BIOCYBERNETICS AND BIOMEDICAL ENGINEERING 42 (2022) 1051–1065

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Predicting hospital emergency department visits with deep learning approaches

Xinxing Zhao [a], Joel Weijia Lai [a], Andrew Fu Wah Ho [b,c], Nan Liu [d], Marcus Eng Hock Ong [c,d], Kang Hao Cheong [a,*]

[a] Science, Mathematics and Technology Cluster, Singapore University of Technology and Design, Singapore
[b] Pre-hospital & Emergency Research Centre, Duke-National University of Singapore Medical School, Singapore
[c] Department of Emergency Medicine, Singapore General Hospital, Singapore
[d] Health Services and Systems Research, Duke-NUS Medical School, Singapore

ARTICLE INFO

Article history:
Received 27 January 2022
Received in revised form
23 June 2022
Accepted 20 July 2022
Available online 2 August 2022

Keywords:
Deep learning
Deep stacked architecture
ED Prediction
LSTM
RNN
GRU

ABSTRACT

Overcrowding in emergency department (ED) causes lengthy waiting times, reduces adequate emergency care and increases rate of mortality. Accurate prediction of daily ED visits and allocating resources in advance is one of the solutions to ED overcrowding problem. In this paper, a deep stacked architecture is being proposed and applied to the daily ED visits prediction problem with deep components such as Long Short Term Memory (LSTM), Gated Recurrent Units (GRU) and simple Recurrent Neural Network (RNN). The proposed architecture achieves very high mean accuracy level (94.28–94.59%) in daily ED visits predictions. We have also compared the performance of this architecture with non-stacked deep models and traditional prediction models. The results indicate that deep stacked models outperform (4–7%) the traditional prediction models and other non-stacked deep learning models (1–2%) in our prediction tasks. The application of deep neural network in ED visits prediction is novel as this is one of the first studies to apply a deep stacked architecture in this field. Importantly, our models have achieved better prediction accuracy (in one case comparable) than the state-of-the-art in the literature.

1. Introduction

The emergency departments (ED) overcrowding is a global healthcare challenge [1–5]. It is one of the major hindrances that stops patients from receiving timely and adequate emergency care, as there are generally limited resources in hospitals, such as manpower (medical, nursing and other professionals), facilities and equipment, and the coordination among human and material elements to cope with surge in the patient visits. Overcrowding causes lengthy waiting times [6,7], which in turn causes reduced quality of care [8,9], patients leaving the ED without receiving treatment [10] and

* Corresponding author at: Science, Mathematics and Technology Cluster, Singapore University of Technology and Design, 8 Somapah Road, 487372, Singapore
E-mail address: kanghao_cheong@sutd.edu.sg (K.H. Cheong).
https://doi.org/10.1016/j.bbe.2022.07.008

increased levels of mortality [11,12]. It also adversely impacts both patients and medical staff emotionally, such as dissatisfaction [13] with prolonged waiting time and decreased morale and productivity [14]. To shorten the waiting time, accurate forecasting of ED demand such as the number of patients that may be visiting [15,16] or estimating future bed demand [17] will allow hospitals to plan resource allocation in advance (eg. manpower scheduling), improve operational effectiveness, and can potentially alleviate overcrowding problems effectively.

Deep learning models have achieved excellent results in many fields, such as detection of abnormalities in medical images [18–20], disease detection [21,22], classification of network traffics and attacks [23,24], recognition of movement and activity [25,26], object detection [27], future trend prediction [28,29], even for mental states analysis [30]. Taking Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) for example: CNN models are popularly used for image processing while RNN models such as Long Short Term Memory (LSTM) and Gated Recurrent Units (GRU) are used more often for time series data predictions. For instance, Kassania et al. [31] utilized chest X-ray and CT images with CNN for automatic COVID-19 detection with a classification accuracy of 99%, while [32] Mishra used CT images and CNN models with transfer learning achieved similar performance for COVID-19 and normal classification. Hashemzehi et al. [33] used hybrid deep CNN for automated brain tumor detection based on MRI images, achieved an accuracy of 95%. Hiransha et al. used RNN and LSTM and other deep learning models for stock market prediction and achieved an accuracy about 94% for 400 days in advance. Karevan et al. used LSTM for weather forecasting [34]for 6 days ahead prediction and achieved good results in terms of mean absolute error and mean squared error. Wang et al. [82] used LSTM for COVID-19 trends prediction with 150-days ahead in Russia, Peru and Iran, and the predictions are largely consistent with the reported cases in different countries. At the current stage, hospitals in general have not adopted deep learning approaches in large scale. However, emergency departments may stand in a unique position to benefit from deep learning. As many machine learning models and deep leaning models have already shown the abilities and potential [35,36] in improving patient care [37,38], in analysing [39,40]. They can also help in predicting during triage stage [41], in helping physicians to balance the risk and decision making with limited information [42], in determining and predicting urgent patient outcomes[43,44], and in predicting revisits [39,45] within a certain time period in the EDs. More deep learning models can be used in more ED scenarios and improve the results of current researches. Deep learning can also be used for more optimised resource allocation, to help emergency departments, or more generally hospitals, in forming their strategic, tactical and operational plan with their different level of decision-making.

In this study, we utilise a deep stacked architecture which has a variable number of layers (2–4) with LSTM, GRU and SimpleRNN models as its components, to predict daily ED visits in Singapore General Hospital, the largest and oldest public hospital in the city-state of Singapore. The used dataset is a five-year historical record with date and daily ED visits information. We have compared deep stacked models with other non-stacked deep learning models such as convolutional neural networks (CNN), Bidirectional LSTM (BiLSTM), Convolutional LSTM Network (ConvLSTM), and with the auto regressive integrated moving average (ARIMA) model, one of the most popular traditional statistical methods, as well as with Facebook Prophet model. The Prophet [46] model bases on an additive model, and it can model non-linear trends with yearly, weekly, daily seasonality, and holiday effects well. The prediction results indicate that our deep stacked models are more accurate than these existing methods (for all 4 prediction tasks) in predicting daily ED visits.

The main contributions of this paper are, first, using deep learning models to predict daily ED visits is new in the field. Our work here is one of the first studies to use a deep stacked architecture with RNN and its variants as deep components in the ED visits prediction. Our model has achieved a very high prediction accuracy—better, or have similar performance over the state-of-the-art in the literature. Second, our study confirms that deep learning models have better performance than traditional models such as ARIMA and Prophet. Third, our study has provided some research insights into achieving good predictive results in such problem: neither the number of hidden layers nor the number of hidden units needs to be very large in the deep stacked architecture. Fourth, our study has demonstrated that deep stacked models can achieve very high prediction accuracy (94%) in daily ED visits prediction with data over a period of 6 months. Note that although we only considered the prediction in one country, the models we provided and the insights we obtained from this study are generally applicable across other EDs in different regions and countries. The rest of this paper is organized as follows. Section 2 discusses the related work. The details of the dataset are given in Section 3. The proposed architecture, as well as the evaluation metrics and configurations of the deep models are also discussed in Section 3. Section 4 and Section 5 provides the analysis and discussion of our results. Finally, we give our conclusion in Section 6.

2. Related Work

The ED visits prediction, such as predicting daily number of patient visits, has long been regarded as one of the key strategies in mitigating ED crowding problem. Over the past few decades, many ED prediction models have been developed to improve hospital management efficiency. These developed approaches can be grouped into two categories, i.e., parametric and non-parametric models. Parametric models include such as linear regression [15], exponential smoothing (ES) [47] and ARIMA [48] models. Nonparametric models range from artificial neural networks (ANNs) [49], support vector machine (SVM) [50] to deep learning models such as LSTM [51].

Parametric models (or the traditional models) are relatively simple and easy to implement. Table 1 gives some of the studies in the recent years for ED patients forecasting. Kam et al. [52] used moving average, seasonal autoregressive integrated moving average (SARIMA) and multivariate SARIMA to forecast ED daily visits, achieved best

Table 2 – Some of recent studies involving ED patient arrivals forecasting with traditional machine learning methods.

