RISK PREDICTION OF CRITICAL VITAL SIGNS FOR ICU PATIENTS USING RECURRENT NEURAL NETWORK

A Thesis

Presented to

The Faculty of the Department of Computer Science

California State University, Los Angeles

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Computer Science

By

Daniel Chang

December 2019

The thesis of Daniel Rich Chang is approved.

Dr. Mohammad Pourhomayoun, Committee Chair

Dr. Russ Abbot

Dr. YuQing Zhu

Dr. Zilong Ye

Dr. Eun-Young Elaine Kang, Department Chair

California State University, Los Angeles

December 2019

ABSTRACT

Risk Prediction of Critical Vital Signs for ICU Patients Using Recurrent Neural Network

By

Daniel Chang

Monitoring vital signs for Intensive Care Unit (ICU) patients is an absolute necessity to help assess the general physical health. In this research, we use machine learning to make a classification forecast that uses continuous ICU vital signs measurements to predict whether the vital signs of the next hour would reach the critical value or not. With the early warning, nurses and doctors can prevent emergency situations that may cause organ dysfunction or even death before it is too late.

In this study, the data includes vital sign measurements, laboratory test results, procedures, medications collected from over 40,000 patients. After data preprocessing, bias data balancing, feature extraction, and feature selection, we have a clean dataset with informative and discriminating features. Then, various machine learning algorithms including Random Forest, XGBoost, Artificial Neural Networks (ANN), and LSTM were developed to predict critical vital signs of ICU patients 1-hour beforehand. We particularly developed predictive models to predict Heart Rate, Blood Oxygen Level (SpO2), Mean Arterial Pressure (MAP), Respiratory Rate (RR), Systolic Blood Pressure (SBP). The results demonstrated the accuracy of the developed methods.

ACKNOWLEDGMENTS

I would like to appreciate MIT Laboratory for Computational Physiology and collaborating research groups for providing valuable data. I would also like to thank Dr. Mohammad Pourhomayoun Ph.D. for guiding me improving the performance of this thesis. And, I want to thank Dr. David Ray Chang M.D. for giving the medical knowledge and suggestions.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

CHAPTER 1

Introduction

In an intensive care unit (ICU), vital signs indicate the status of the patient's life-sustaining functions. If any of the vital signs has reached the critical value, the patient is in danger and needs immediate help from nurses or doctors. Prolonged hypoperfusion in critically ill patients leads to multiple organ failure, which further increases mortality. Hypoperfusion usually presents as unstable vital signs, which is instantly warned via the ICU alarm [2],[3].

Frequently in the critically ill setting, the alarm from the ICU monitor is too slow for intervention. For example, the ICU monitor will only alarm when tachycardia occurs in stage 2 hypovolemic shock patients with blood loss 15-30%, or when hypotension occurs in stage 3 with blood loss over 30%. The optimal timing of treatment should be stage 1 when the patient only has blood loss under 15%, but the vital signs will mostly be normal at this stage [4].

TABLE 1. STAGES OF HYPOVOLEMIC SHOCK

	Stage I	**Stage II**	**Stage III**	**Stage IV**
Blood Loss	<750	750 - 1500	1500 - 2000	> 2000
% Blood Vol.	<15%	15 – 30%	30 – 40%	> 40%
Heart Rates	<100	>100	>120	>140
Blood Pressure	Normal	Increased diastolic BP	Systolic BP < 100	Systolic BP < 70
Resp. Rate	14 – 20	20 – 30	30 – 40	> 40
Mental Status	Normal	Anxious	Confused	Lethargic

Purpose of Project

An early prediction could help doctors and nurses be aware of critical vital signs in advance and prevent the patient from deteriorating. By using an accurate predictive model, in a list of patients with normal vital signs, doctors and nurses will know which patients to look into first. Thus, the medical team will have more time for diagnosis and intervention.

The rapid advances in data science, body sensors, and artificial intelligence have led to the development of effective patient monitoring and data-driven analytics systems that allow for gathering information from patients and analyze it to predict health conditions [8]. These systems have shown potential effectiveness in decreasing healthcare costs and reducing morbidity and mortality [8]-[22].

In this paper, we develop an accurate predictive model using several machine learning algorithms including Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) networks to predict critical vital signs of ICU patients 1-hour beforehand.

