The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200528

Using Long Short-Term Memory (LSTM) Neural Networks to Predict Emergency Department Wait Time

Nok CHENG [a,1] and Alex KUO [a]
[a] *School of Health Information Science, University of Victoria, Canada*

Abstract. Emergency Department (ED) overcrowding is a major global healthcare issue. Many research studies have been conducted to predict ED wait time using various machine learning prediction models to enhance patient experience and improve care efficiency and resource allocation. In this paper, we used Long Short-Term Memory (LSTM) recurrent neural networks to build a model to predict ED wait time in the next 2 hours using a randomly generated patient timestamp dataset of a typical patient hospital journey. Compared with Linear Regression model, the average mean absolute error for the LSTM model is decreased by 9.7% (3 minutes) ($p < 0.01$). The LSTM model statistically outperforms the LR model, however, both models could be practically useful in ED wait time prediction.

Keywords. ED overcrowding, ED wait time prediction, recurrent neural network, Long Short-Term Memory (LSTM)

1. Introduction

Emergency Department (ED) overcrowding is a complex, systemic and global healthcare issue resulting in decreased care efficiency and worsening patient outcomes [1]. Due to the aging population and the lack of inpatient capacity, admitted patients occupying ED beds create an access block for incoming emergency patients [1,2]. Hence, ED overcrowding is a problem that stems from the system as a whole rather than solely the ED [1,2].

ED wait time, defined as the time from patient arrival to physician initial assessment, can support decision making to enhance patient experience and improve hospital resource allocation [1]. Machine learning prediction models have been used to predict ED length of stay and wait time based on the patient flow in the ED [3,4]. Very limited research studies explore the use of *Long Short-Term Memory* (LSTM) *Recurrent Neural Networks* [5] in Emergency Medicine research, while LSTM has been commonly used in fields such as Business consistently providing more accurate sequence predictions [6]. Compared to the LSTM, *Linear Regression* (LR) [7] is a more popular method for prediction due to its easy implementation and less processing time, however, such a method may not offer accurate prediction because some complex real-world problems such as ED overcrowding may not be explainable with a linear approach.

[1] Corresponding Author, Nok Cheng, School of Health Information Science, University of Victoria, HSD Building, A202, Victoria, B.C., Canada; E-mail: nokcheng@uvic.ca.

Therefore, the aim of this paper is to preliminarily investigate the application of the LSTM method in ED overcrowding analysis by comparing the accuracy of ED wait time prediction of an LSTM machine learning model and that of an LR model using a randomly generated patient timestamp dataset.

2. Methods

A dataset of randomly generated patient instances was created. Two models utilizing LR and LSTM methods, were created, respectively. Using K-Fold Cross Validation [8], a training-validating process followed by a training-testing process were performed to each model to predict the ED wait time. The mean absolute error of the ED wait time predicted by the model is studied by comparing to the true ED wait time. The performance of the two models are evaluated.

2.1. Patient Timestamp Dataset

A dataset was created to represent 50,000 patient instances of timestamp differences randomly generated over a pre-defined range uniformly. The use of the uniform distribution enables maximum entropy probability to ensure the consideration of all possible circumstances given the lack of clinical data [9]. A sequence of five (5) timestamp differences were used to illustrate the six (6) timestamps in a patient journey, including: patient arrival time, triage time, registered time, physician initial assessment time, ED exit time, and inpatient unit discharge time. At a specific time t, each of the five sequential timestamp differences is labelled as x_t. The ED wait time for each patient instance was calculated in minutes by adding the first 3 timestamp differences. Each patient instance contains an arrival timestamp including the month, week, day, hour and minute that was randomly generated using uniform distribution. The average ED wait time for the next 2 hours labelled as y was calculated for each patient instance by taking the average of the ED wait time of the patients arriving in the next 2 hours. The dataset was then scaled using MinMaxScaler of the Scikit-learn 0.20.1 library [10].

