the output of -1.

6.3 24Hr Sliding Window Model

To address the limitations of the first method, we introduced an alternative algorithm, the 24Hr model. This model abandons the patient-centric approach and instead utilizes a sliding window technique to extract continuous 24-hour urine output data, as depicted in Figure 4. The model then predicts the likelihood of AKI occurring six hours later. The dataset is divided using a 7:2:1 ratio for training, validation, and testing, respectively. The 24Hr model architecture incorporates LSTM5, a five-layer stacked LSTM model shown in Figure 2, as its core component. While the accuracy of our models does not surpass that of existing research, we observed a significantly higher recall rate. This indicates an enhanced capability of our models in identifying AKI cases within unevenly distributed datasets, a crucial advantage in the context of acute medical conditions such as AKI in ICU settings.

7 Discussion

In our analysis, we found that while the Longest Length of Stay padding method was initially utilized, it may not be optimal due to the introduction of bias from outlier data representing extended ICU stays. The study revealed a potential area for improvement by suggesting a padding threshold that reflects the average duration of ICU stays. Furthermore, the practice of uniform padding of urine output data was reevaluated in light of clinical insights regarding the correlation between urine output and body weight, indicating a need for a more individualized padding approach.

The prediction bias towards non-AKI and stage 3 AKI due to imbalanced data distribution was another significant observation. Our findings suggest that a more balanced representation of AKI stages could improve model performance, particularly for those stages not typically predicted with high accuracy.

8 Conclusion

Our research indicates that careful consideration of data preprocessing methods is vital in developing robust predictive models for AKI in ICU settings. Adjusting padding strategies to avoid overrepresentation of outlier data and incorporating patient-specific information such as body weight could significantly enhance model accuracy. Equally important is the need to address data imbalances to ensure uniformly high predictive performance across all AKI stages. While our model showed promise with a high recall rate, future work should aim to refine these aspects of model training to further improve the accuracy and reliability of AKI predictions, ultimately contributing to better patient outcomes in critical care.

9 Author Contribution

- Stephanie Wang - 30%
- Jan-Yue Lin - 30%
- Xun Zhe Yan - 30%
- Ming-Chun Yeh - 10%

References

[1] Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., & Mark, R. (2022). MIMIC-IV (version 2.1).

[2] Alfieri F, Ancona A, Tripepi G, Crosetto D, Randazzo V, Paviglianiti A, Pasero E, Vecchi L, Cauda V, Fagugli RM. A deep-learning model to continuously predict severe acute kidney injury based on urine output changes in critically ill patients. *J Nephrol.* 2021 Dec;34(6):1875-1886. doi: 10.1007/s40620-021-01046-6. Epub 2021 Apr 26. PMID: 33900581; PMCID: PMC8610952.

[3] Zhang X, Chen S, Lai K, Chen Z, Wan J, Xu Y. Machine learning for the prediction of acute kidney injury in critical care patients with acute cerebrovascular

disease. *Ren Fail.* 2022 Dec;44(1):43-53. doi: 10.1080/0886022X.2022.2036619. PMID: 35166177; PMCID: PMC8856083.

[4] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.