CHAPTER 2

The Dataset

The data of this paper, MIMIC-III (Medical Information Mart for Intensive Care III) [5], includes patient general information, vital sign measurements, laboratory test results, procedures, medications, and mortality collected from over 40,000 patients between 2001 and 2012. In total, we collected 1,323,067 (hours) data samples from all patients after the data preprocessing.

Features and Label

In general ICU, if primary vital signs are out of normal range, then ICU monitors would warn the nurses that patients have abnormal situations. Thus, we want to predict whether the vital signs of the next hour will change to abnormal (above upper bound, below lower bound) or still stays in the normal range, which are our two target classifications. In other word, the goal is to predict the transition from currently normal status to future abnormal status.

Table 2 shows the list of primary vital sign measurements and its normal range based on the National Early Warning Score (NEWS) and AHA ACLS guidelines [6], [7]. We are using 6 vital signs which are the features in 6 different individual models to predict 0 as normal or 1 as abnormal for each vital sign.

TABLE 2. SOME VITAL SIGN MEASUREMENTS [6], [7]

Feature	**Normal Range**
Heart Rate (HR)	50 to 130 (beats per minute)
Blood Oxygen Level (SpO_2)	90% to 100%
Mean Arterial Pressure (MAP)	65 to 110 (mm Hg)
Respiratory Rate (RR)	8 to 30 (Breaths per minute)
Systolic Blood Pressure (SBP)	90 to 160 (mm Hg)
Diastolic Blood Pressure (DBP)	60 to 110 (mm Hg)

Combine Duplicate Measurements

Because this 45 Gig Big Dataset is collected from over 600 hospitals and each hospital using different LABEL name, we have the same measurements with different item id. As shown in Fig. 2, we can notice that “abp [systolic]” is an abbreviation of “arterial bp [systolic]”, but the ITEMID is different. We also can notice that blood pressure has been measured from different part of body.

In order to deal with this situation, we combine all the vital sign measurements if they represent the same. Thus, “abp [systolic]” and “arterial bp [systolic]” would have the same ITEMID.

	ROW_ID	ITEMID	LABEL	ABBREVIATION	DBSOURCE	LINKSTO	CATEGORY	UNITNAME	PARAM_TYPE	CONCEPTID
295	32	6	abp [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
320	57	51	arterial bp [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
671	408	442	manual bp [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
682	419	455	nbp [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
705	442	480	orthostat bp sitting [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
707	444	482	orthostatbp standing [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
709	446	484	orthostatic bp lying [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
715	452	492	pap [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1437	618	666	systolic unloading	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1748	929	3313	bp cuff [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1750	931	3315	bp left arm [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1752	933	3317	bp left leg [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1754	935	3319	bp pal [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1756	937	3321	bp right arm [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1758	939	3323	bp right leg [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN
1760	941	3325	bp uac [systolic]	NaN	carevue	chartevents	NaN	NaN	NaN	NaN

Fig. 1. Item Table of Systolic Blood Pressure

Timeline Alignment

Although the raw data is collected by ICU monitors, the timestamps of all measurements are not unified. Since more than 80% of the raw data are measured on hourly basis, we fixed the data collection rate to one sample per hour.

	SUBJECT_ID	HADM_ID	ICUSTAY_ID	ITEMID	CHARTTIME	STORETIME	CGID	VALUE	VALUENUM	VALUEUOM	WARNING
34061542	3	145834	211552.0	807	2101-10-21 06:00:00	2101-10-21 05:46:00	21570.0	306	306.0	NaN	NaN
34028522	4	185777	294638.0	807	2191-03-16 08:00:00	2191-03-16 07:49:00	17144.0	266	266.0	NaN	NaN
34297591	6	107064	228232.0	807	2175-05-30 21:00:00	2175-05-30 21:54:00	15443.0	254	254.0	NaN	NaN
34086330	9	150750	220597.0	807	2149-11-13 12:00:00	2149-11-13 19:10:00	20855.0	212	212.0	NaN	NaN
34062720	13	143045	263738.0	807	2167-01-08 23:30:00	2167-01-08 23:21:00	21205.0	205	205.0	NaN	NaN
34049478	18	188822	298129.0	807	2167-10-03 11:00:00	2167-10-03 11:01:00	19667.0	243	243.0	NaN	NaN
34035552	21	111970	216859.0	807	2135-02-04 06:00:00	2135-02-04 06:05:00	15907.0	200	200.0	NaN	NaN
34033889	25	129635	203487.0	807	2160-11-02 04:45:00	2160-11-02 05:23:00	18654.0	353	353.0	NaN	NaN
34406341	43	146828	225852.0	807	2186-10-04 07:00:00	2186-10-04 07:29:00	17435.0	210	210.0	NaN	NaN
34510579	56	181711	275642.0	807	2104-01-03 16:00:00	2104-01-03 16:19:00	21347.0	201	201.0	NaN	NaN
34486585	62	116009	216609.0	807	2113-02-15 22:00:00	2113-02-15 22:55:00	15023.0	221	221.0	NaN	NaN
34483962	68	108329	272667.0	807	2174-01-08 14:00:00	2174-01-08 14:05:00	14375.0	212	212.0	NaN	NaN