Refernce	Data& Predicting variables	Models & Horizon of forecasting	Performance
[59] (2013)	1 year; Calendar, holiday, climate, influenza level	ANN; Daily	MAPE of 5.10–14.97%
[60] (2014)	2.75 years; Weather, days of week, air quality, and special events	ANN; Daily	R^2 of 0.957
[61] (2015)	8 years; Calendar, holiday and week number	SVR, M5P; Weekly	—
[62] (2016)	2 years; Calendar, holiday, temperature	ARIMA-ANN, LR; Daily	MAPE of 7.2–12.5%
[57] (2019)	7 years; Calendar, holiday, temperature, precipitation	MLP, SARIMA, SARIMAX, Regression; Daily	MAPE of 8.4%
[63] (2019)	7 years, week data	EEMD-ANN, DWT-ANN and ANN; Weekly	RMSE of 52.86; MAE of 39.88
[64] (2019)	4 years; Lagged volumes and corresponding time factors	ARIMA-SVR; Daily	MAPE of 7.02%
[51] (2021)	3.5 years; Hour, holiday and climate	RF; Daily and Weekly	MAPE of 9.81% for 3 days and 10.61% for 7 days

mean absolute percentage error (MAPE) of 7.4% with calendar and weather variables while Boyle et al.[53] developed ES, ARIMA and regression models for predicting hourly, daily and monthly, with best MAPE about 2% for forecasting monthly admissions and 7% for daily ED presentation. Marcilio et al. [54] developed generalized linear models (GLM), generalized estimating equations (GEE), and SARIMA models to forecast daily number of patients in an ED. They achieved high accuracy for 7-day (best MAPE 4.5%) and 30-day forecasting (best MAPE 8.7%). They reported that ambient temperature readings included in the models did not improve forecasting accuracy. Calegari et al. [55] tested simple seasonal exponential smoothing (SS) model, seasonal multiplicative Holt-Winters (SMHW) model, SARIMA model, and multivariate ARIMA (MSARIMA) model and Whitt et al. [57] compared SARIMA, SARIMA with exogenous regressors (SARIMAX) models in forecasting daily visits of patients in EDs. The performance depends on many factors, such as the models, the data length, weather and holidays information.

As the nature of ED flow is stochastic and non-linear, nonparametric methods have gained momentum in the last decade. Researchers have developed and adopted many machine learning models in the ED visits forecasting field. For instance, Xu et al. [59] modeled daily patient arrivals at an ED with ANN and tried to quantify the importance of different contributing variables. Menke et al. [60] also used an ANN model to predict daily ED visits, whereas Whitt et al. [57] used MLP to achieve MAPE of 8.4% in ED visits prediction. Other machine learning models such as SVR[61,64], M5P[61] and Random Forest (RF) [51] have also been used for ED visits prediction. Machine learning and traditional statistical models have their own advantages and disadvantages. Some studies have demonstrated hybrid models [62–64] can achieve good prediction accuracy, outperforming each single method in either category. Table 2 summarises some of the recent ED visits prediction studies with machine learning models. As we can see, for the daily predictions, the accuracy in terms of MAPE are in the range of 5.10 %- 14.97%. Xu et al. [59] achieved the best MAPE (5.10 %) with ANN models, however, their research only has one year data, and used 90% of them to build the models, their overfitting problem could have been very serious.

Inspired by notable achievements in different domains such as image classification [70], natural language processing [71] and reinforcement learning [72], deep learning based models have attracted widespread attention of the scientific community due to their novel properties and technological applications. CNNs and RNNs (and their variants) have been used [51], modified [73] and adapted [74] for many time series related forecasting problems. For instance, in [51], Sudarshan et al. used LSTM, CNN prediction models with meteorological and calendar parameters. They reported that, CNN outperformed (MAPE of 9.24%) in patient arrivals prediction for 3-day ahead and LSTM performed better (MAPE 8.91%) for 7-day ahead with weather information; LSTM model (MAPE 8.04%) outperformed CNN (MAPE 9.53%) models with past weather information. Harrou et al. [66] developed a variational autoencoder (VAE) model to forecast daily and hourly visits at a ED with two-year data. RNN, LSTM, Bidirectional LSTM (BiLSTM), Convolutional LSTM Network (ConvLSTM) were also used for the forecasting in the same scenario. These models achieved high performance in terms of coefficient of determination (R^2) and Root Mean Square Error (RMSE). Kadri et al. [67] used only one LSTM layer for forecasting daily and hourly admissions and achieved good performance (R^2: 0.972) for a horizon of 78 days forecasting with two-year dataset. Table 3 summarises some of recent studies for ED visits prediction with deep learning models.

One interesting trend that we notice is that, the forecasting horizons are usually daily, weekly or even monthly. Hourly predictions are less frequently studied. This is because hourly predictions usually return poor prediction performance. For example, with more traditional methods (regression, ARIMA, exponential smoothing), the best MAPE was approximately 50% [53,56] for hourly predictions, similarly, for more advanced machine learning methods [65] such as SVR, KNN,

Table 2 – Some of recent studies involving ED patient arrivals forecasting with traditional machine learning methods.

Refernce	Data& Predicting variables	Models & Horizon of forecasting	Performance
[59] (2013)	1 year; Calendar, holiday, climate, influenza level	ANN; Daily	MAPE of 5.10–14.97%
[60] (2014)	2.75 years; Weather, days of week, air quality, and special events	ANN; Daily	R^2 of 0.957
[61] (2015)	8 years; Calendar, holiday and week number	SVR, M5P; Weekly	—
[62] (2016)	2 years; Calendar, holiday, temperature	ARIMA-ANN, LR; Daily	MAPE of 7.2–12.5%
[57] (2019)	7 years; Calendar, holiday, temperature, precipitation	MLP, SARIMA, SARIMAX, Regression; Daily	MAPE of 8.4%
[63] (2019)	7 years, week data	EEMD-ANN, DWT-ANN and ANN; Weekly	RMSE of 52.86; MAE of 39.88
[64] (2019)	4 years; Lagged volumes and corresponding time factors	ARIMA-SVR; Daily	MAPE of 7.02%
[51] (2021)	3.5 years; Hour, holiday and climate	RF; Daily and Weekly	MAPE of 9.81% for 3 days and 10.61% for 7 days

LSTM, the reported performances were also poor (MAPEs larger than 50%).

3. Methodology

In this section, we first present the dataset that has been collected and then provide details about the proposed architecture for our purpose here—predicting the daily ED visits. We then give a short introduction to the evaluation metrics that used to compare the prediction performance between the different models. Finally, we present the configuration details for these deep models.

3.1. The Dataset

The dataset used in this study was collected at Singapore General Hospital (SGH), one of the largest and oldest hospital in the city-state of Singapore. The collected period spans five years from January 2015 to December 2019, totalling 1820 consecutive days of ED daily patient numbers. The collected data include the date, the day of the week and the number of ED visits on that day, in all, there were a total of 648,803 ED visits (mean 356.48, std 36.73, min 174, 25th percentile 331, 50th percentile 352, 75th percentile 378, max 516). Fig. 1 gives the general daily visits pattern of the ED (presented at the ED regardless of urgency of medical conditions) for the 5 years. Fig. 2 gives the frequency histogram and the pattern by the day of the week of the ED visits. In this study, the 5-year dataset will be segmented into several slices with different length, each slice will be further divided into a training set and a testing set with a ratio of 7:3 (in chronological order, with no data appearing in both of the training and test sets), similar to other deep learning based forecasting studies [66,75].

3.2. The Proposed Architecture

Studies have shown that, the overall performance of deep learning models can be improved by increasing the depth of a neural network [76,77]. Here, we propose a deep stacked architecture with LSTM, RNN [78]and GRU[79] models as components to predict daily ED visits, where the depth of deep learning networks can be easily adjusted. We compared the performance of the architecture with CNN [51], BiLSTM and ConvLSTM [66]) and more traditional models (ARIMA and Prophet). For more details about these deep learning models and more traditional models can be found in the referenced papers. In our proposed architecture for the prediction of daily ED visits (as shown in Fig. 3), the input X_t reaches the first deep block at time t, the new hidden state $h_t^{(1)}$ at time t at the first deep block will be calculated. The new hidden state then goes forward to the second deep block and gets updated at deep block 2. This process repeats itself until the last deep block is compiled. With this architecture we achieved excellent predictive capability which we will elaborate over the next few sections.

3.3. Evaluation Metrics of Models

To assess the performance of our forecasting for the ED flow, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) denoted in Eqs. (1)–(3), have been adopted to quantify the forecasting models. Here, n corresponds to the number of data, $y_t$is the actual observed value of ED number at time t, and $\hat{y}_t$ corresponds to the predicted ED flow number at time t.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}, \tag{1}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}|(y_t - \hat{y}_t)|, \tag{2}$$

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n}\frac{|(y_t - \hat{y}_t)|}{y_t}. \tag{3}$$

3.4. Models and Configurations

The performance of deep learning models depends not only on the quality and the quantity of training dataset but also on the configuration of models such as the time steps and optimizers. For the structure shown in Fig. 3, we can adjust

Table 3 – Some of recent studies involving ED patient arrivals forecasting with deep learning methods.