2.2. Linear Regression Model

Python 3.7.1 and Scikit-learn 0.20.1 were used to build an LR model with an input array of 1 x 5 for the 5 timestamp differences and an output array of 1 x 1.

2.3. Long Short-Term Memory Model

Long Short-Term Memory (LSTM) is a type of recurrent neural network which when unfolded, contains individual identical units [5]. Each unit receives input data at each time-step of an instance in the dataset [5]. LSTM can learn long-term dependencies without the vanishing gradient problem due to its sigmoid activation function in each unit which produces an output of hidden state, h_t and cell state, c_t. Both of them depend on the input, x at time t, the previous cell state, c_{t-1}, and the previous hidden state, h_{t-1} [5]. Using Python 3.7.1 and PyTorch 1.1.0 [11], a single-layered LSTM model using 5 input units with each unit receiving an element of the patient instance sequence, x_t, as input was built. The output from the final LSTM unit passes through an LR layer with a

1 x 1 input array, which produces a 1 x 1 predicted output, $\hat{y}$, of ED wait time in the next 2 hours as illustrated in Figure 1.

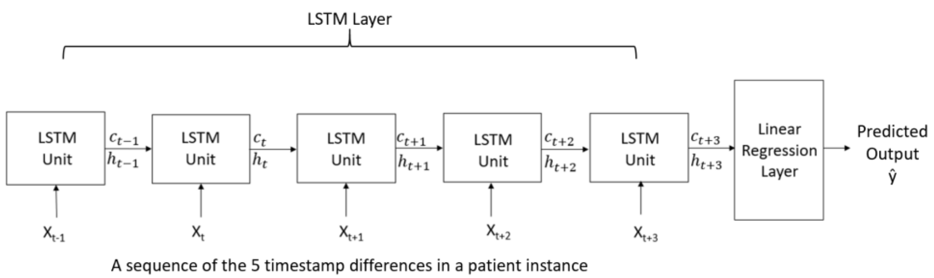

Figure 1. Single-layered LSTM model using 5 input units with a linear regression layer.

2.4. Training and Evaluation

For both the LR and LSTM model, 80% of the data was first used to train and validate the model using 100-Fold Cross-Validation for 500 epochs to observe the behavior of the hyper-parameters of the models and ensure no overfitting occurs prior to the exposure of the testing data [8]. Then, 100% of the data was used to train and test both models with 100-Fold Cross-Validation for 500 epochs. Both models include an early stopping to prevent overtraining and use mean absolute error (MAE) as loss function. The LR model was optimized using Stochastic Gradient Descent [12], while the LSTM model used ADMA, Adaptive Moment Estimation, optimization [13]. A two-sample t-test was performed on the MAE values from the testing phase in both models with a significant difference of $p < 0.01$ and 99% confidence level. The lowest MAE value for each of the models and their overall training-testing time were observed during the process.

3. Results

As shown in Table 1, the LSTM model achieved a 9.7% (3 minutes) lower average MAE (p < 0.01), 7.49×10^{-2} (28 minutes) compared to 8.84×10^{-2} (31 minutes) in the LR model. The minimum MAE in the LSTM model is 12% lower than in the LR model. The training-testing time of the LSTM model is 2.5 hours compared to 2.17 seconds for the LR model.

Table 1. The training-testing time (TTT), Average MAE, standard deviation (SD), 99% confidence interval (99% CI), minimum MAE (Min MAE) and p-value of the LR and LSTM in the testing phase are shown below.