Fig. 2. Non-hourly measurement raw data

Remove Error Measurements

After visualizing the features, we found that some of the measurements of ICU monitor was incorrect. As shown in Fig. 3, for patient 6, respiratory rate has suddenly fallen to 0 between 0 to 10 hours. This was an abnormal situation in ICU, so we considered it as an error from ICU monitor. Thus, we removed all the sudden 0 value of RR. As shown in Fig. 4, the percentage of oxygen in blood (SpO_2) would not be over 100%. We think this was a kind of typo from nurses. However, after talking to the ICU nurses, they said if the measurements values are error or incorrect in ICU, then they would record above 100%. Thus, we remove over 100% value of SpO_2 as well.

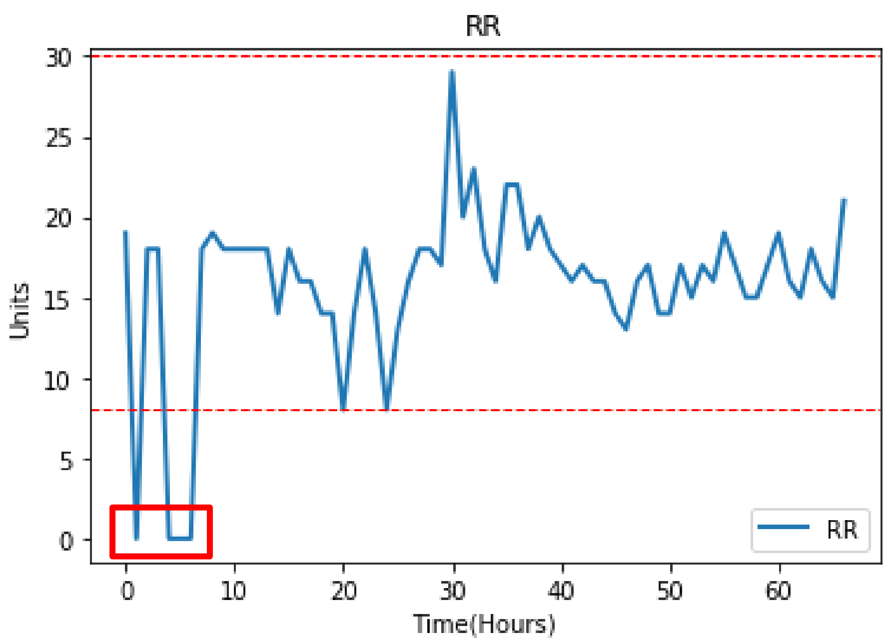

Fig. 3. Continuously Measured Respiratory Rate of Patient 6

	SUBJECT_ID	ICUSTAY_ID	CHARTTIME	heart rate	spo2	arterial bp mean	respiratory rate	arterial bp [diastolic]	arterial bp [systolic]
112892	28187	297092.0	2109-04-10 13:00:00	101.0	101.00	76.0	31.0	60.0	107.0
698209	13183	297445.0	2117-06-11 13:00:00	56.0	101.00	83.0	28.0	63.0	131.0
874907	17699	211512.0	2106-10-26 19:00:00	96.0	100.25	84.0	10.0	61.0	140.0
1068157	22596	270858.0	2133-06-02 20:00:00	83.0	102.00	104.0	77.0	77.0	124.0

Fig. 4. Error measurements of SpO_2

CHAPTER 3

Types of Dataset

Situations in ICU

In ICU, we will have 9 different situations. As Shown in fig. 5, current value of vital sign could be normal, low (below lower bound of normal range), and high (above upper bound of normal range), and the next hour value could be normal, low, high as well. Thus, we would get 9 different situations in ICU. However, we only want to predict for the normal value of current hour of vital sign because the main purpose of this study is to warn the nurses and the doctors that this patient is going to deteriorate in this hour. As shown in fig 6, we only need to consider about 3 situations in ICU.