Reference	Data& Predicting variables	Models & Horizon of forecasting	No. of Layers & Hidden units	Performance
This study (2022)	5 years, 3 years, 1 year, 6 months; Calendar and meteorological	Deep stacked architecture with RNN, LSTM, GRU; ConvLSTM, BiLSTM, CNN, ARIMA and Prophet; Daily	RNN layers [3,4], units 50; LSTM layer [2,3], units [200,300]; GRU layers 3, units [200,300]; BiLSTM layer 1,units 64; ConvLSTM layer1, units 64	MAPE of 5.41–6.12% using deep stacked models; 5.79–7.74% non-stacked deep models; 9.15–12.66% for more traditional models
[65](2022)	396 days; Meteorological and calendar and constructed features	KNN, LSTM and SVR and others; Daily and Hourly	- -	MAPE of 40–70% hourly prediction, daily 9–13%
[51](2021)	3.5 years; Meteorological and calendar	CNN, LSTM and RF; Daily and Weekly	CNN layers 32; LSTM layer 1, units 75	MAPE of LSTM 9.31% (3 days), 8.91% (7 days); CNN 9.24% (3 days), 10.69% (7 days)
[66] (2020)	3.0 years; Day of week, month of year	RNN, LSTM, BiLSTM, ConvLSTM, RBM, CNN, GRU, VAE; Hourly and Daily	VAE layers 5, units 32; CNN layers 3; RBM layers 2, units 16; GRU units 32; ConvLSTM units 32; BiLSTM units 32; LSTM units 32	R^2 of LSTM 0.905; GRU 0.907; RNN 0.757; BiLSTM 0.916; ConvLSTM 0.812; RBM 0.883; VAE 0.925; CNN 0.909
[67] (2020)	2 years; Calendar	LSTM; Daily	LSTM layer 1 units 50	R^2 of 0.972
[68](2019)	2.5 years;–	LSTM; Monthly	LSTM layers 2, units 4	MAPE of monthly 4.71%;
[69] (2019)	3 years; weekends, holidays, soccer match days, day before holiday and day after holiday	LSTM and other statistical methods	- -	average MAPE of 5.55 %

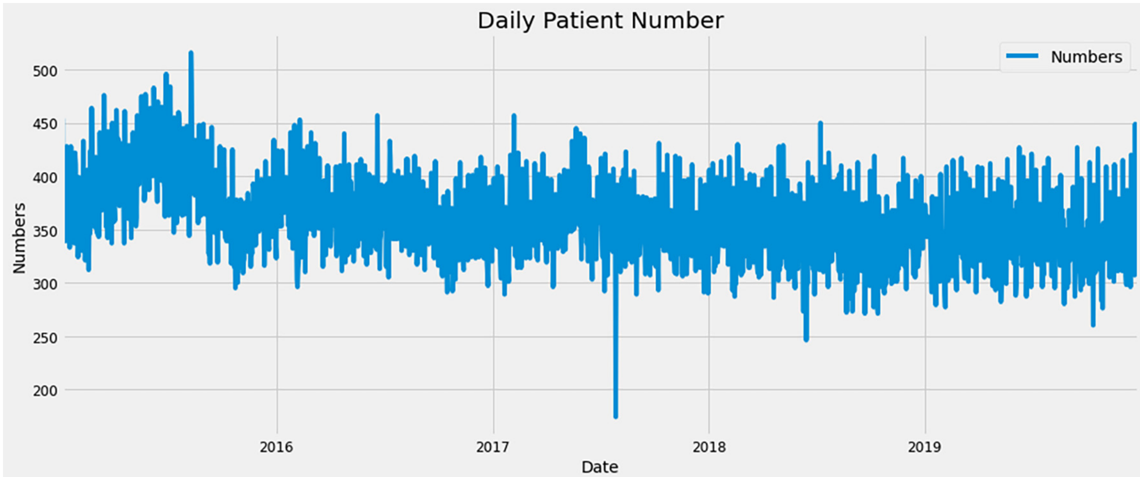

Fig. 1 – Daily ED visits over the 5 years (2015–2019).

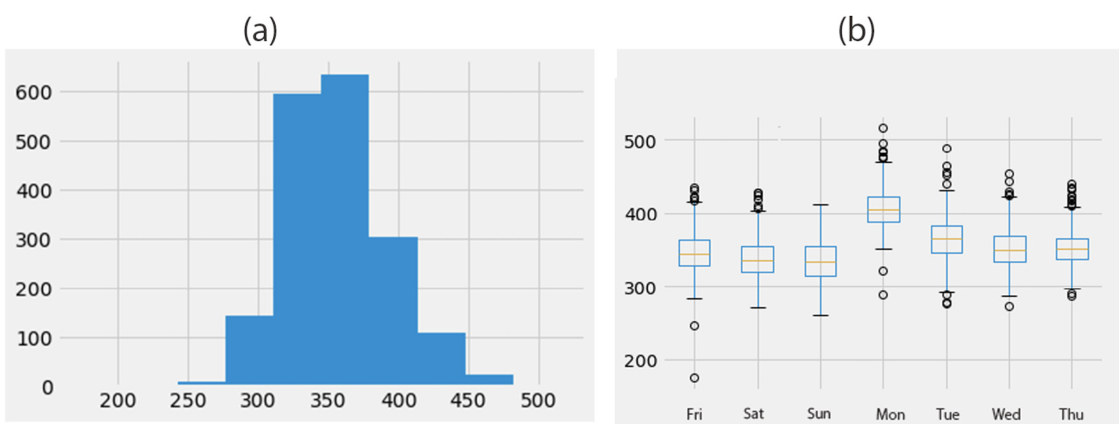

Fig. 2 – (a) Frequency histogram of the 5-year dataset; (b) Daily visits pattern by the day of the week.

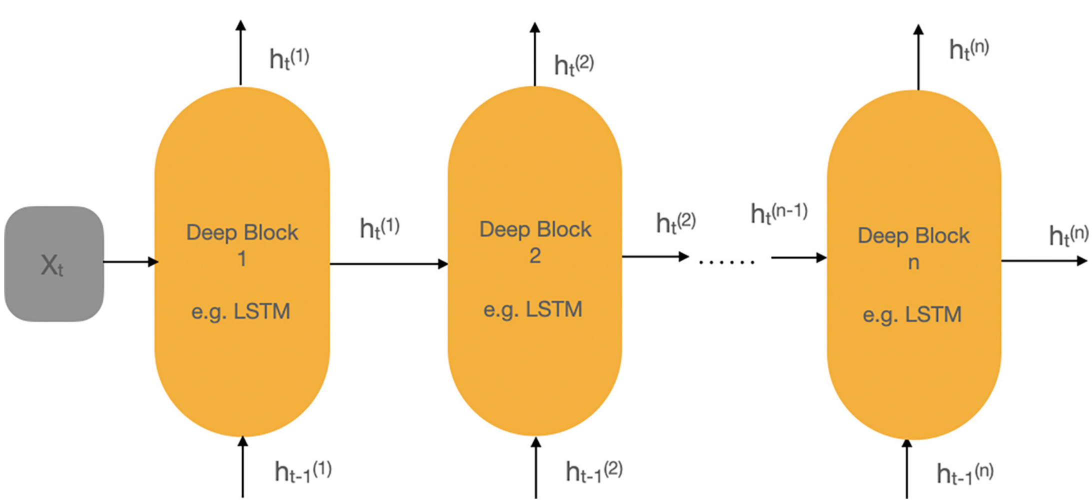

Fig. 3 – The architecture of the proposed method.

the number of hidden layers and the number of hidden units in each hidden layer, to achieve the best forecasting accuracy. In order to predict the ED visits (y_t) at date t, we also need to consider the temporal relationship, that is how many previous days (time steps) will be used for the prediction, that is, $y_{t-1}, y_{t-2}, y_{t-3}, \ldots, y_{t-r}$. Here, we choose r from $\{10, 20, 30, 40, 50, 60\}$. For example, when time steps $= 50$, we use the previous 50 day to predict the current day, and this process will repeat itself to move forward to predict the following days. We set the layer size from 1 to 6, and hidden units number from $\{50, 100, 200, 300, 400, 500, 600\}$. We use tanh as gate activation, dropout rate is 20% for each layer. We set the output of the recurrent layers to return a sequence. Each model is trained independently until convergence by using Mean Square Error (MSE) as the loss function and by the different optimizers such as Adam, SGD and RmsProp.

Table 4 gives the possible configurations for the deep stacked models. After studying the different combinations, we choose the best configuration (that achieves best performance) for the different prediction tasks, as shown in Table 5. It is worth noting that with RmsProp as the optimizer, the deep stacked models have better performance, which is consistent with a previous study [66]. We configure other non-stacked deep learning models such as ConvLSTM, BiLSTM and CNN in the ways similar to a previous study [66]. For the Prophet model, we have added weekly seasonality and the corresponding Fourier order to 3, others with default settings. For ARIMA model, we have used similar configurations from a previous study [80].

To further improve the applicability of our work here, we have also determined a variety of prediction tasks (as different hospitals may require different planning strategies). Task I utilises the entire five-year dataset, task II utilises the most recent three-years worth of data, task III utilises the most recent one-year worth of data, and task IV utilises the most recent six-months worth of data. In each task, we divided the data into 7:3 ratio of training and testing set for model building and testing.