Model	TTT	Average MAE (minutes) $\pm$ SD (minutes)	99% CI (minutes)	Min MAE (minutes)
LR	2.17s	8.84×10^{-2} (31) $\pm 6.10 \times 10^{-3}$ (1.41)	$(7.27 - 10.41) \times 10^{-2}$ (27.79 – 35.05)	7.25×10^{-2} (27.75)
LSTM	**2.5hr**	$\mathbf{7.49 \times 10^{-2}}$ **(28)** $\pm \mathbf{7.43 \times 10^{-3} (1.72)}$	$\mathbf{(5.57 - 9.40) \times 10^{-2}}$ **(23.87 – 32.71)**	$\mathbf{5.82 \times 10^{-2}}$ **(24.44)**
p-value	$\mathbf{1.54 \times 10^{-31}}$			

4. Discussion and Conclusions

Supported by Tax et al. 2017 [6], the LSTM model statistically outperforms the LR model due to its ability to learn the relationships between consecutive time-steps. The relatively small 3-minute prediction difference between the 2 models suggests that both models could be practically useful in wait time prediction. An important aspect of this study is the use of a complex neural network model in the ED setting to construct a wait time prediction tool. This tool serves for clinical decision support to provide patient control in ED care [14], and improves care efficiency with proactive resource allocation such as increasing the number of nurses and overflow units [1].

A major limitation of the study is the use of a randomly generated dataset without real patient flow and data quality representation. Since ED overcrowding is a system issue involving the entire hospital [1,2], it is crucial to use data that extends beyond the ED. The validity and reliability of the data may vary at hospitals which may impact the final model and its accuracy. However, incomplete patient data may offer new understanding of ED wait time. As next steps, we plan to use hospital patient timestamps as well as clinical data and explore different model architectures and accuracy evaluations.

In summary, this paper utilized the LSTM neural networks to build a model to predict the average ED wait time in the next 2 hours. Though the LSTM model statistically outperforms the LR model, both models could be practically useful in wait time prediction. Future work will include broader patient data from a hospital and variations of the model architecture.

References

[1] Affleck A, Parks P, Drummond A, Rowe BH, Ovens HJ, Emergency department overcrowding and access block, Canadian Journal of Emergency Medicine 15 (2013), 359-370.

[2] Innes GD et al., Emergency overcrowding and access block: A smaller problem than we think, Canadian Journal of Emergency Medicine 2 (2019), 177-185.

[3] Gill SD, Lane SE, Sheridan M, Ellis E, Smith D, Stella J, Why do 'fast track' patients stay more than four hours in the emergency department? An investigation of factors that predict length of stay, Emergency Medicine Australasia 30 (2018), 641-647.

[4] Gonçalves F, Pereira R, Ferreira J, Vasconcelos JB, Melo F, Velez I, Predictive Analysis in Healthcare: Emergency wait time prediction, Advances in Intelligent Systems and Computing 806 (2018), 138-145.

[5] Hochreiter S et al., Long short-term memory, Neural computation 9 (1997): 1735-1780.

[6] Tax N, Verenich I, La Rosa M, Dumas M, Predictive Business Process Monitoring with LSTM Neural Networks, Lecture Notes in Computer Science (LNCS) 10253 (2017), 477- 492.

[7] Grömping U, Variable Importance Assessment in Regression: Linear Regression versus Random Forest, The American Statistician 63 (2009), 308–319.

[8] Hastie T, Tibshirani R, Friedman J, The Elements of Statistical Learning: Data mining, inference, and prediction, 2nd ed., Springer Science & Business Media, New York, 2009.

[9] Trębicki J, Sobczyk K, Maximum entropy principle and non-stationary distributions of stochastic systems, Probabilistic Engineering Mechanics 11 (1996), 169-178.

[10] Pedregos F et al., Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011), 2825–2830.

[11] Ketkar N, Deep Learning with PyTorch, Apress, California, 2017.

[12] Ruder S, An overview of gradient descent optimization algorithms, arXiv (2014), 1-12.

[13] Kingma DP et al., Adam: A method for stochastic optimization, arXiv (2015), 1-15.

[14] Navarin N, Vincenzi B, Polato M, Sperduti A, LSTM Networks for data-aware remaining time prediction of business process instances, In 2017 IEEE Symposium Series on Computational Intelligence (2017), 1-7.