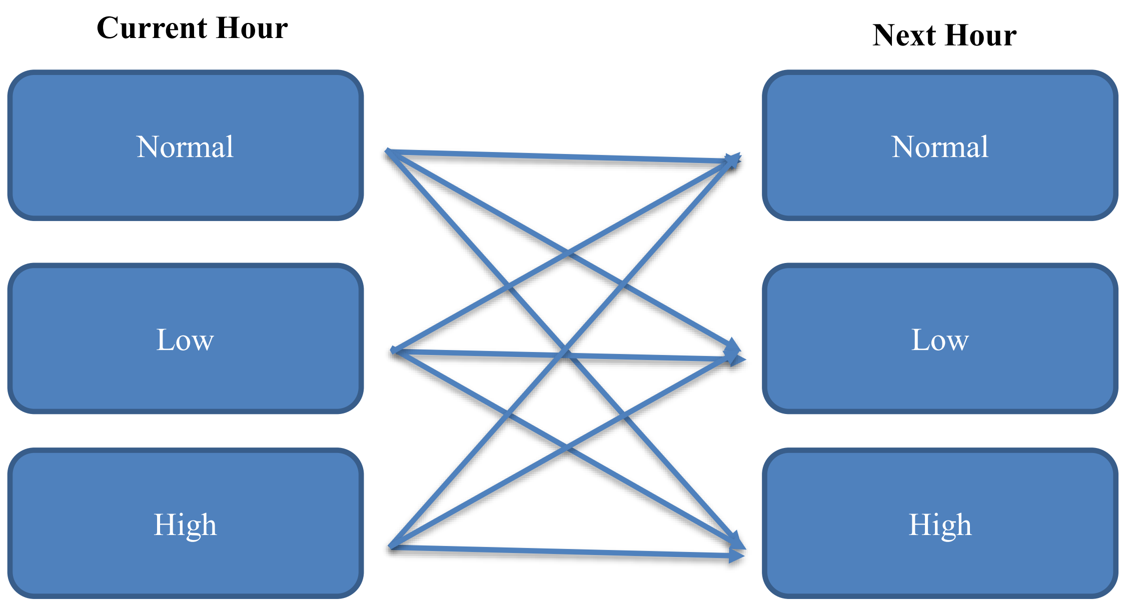

Fig. 5. 9 Situations in ICU

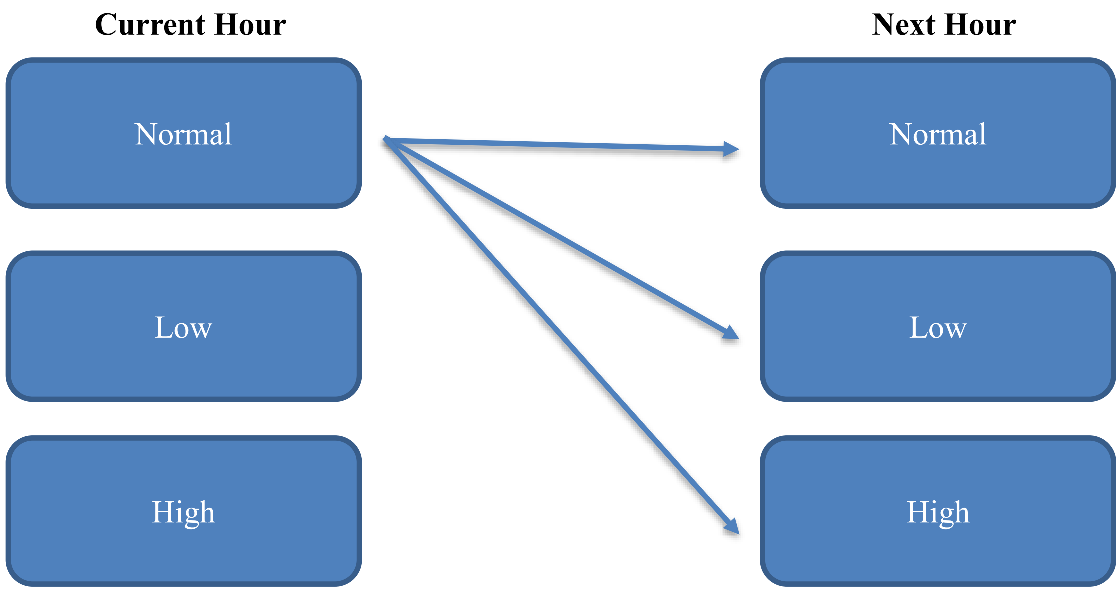

Fig. 6. Current Normal Situations in ICU

Bias of Data

Bias is a common machine learning problem, especially in the medical domain. Machine learning would keep predicting the biased result because the occurrence rate of the healthy state is much higher than the diseased. In our dataset, 99% of HR vital sign is in the normal range and the next hour is still in the normal range as well. The machine learning algorithm trained by the original unbalanced dataset would only predict "normal" for the next hour and would still have a 99% accuracy. In order to avoid this bias problem, we first filtered out all currently abnormal data samples because the goal is to predict the transition from currently normal status to future abnormal status (please note that the patients whose vital signs are currently in abnormal status are not of our prediction