4. Results

Table 6 gives the prediction performances of the various models. Task I to Task IV are the four tasks with 5 years to 6 months data length. As we can see, for the 4 prediction tasks, the deep learning stacked models (DRNN, DLSTM and DGRU networks) are all consistently better (4–7%) than ARIMA models and Prophet models in terms of mean accuracy. In almost all cases, they also outperform typical non-stacked deep models such as ConvLSTM, BiLSTM and CNN(1–2%). Mean accuracy (1-MAPE) metrics for the different deep learning models are as follows: for deep RNN networks, it is in the range of 93.82–94.59%; for deep LSTM networks, 94.02–94.33%; for deep GRU networks, 93.88–94.40%; for ConvLSTM net-

Table 4 – Configuration Candidates of Deep Stacked Networks for ED Visits Forecasting.

Hyper-Parameters	DeepRNN	DeepLSTM	DeepGRU
Recurrent layer	1 to 6	1 to 6	1 to 6
Hidden Units	50 to 600	50 to 600	50 to 600
time steps (r)		10 to 60	
Epochs		200	
Batch size		32	
Activation function		tanh	
Dropout rate		20%	
Loss function		MSE	
Optimiser		Adam, SGD, RmsProp	

Table 5 – Configuration of the deep stacked architecture for ED visits prediction.

Task	DRNN	DLSTM	DGRU
Task I (5-year dataset, training set vs. testing set: 70% vs 30%)	3 SimpleRNN layers, 50 hidden units per layer, RmsProp optimizer	3 layers, $r = 60$, 200 units per layer, RmsProp optimizer	3 layers, $r = 60$, 200 units per layer, RmsProp optimizer
Task II (3-year dataset,training set vs. testing set: 70% vs 30%)	3 SimpleRNN layers, 50 units per layer, RmsProp optimizer	2 layers, $r = 60$, 200 units per layer, RmsProp optimizer	3 layers, $r = 60$, 200 units per layer, RmsProp optimizer
Task III (1-year dataset, training set vs. testing set: 70% vs 30%)	3 SimpleRNN layers, 50 units per layer, RmsProp optimizer	3 layers, $r = 20$, 200 units per layer, RmsProp optimizer	3 layers, $r = 20$, 300 units per layer, RmsProp optimizer
Task IV (6-month dataset, training set vs. testing set: 70% vs 30%)	4 SimpleRNN layers, 50 units per layer, RmsProp optimizer	3 layers, $r = 10$, 300 units per layer, RmsProp optimizer	3 layers, $r = 10$, 200 units per layer, RmsProp optimizer

Table 6 – Prediction Performance of different Models for Daily ED Visits Forecasting.

Task	Metrics	DRNN	DLSTM	DGRU	ConvLSTM	BiLSTM	CNN	Prophet	ARIMA
Task I	RMSE	25.26	**24.71**	26.07	33.43	27.15	28.54	41.80	30.51
	MAPE %	6.02	**5.72**	6.12	7.74	6.00	6.80	12.06	9.28
	MAE	20.03	**19.22**	20.36	25.66	20.29	22.59	34.39	24.01
Task II	RMSE	24.92	26.69	25.71	33.95	31.21	28.75	26.05	**24.52**
	MAPE %	**5.70**	5.98	5.92	7.73	6.81	6.62	9.40	9.09
	MAE	19.34	20.64	19.84	26.04	23.23	22.51	20.24	**19.01**
Task III	RMSE	25.78	25.29	25.61	32.18	31.51	30.69	24.98	**24.40**
	MAPE %	5.70	**5.67**	6.06	7.14	7.16	6.95	9.27	9.15
	MAE	19.69	19.45	20.69	24.17	24.41	23.75	**18.44**	19.14
Task IV	RMSE	26.22	28.50	**25.48**	32.36	32.47	26.79	32.03	40.91
	MAPE %	**5.41**	5.85	5.60	6.98	7.10	5.79	10.37	12.66
	MAE	**19.68**	21.31	19.95	25.07	25.59	20.47	25.09	32.76

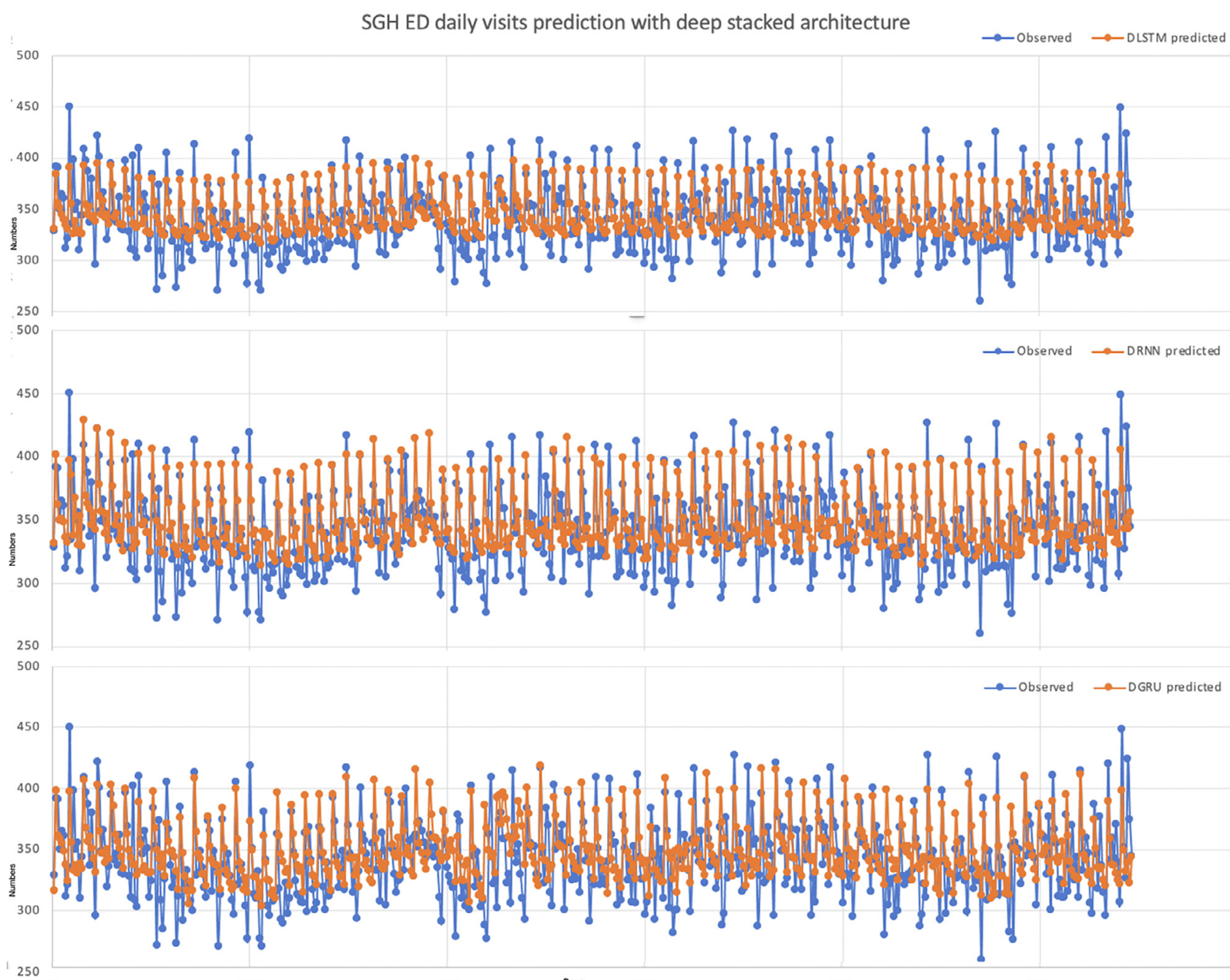

Fig. 4 – Prediction with DLSTM, DGRU and DRNN models with 5 years dataset.

works, 92.27–93.12%; for BiLSTM 92.84–94.00%; for CNN 93.05–94.21%; for ARIMA models 87.34–90.91%; for Prophet models 87.94–90.73%.

Specifically for the deep stacked models, DLSTM achieves the best performance with an MAPE of 5.72% (therefore mean accuracy 94.28%) for Task I; for Task II, DRNN has the best performce with an MAPE of 5.70%; for Task III, DLSTM achieves the best performance with an MAPE of 5.67%, whereas for Task IV, DRNN outperforms others with an MAPE of 5.41%. We can also observe that, for each prediction task, for the

Fig. 5 – Prediction with ConvLSTM, BiLSTM and CNN models with 5 years data set.

deep stacked models, the performance are quite close, for example, for Task II, MAPE of DRNN is 5.7%, DLSTM is 5.98% and DGRU is 5.92%. Another interesting fact we can observe is that, with different configurations, the deep architecture can provide consistently good prediction performance, with MAPE of 5.72%, 5.70%, 5.67%, 5.41% for the 4 different prediction tasks.

Fig. 4–11 are the graphical representation of our predictions with different length of the dataset for different tasks that with deep learning models (the more traditional models are not presented, because their prediction performance, in terms of mean accuracy, are much worse than the deep models, as we can see from the Table 6). We observe that the deep models can predict the future trend reasonably well, especially these deep stacked models.

5. Discussion

Table 3 also gives comparisons (in terms of models and methods used, number of layers, number of hidden units and performance) between our study and the most recent research work with deep learning techniques on ED visits prediction. Using deep learning techniques to predict ED visits is worth pursuing as seen in our current study here. LSTM is one of the most popular deep learning methods used in the ED visits predictions; other deep leaning models such as RNN, GRU, CNN, VAE, ConvLSM, BiLSM are also being used in the studies. For our study, we have utilised the LSTM, RNN, GRU models, however, the most important difference between our study and those of previous studies is the deep stacked architecture that we have used. Therefore, for different prediction tasks with different length of data, we can choose different number of layers and with different hidden units to achieve the best performance. As can be seen form Table 3, the size of deep layers are 2–4 for the different tasks and for different deep stacked models (DRNN, DLSTM and DGRU). Past studies have suggested that hidden layer size should neither be too large nor too small [75,81] and experiments in this study have confirmed these insights. We have also observed that, some studies [67,69] only use one hidden layer (LSTM), while in other models [51,66] more hidden layers have been used. Our deep stacked architecture uses a different number of hidden layers to achieve the best prediction performance depending on the length of the dataset and the used deep model. In particular,

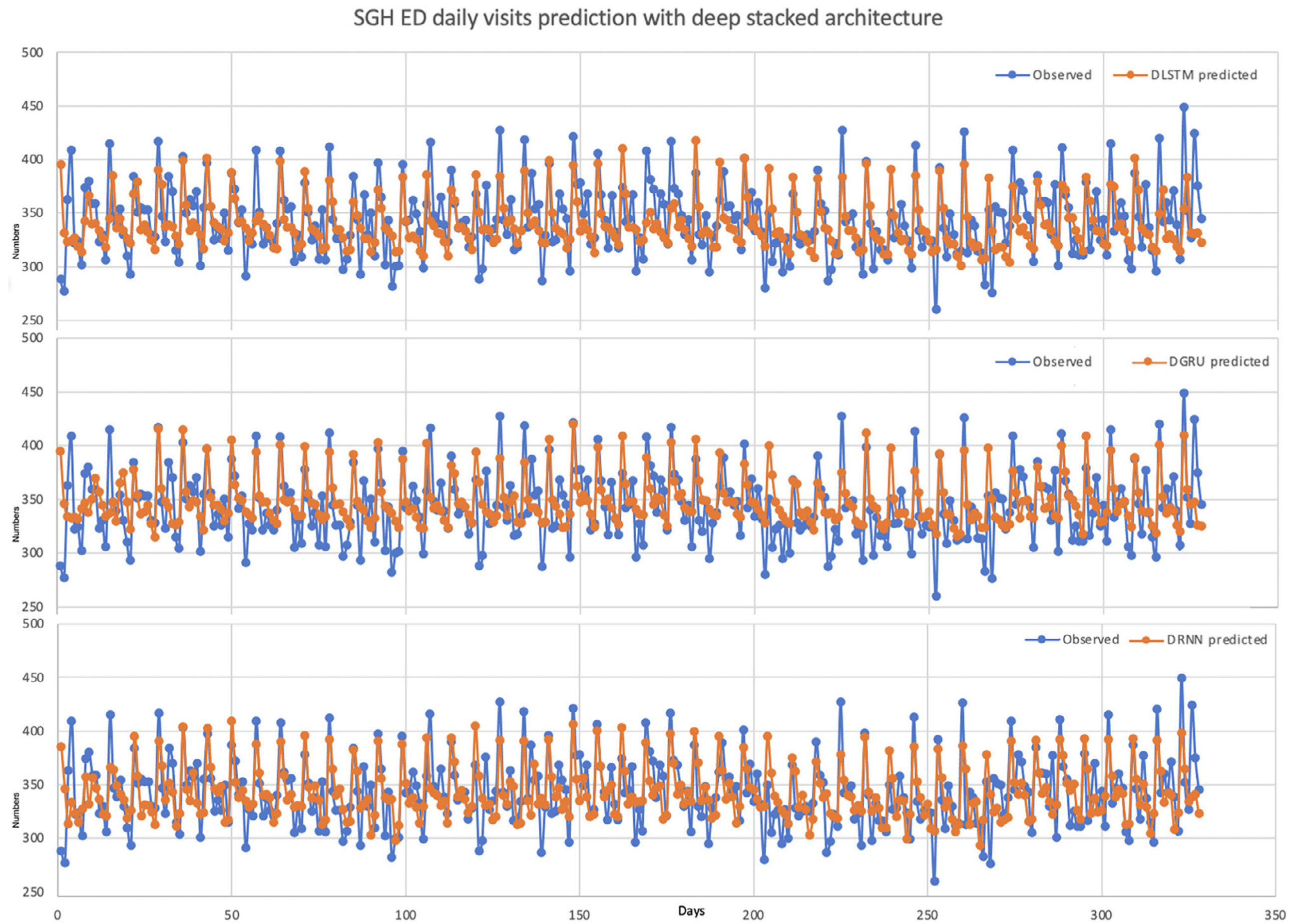

Fig. 6 – Prediction with DLSTM, DGRU and DRNN models with 3 years data set.

we notice that for different length of the dataset (5 years, 3 years, 1 year and 6 months), different deep stacked RNN-variants networks have comparable performance in terms of RMSE, MAPE and MAE (can be seen in Table 6). It appears that it is not necessarily the longer the historical dataset, the better the performance for the different deep stacked models that are being employed here. In our case, a 6-month dataset can already allow the deep models to achieve good (more than 94% mean prediction accuracy) prediction performance. It is worth noting that the number of hidden units in each hidden layer also does not need to be very large, for our cases 50–300, other studies 10–75. It is also interesting to note that the time steps r used for the four tasks are different. For the first two tasks, $r = 60$, for the third and fourth task, $r = 20$ and $r = 10$ respectively. As the length of each dataset for the four tasks are increasingly shorter, the time steps r are also getting shorter. This could because with time steps of 60 days, the proposed deep stacked models can learn long-term trends hidden within the dataset well, whereas with 20 and 10 days, the models can learn more internal patterns from the short-term variations.

For the performance, our deep stacked architecture achieves an MAPE range of 5.41–5.72% for the different predictions (all in daily), with an average MAPE 5.62%. It is better than and comparable to the existing studies [51,69]. Karsanti et al. (2019)[68] used an LSTM model and achieved a very good performance with an MAPE of 4.71%, however, the work focused on monthly ED prediction, while our work focus on daily prediction. Harrou et al. [66] used a VAE model and compared its performance with other deep learning models and achieved a good performance, 0.925, in terms of R^2.

By considering meteorological and calendar factors[51,65], studies have shown that the forecasting performance of certain models can be improved. There are also other researches that have considered other factor like temperature, but only with limited improvements in their forecasting accuracy [54,47,58]. The performance of our deep stacked models has not improved despite taking into account the temperature and daily relative humidity factors.

It is worth noting that, we use the longest dataset (5 years daily data) among the different deep learning studies, while other studies range from 2 years to 3 years 5 months data. We also test more prediction tasks in the work using different lengths of data (5 years, 3 years, 1 year and 6 months) and achieve consistently good performance with different configurations, while other studies only focused on one length. This implies for different length of data, the configurations should be fine tuned to achieve the best performance, and these configurations provided by this study can provide a guidance for different EDs with different lengths of available data. Interestingly, there is another implication here, that this study has answered one important question for building practical deep