interest. We would like to predict whether a patient who is currently in normal situation will change to abnormal range in the next hour or still stays in the normal range).

Secondly, we balanced the binary-target classification in the Training dataset. Thus, we have 50% next hour data within the normal range, 50% next hour data greater than the upper bound and less than the lower bound in the training dataset. (Fig. 7).

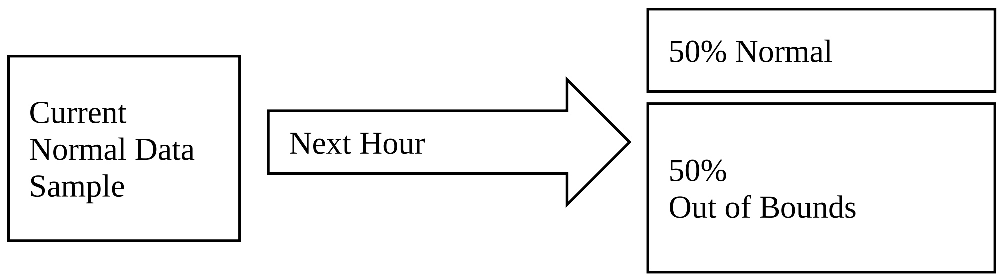

Fig. 7. Balanced Training Dataset

CHAPTER 4

Feature Extraction

As shown in fig.8, with the hourly-measured dataset, we extracted new features from the past 3- and 5-hours window to increase the accuracy, such as the mean and standard deviation of every measurement of the past 3 hours and the past 5 hours. These features can perform how vital signs varied in over the past few hours which are relevant to the next hour value. We derive 24 features from the 3-hour window and 5-hour window.