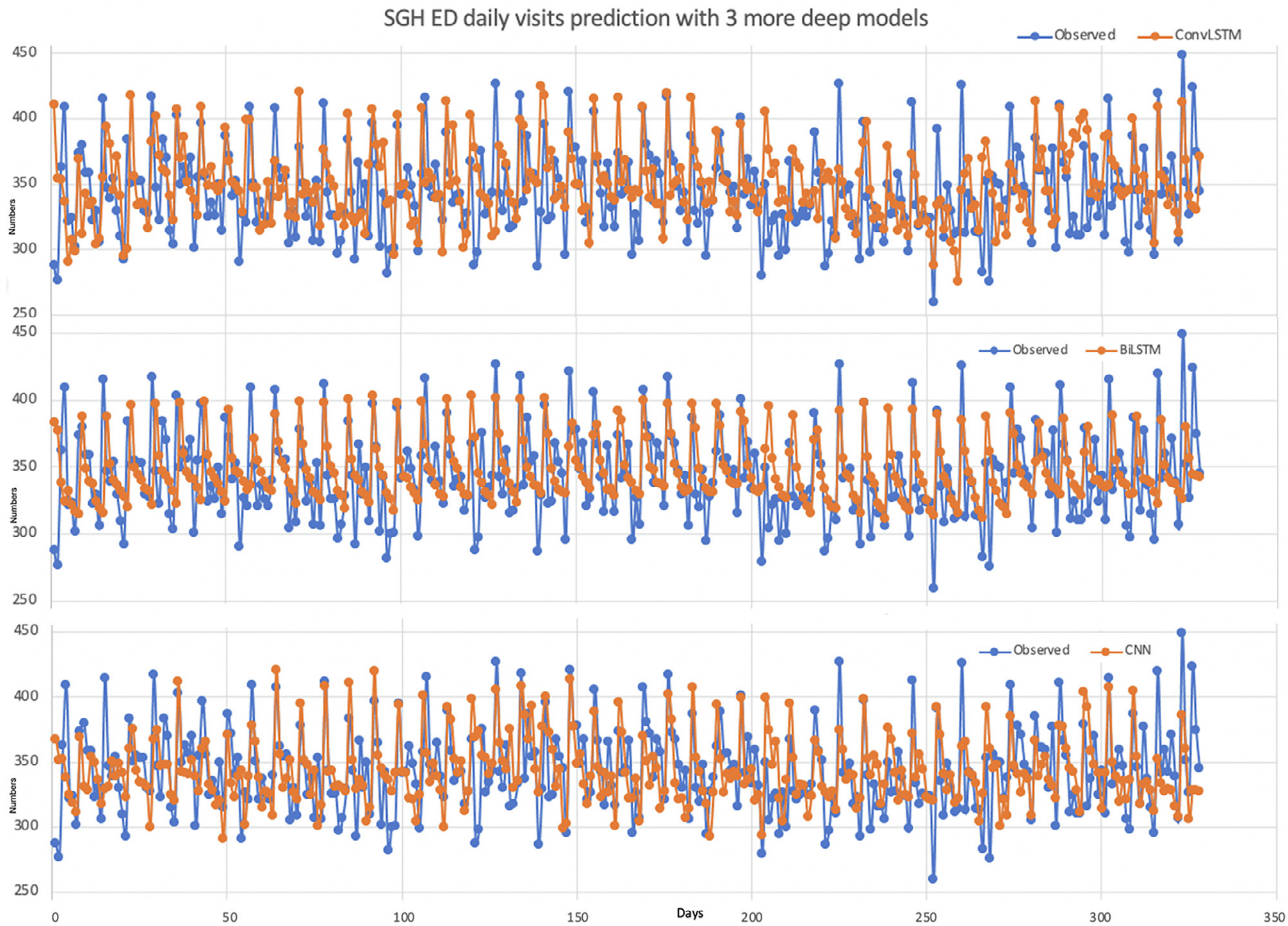

Fig. 7 – Prediction with ConvLSTM, BiLSTM and CNN models with 3 years data set.

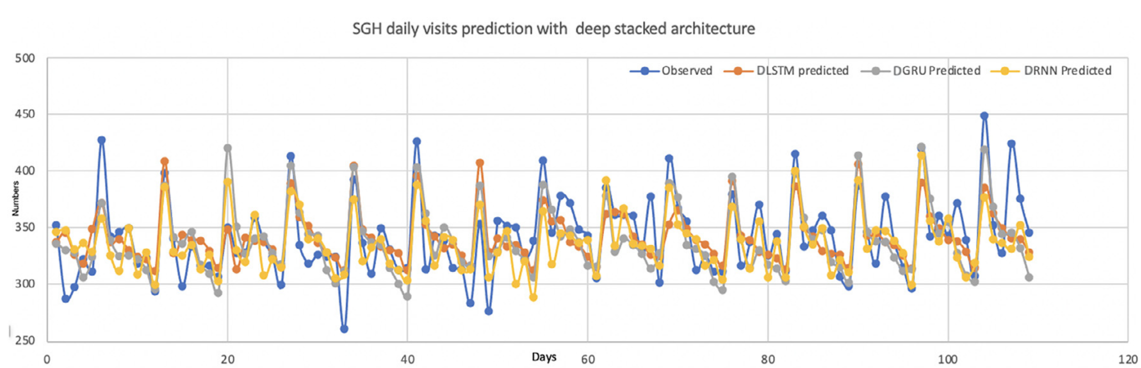

Fig. 8 – Prediction with DLSTM, DGRU and DRNN models with 1 year data set.

leaning models for ED patients prediction, that is, how much data that is needed to achieve good prediction performance? The experiments we provided in this study can be used as a guide to the research community and can be used as references when building the deep learning models for these technical experts.

Our study here and elsewhere (Table 3) have shown the effectiveness of the deep learning techniques in this area of research. It is our belief that these results will inspire and encourage medical practitioners to use deep learning techniques in forecasting ED crowding. As part of our future work, we will be focusing on forecasting the categories of the ED patients, such as acuity levels, and/or with specific conditions such as respiratory and circulatory related problems. These will assist the hospitals in resource planning and manpower deployment at both macro and micro levels.

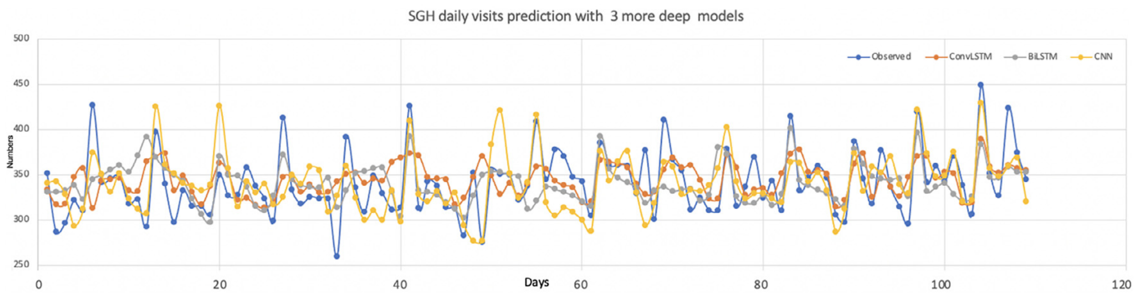

Fig. 9 – Prediction with ConvLSTM, BiLSTM and CNN models with 1 year data set.

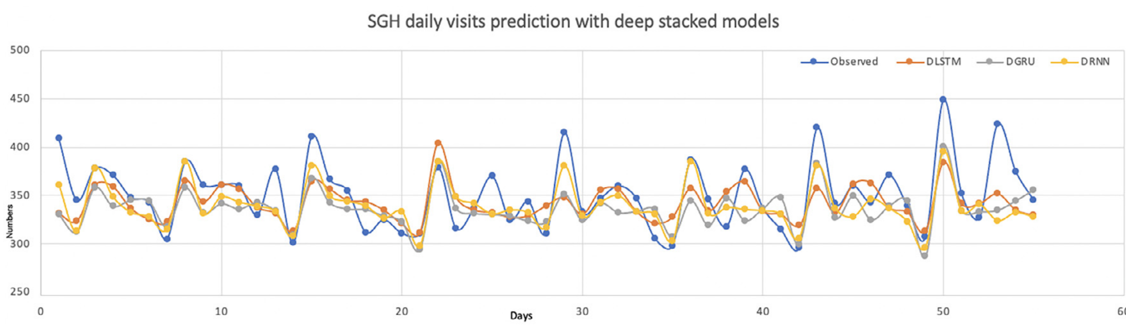

Fig. 10 – Prediction with DLSTM, DGRU and DRNN models with half year data set.

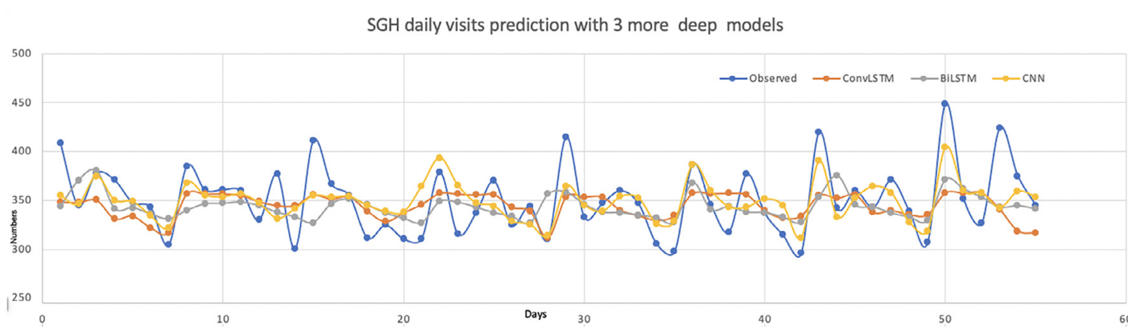

Fig. 11 – Prediction with ConvLSTM, BiLSTM and CNN models with half year data set.

6. Conclusion

The ability to predict ED demand is of utmost significance for hospital resource planning. The deep stacked architecture proposed in this study provides with a very high mean accuracy level (94.28–94.59%) in predicting the daily ED patient numbers. With the ability to predict ED visits accurately in advance, our models can assist in hospital resource planning with cost-effective solutions for manpower, materials and to improve outcomes in healthcare systems. We have also made a few important observations from this study: firstly, the deep stacked models which have RNN, LSTM and GRU as the components all outperformed more traditional models such as ARIMA and Prophet models. It has also been observed that in almost all cases, they have better performance than the non-stacked deep models such as ConvLSTM, BiLSTM and CNN in daily ED visits prediction. Secondly, with different configurations, deep learning models can achieve consistent high performance in predicting the daily ED visits using different lengths of period. As shown in our last prediction task, the dataset only requires a 6-month length of data to achieve a high accuracy. Thirdly, to achieve a high performance in prediction, neither the layer size nor the hidden units in each layer need to be very large in the deep stacked architecture. These research findings and insights can also be extended to develop deep learning forecasting methods for other domains as well.

Funding

This project was funded by the SingHealth-SUTD Population Health Innovation Fund (Grant No. PHIF 2021–02).

CRediT authorship contribution statement

Xinxing Zhao: Software, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Joel Weijia Lai:** Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. **Andrew Fu Wah Ho:** Methodology, Formal analysis, Writing - review & editing, Visualization. **Nan Liu:** Methodology, Formal analysis, Writing - review & editing, Visualization. **Marcus Eng Hock Ong:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Kang Hao Cheong:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

[1] Di Somma S, Paladino L, Vaughan L, Lalle I, Magrini L, Magnanti M. Overcrowding in emergency department: an international issue. Intern Emerg Med 2015;10(2):171–5.

[2] Horwitz LI, Green J, Bradley EH. Us emergency department performance on wait time and length of visit. Ann Emerg Med 2010;55(2):133–41.

[3] Schoenenberger LK, Bayer S, Ansah JP, Matchar DB, Mohanavalli RL, Lam SS, et al. Emergency department crowding in singapore: insights from a systems thinking approach. SAGE Open Med 2016;4. 2050312116671953.

[4] Lau H, Dadich A, Nakandala D, Evans H, Zhao L. Development of a cost-optimization model to reduce bottlenecks: A health service case study. Expert Systems 2018;35(6) e12294.

[5] Tang KJW, Ang CKE, Constantinides T, Rajinikanth V, Acharya UR, Cheong KH. Artificial intelligence and machine learning in emergency medicine. Biocybern Biomed Eng 2020;41 (1):156–72.

[6] Bobrovitz N, Lasserson DS, Briggs AD. Who breaches the four-hour emergency department wait time target? a retrospective analysis of 374,000 emergency department attendances between 2008 and 2013 at a type 1 emergency department in england. BMC Emergency Med 2017;17(1):1–10.

[7] Khanna S, Boyle J, Good N, Bell A, Lind J. Analysing the emergency department patient journey: discovery of bottlenecks to emergency department patient flow. Emergency Med Australasia 2017;29(1):18–23.

[8] P. Vezyridis, S. Timmons, National targets, process transformation and local consequences in an nhs emergency department (ed): a qualitative study, BMC Emergency Medecine 14 (1) (2014) 1–11.

[9] Chang AM, Lin A, Fu R, McConnell KJ, Sun B. Associations of emergency department length of stay with publicly reported quality-of-care measures. Acad Emerg Med 2017;24(2):246–50.

[10] Hobbs D, Kunzman SC, Tandberg D, Sklar D. Hospital factors associated with emergency center patients leaving without being seen. Am J Emerg Med 2000;18(7):767–72.

[11] Singer AJ, Thode Jr HC, Viccellio P, Pines JM. The association between length of emergency department boarding and mortality. Acad Emerg Med 2011;18(12):1324–9.

[12] Jo S, Jeong T, Jin YH, Lee JB, Yoon J, Park B. Ed crowding is associated with inpatient mortality among critically ill patients admitted via the ed: post hoc analysis from a retrospective study. Am J Emerg Med 2015;33(12):1725–31.

[13] Abidova A, da Silva PA, Moreira S. Predictors of patient satisfaction and the perceived quality of healthcare in an emergency department in portugal. Western J Emergency Med 2020;21(2):391–403.

[14] Derlet RW, Richards JR. Overcrowding in the nation's emergency departments: complex causes and disturbing effects. Ann Emerg Med 2000;35(1):63–8.

[15] Ho AFW, To BZYS, Koh JM, Cheong KH. Forecasting hospital emergency department patient volume using internet search data. IEEE Access 2019;7:93387–95.

[16] Kadri F, Harrou F, Chaabane S, Tahon C. Time series modelling and forecasting of emergency department overcrowding. J Med Syst 2014;38(9):1–20.

[17] Lucini FR, Fogliatto FS, Da Silveira GJ, Neyeloff JL, Anzanello MJ, Kuchenbecker RS, et al. Text mining approach to predict hospital admissions using early medical records from the emergency department. Int J Med Informatics 2017;100:1–8.

[18] Yang T, Song J, Li L. A deep learning model integrating sk-tpcnn and random forests for brain tumor segmentation in mri. Biocybern Biomed Eng 2019;39(3):613–23.

[19] Panicker RO, Kalmady KS, Rajan J, Sabu M. Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybern Biomed Eng 2018;38(3):691–9.

[20] Hegde RB, Prasad K, Hebbar H, Singh BMK. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybern Biomed Eng 2019;39(2):382–92.

[21] Kok C, Jahmunah V, Oh SL, Zhou X, Gururajan R, Tao X, et al. Automated prediction of sepsis using temporal convolutional network. Comput Biol Med 2020;127:103957.

[22] Lai JW, Ang CKE, Acharya UR, Cheong KH. Schizophrenia: A survey of artificial intelligence techniques applied to detection and classification. Int J Environ Res Public Health 2021;18(11):6099.

[23] Jiang F, Fu Y, Gupta BB, Liang Y, Rho S, Lou F, et al. Deep learning based multi-channel intelligent attack detection for data security. IEEE Trans Sustainable Computing 2018;5 (2):204–12.

[24] D'Angelo G, Palmieri F. Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial–temporal features extraction. J Network Computer Appl 2021;173:102890.

[25] Altuve M, Lizarazo P, Villamizar J. Human activity recognition using improved complete ensemble emd with adaptive noise and long short-term memory neural networks. Biocybern Biomed Eng 2020;40(3):901–9.

[26] Rajapriya R, Rajeswari K, Thiruvengadam S. Deep learning and machine learning techniques to improve hand movement classification in myoelectric control system. Biocybern Biomed Eng 2021;41(2):554–71.

[27] Liang Y, Tang Z, Yan M, Liu J. Object detection based on deep learning for urine sediment examination. Biocybern Biomed Eng 2018;38(3):661–70.

[28] Jiang W. Applications of deep learning in stock market prediction: recent progress. Expert Syst Appl 2021;184 115537.

[29] Vijayalakshmi B, Ramar K, Jhanjhi N, Verma S, Kaliappan M, Vijayalakshmi K, et al. An attention-based deep learning

model for traffic flow prediction using spatiotemporal features towards sustainable smart city. Int J Commun Syst 2021;34(3).
[30] Han S-Y, Kwak N-S, Oh T, Lee S-W. Classification of pilots' mental states using a multimodal deep learning network. Biocybern Biomed Eng 2020;40(1):324–36.
[31] Kassania SH, Kassanib PH, Wesolowskic MJ, Schneidera KA, Detersa R. Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: a machine learning based approach. Biocybern Biomed Eng 2021;41(3):867–79.
[32] Mishra NK, Singh P, Joshi SD. Automated detection of covid-19 from ct scan using convolutional neural network. Biocybern Biomed Eng 2021;41(2):572–88.
[33] Hashemzehi R, Mahdavi SJS, Kheirabadi M, Kamel SR. Detection of brain tumors from mri images base on deep learning using hybrid model cnn and nade. Biocybern Biomed Eng 2020;40(3):1225–32.
[34] Karevan Z, Suykens JA. Transductive lstm for time-series prediction: An application to weather forecasting. Neural Networks 2020;125:1–9.
[35] Z. King, J. Farrington, M. Utley, E. Kung, S. Elkhodair, S. Harris. et al. Machine learning for real-time aggregated prediction of hospital admission for emergency patients, medRxiv.
[36] Pianykh OS, Guitron S, Parke D, Zhang C, Pandharipande P, Brink J, et al. Improving healthcare operations management with machine learning. Nature Mach Intell 2020;2(5):266–73.
[37] Davazdahemami B, Peng P, Delen D. A deep learning approach for predicting early bounce-backs to the emergency departments. Healthcare Anal 2022:100018.
[38] van Klaveren D, Rekkas A, Alsma J, Verdonschot RJ, Koning DT, Kamps MJ, et al. Covid outcome prediction in the emergency department (cope): using retrospective dutch hospital data to develop simple and valid models for predicting mortality and need for intensive care unit admission in patients who present at the emergency department with suspected covid-19. BMJ Open 2021;11(9) e051468.
[39] Hsu C-C, Chu C-C, Lin C-H, Huang C-H, Ng C-J, Lin G-Y, et al. A machine learning model for predicting unscheduled 72 h return visits to the emergency department by patients with abdominal pain. Diagnostics 2021;12(1):82.
[40] Cheong KH, Tang KJW, Zhao X, Koh JEW, Faust O, Gururajan R, et al. An automated skin melanoma detection system with melanoma-index based on entropy features. Biocybern Biomed Eng 2021;41(3):997–1012.
[41] Liang W, Yao J, Chen A, Lv Q, Zanin M, Liu J, et al. Early triage of critically ill covid-19 patients using deep learning. Nature Commun 2020;11(1):1–7.
[42] Joseph JW, Leventhal EL, Grossestreuer AV, Wong ML, Joseph LJ, Nathanson LA, et al. Deep-learning approaches to identify critically ill patients at emergency department triage using limited information. J Am College Emergency Physicians Open 2020;1(5):773–81.
[43] Lee J-T, Hsieh C-C, Lin C-H, Lin Y-J, Kao C-Y. Prediction of hospitalization using artificial intelligence for urgent patients in the emergency department. Sci Reports 2021;11 (1):1–8.
[44] Feretzakis G, Sakagianni A, E. Loupelis, Karlis G, Kalles D, Tzelves L, et al., Predicting hospital admission for emergency department patients: A machine learning approach, in: Informatics and Technology in Clinical Care and Public Health, IOS Press, 2022, pp. 297–300.
[45] Beiser DG, Jarou ZJ, Kassir AA, Puskarich MA, Vrablik MC, Rosenman ED, et al. Predicting 30-day return hospital admissions in patients with covid-19 discharged from the emergency department: A national retrospective cohort study. J Am College Emergency Physicians Open 2021;2(6) e12595.
[46] Taylor SJ, Letham B. Forecasting at scale. Am Stat 2018;72 (1):37–45.
[47] Carvalho-Silva M, Monteiro MTT, de Sá-Soares F, Dória-Nóbrega S. Assessment of forecasting models for patients arrival at emergency department. Oper Res Health Care 2018;18:112–8.
[48] Juang W-C, Huang S-J, Huang F-D, Cheng P-W, Wann S-R. Application of time series analysis in modelling and forecasting emergency department visits in a medical centre in southern taiwan. BMJ Open 2017;7(11) e018628.
[49] Golmohammadi D. Predicting hospital admissions to reduce emergency department boarding. Int J Prod Econ 2016;182:535–44.
[50] Araz OM, Olson D, Ramirez-Nafarrate A. Predictive analytics for hospital admissions from the emergency department using triage information. Int J Prod Econ 2019;208:199–207.
[51] Sudarshan VK, Brabrand M, Range TM, Wiil UK. Performance evaluation of emergency department patient arrivals forecasting models by including meteorological and calendar information: A comparative study. Comput Biol Med 2021;135 104541.
[52] Kam HJ, Sung JO, Park RW. Prediction of daily patient numbers for a regional emergency medical center using time series analysis. Healthcare Inform Res 2010;16(3):158–65.
[53] Boyle J, Jessup M, Crilly J, Green D, Lind J, Wallis M, et al. Predicting emergency department admissions. Emergency Med J 2012;29(5):358–65.
[54] Marcilio I, Hajat S, Gouveia N. Forecasting daily emergency department visits using calendar variables and ambient temperature readings. Acad Emerg Med 2013;20(8):769–77.
[55] Calegari R, Fogliatto FS, Lucini FR, Neyeloff J, Kuchenbecker RS, Schaan BD. Forecasting daily volume and acuity of patients in the emergency department. Comput Math Methods Med 2016.
[56] M. Hertzum, Forecasting hourly patient visits in the emergency department to counteract crowding, The Ergonomics Open Journal 10 (1).
[57] Whitt W, Zhang X. Forecasting arrivals and occupancy levels in an emergency department. Operations Res Health Care 2019;21:1–18.
[58] Erkamp NS, van Dalen DH, de Vries E. Predicting emergency department visits in a large teaching hospital. Int J Emergency Med 2021;14(1):1–11.
[59] Xu M, Wong T-C, Chin K-S. Modeling daily patient arrivals at emergency department and quantifying the relative importance of contributing variables using artificial neural network. Decis Support Syst 2013;54(3):1488–98.
[60] N.B. Menke, N. Caputo, R. Fraser, J. Haber, C. Shields, M.N. Menke, A retrospective analysis of the utility of an artificial neural network to predict ed volume, American Journal of Emergency Medicine 32 (6) (2014) 614–617.
[61] Zlotnik A, Gallardo-Antolin A, Alfaro MC, Pérez MCP, Martínez JMM, et al. Emergency department visit forecasting and dynamic nursing staff allocation using machine learning techniques with readily available open-source software. CIN: Computers Inform, Nursing 2015;33(8):368–77.
[62] Xu Q, Tsui K-L, Jiang W, Guo H. A hybrid approach for forecasting patient visits in emergency department. Quality Reliability Eng Int 2016;32(8):2751–9.
[63] Khaldi R, El Afia A, Chiheb R. Forecasting of weekly patient visits to emergency department: real case study. Procedia Computer Science 2019;148:532–41.
[64] Zhang Y, Luo L, Yang J, Liu D, Kong R, Feng Y. A hybrid arima-svr approach for forecasting emergency patient flow. J Ambient Intelligence Humanized Computing 2019;10 (8):3315–23.

[65] Zhang Y, Zhang J, Tao M, Shu J, Zhu D. Forecasting patient arrivals at emergency department using calendar and meteorological information. Appl Intell 2022:1–12.
[66] Harrou F, Dairi A, Kadri F, Sun Y. Forecasting emergency department overcrowding: A deep learning framework. Chaos, Solitons Fractals 2020;139:110247.
[67] Kadri F, Baraoui M, Nouaouri I. An lstm-based deep learning approach with application to predicting hospital emergency department admissions. In: 2019 International Conference on Industrial Engineering and Systems Management (IESM). IEEE; 2019. p. 1–6.
[68] H.T. Karsanti, I. Ardiyanto, L.E. Nugroho, Deep learning-based patient visits forecasting using long short term memory, in: 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT), IEEE, 2019, pp. 344–349.
[69] Yousefi M, Yousefi M, Fathi M, Fogliatto FS. Patient visit forecasting in an emergency department using a deep neural network approach. Kybernetes 2019;49(9):2335–48.
[70] Chan T-H, Jia K, Gao S, Lu J, Zeng Z, Ma Y. Pcanet: A simple deep learning baseline for image classification? IEEE Trans Image Process 2015;24(12):5017–32.
[71] Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing. IEEE Comput Intell Mag 2018;13(3):55–75.
[72] Arulkumaran K, Deisenroth MP, Brundage M, Bharath AA. Deep reinforcement learning: A brief survey. IEEE Signal Process Mag 2017;34(6):26–38.
[73] Dairi A, Harrou F, Zeroual A, Hittawe MM, Sun Y. Comparative study of machine learning methods for covid-19 transmission forecasting. J Biomed Inform 2021;118:103791.
[74] Hewamalage H, Bergmeir C, Bandara K. Recurrent neural networks for time series forecasting: Current status and future directions. Int J Forecast 2021;37(1):388–427.
[75] Lv Y, Duan Y, Kang W, Li Z, Wang F-Y. Traffic flow prediction with big data: a deep learning approach. IEEE Trans Intell Transp Syst 2014;16(2):865–73.
[76] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521 (7553):436–44.
[77] Telgarsky M. Benefits of depth in neural networks. In: Conference on learning theory, PMLR. p. 1517–39.
[78] Sagheer A, Kotb M. Time series forecasting of petroleum production using deep lstm recurrent networks. Neurocomputing 2019;323:203–13.
[79] Yamak PT, Yujian L, Gadosey PK. A comparison between arima, lstm, and gru for time series forecasting. In: Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence. p. 49–55.
[80] Sun Y, Heng BH, Seow YT, Seow E. Forecasting daily attendances at an emergency department to aid resource planning. BMC Emergency Med 2009;9(1):1–9.
[81] Zhao X, Ang CKE, Acharya UR, Cheong KH. Application of artificial intelligence techniques for the detection of alzheimer's disease using structural mri images. Biocybern Biomed Eng 2021;41(2):456–73.
[82] Wang Peipei, Zheng Xinqi, Ai Gang, Liu Dongya, Zhu. Bangren. Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. Chaos Solitons Fractals 2020;140:110214.