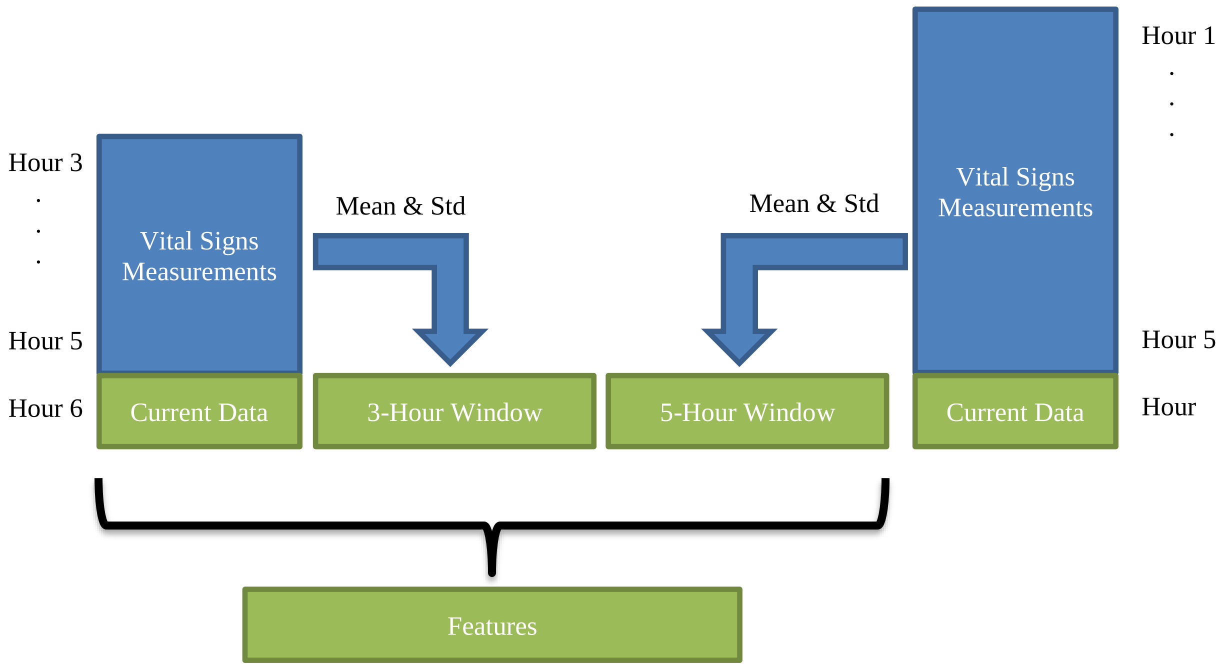

Fig. 8. Concept of Feature Extraction

CHAPTER 5

Deep Learning Method

In this study, we have used several classification machine learning algorithms, such as Random Forest, XGBoost, Artificial Neural Net (ANN), and Recurrent Neural Network (RNN), LSTM. However, we have achieved the best accuracy results with deep learning method, RNN-LSTM.

Artificial Neural Network (ANN)

Artificial Neural Network (ANN) is a computing system like biological neural networks. As shown in fig. 9, the input X are the features, the second layer is neurons, and the output Y as the prediction [26]. The connections between layers are the weights of each nodes. Unlike traditional machine learning, with the calculations of the weights by the neurons, prediction would be more accurate than machine learning.

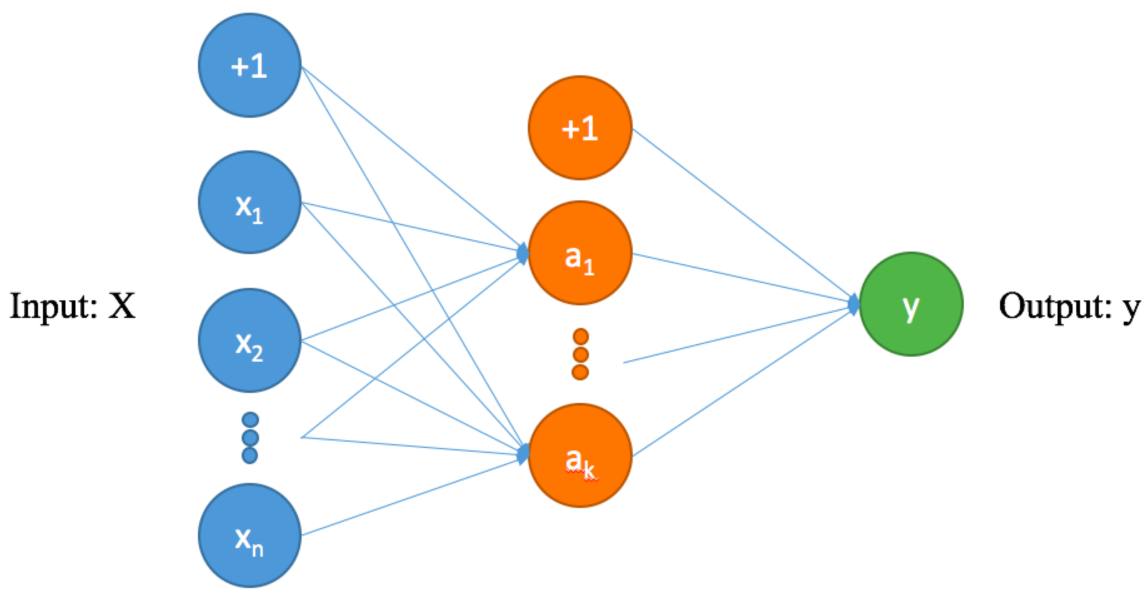

Fig. 9. Structure of Artificial Neural Network (ANN) [26]

Recurrent Neural Network (RNN)

Humans don't start over their thinking when they are making decision. They would make the decision based on their knowledge or experience. For example, you want to predict who would be killed at the next episode of Game of Thrones. It is difficult for traditional ANN because ANN is unclear the time series of all the events.

One of the most effective algorithms for this application is deep RNN due to the strong temporal correlation in the data. As shown in fig. 10, RNN has shown strong potential on forecasting outcomes of time-series or sequential data [24]. Unlike traditional feedforward neural networks, RNN includes feedback connections. It allows the algorithm to memorize historical data and consequently, use a sequence of data samples to make more accurate predictions when there is a temporal correlation in the input data [23].

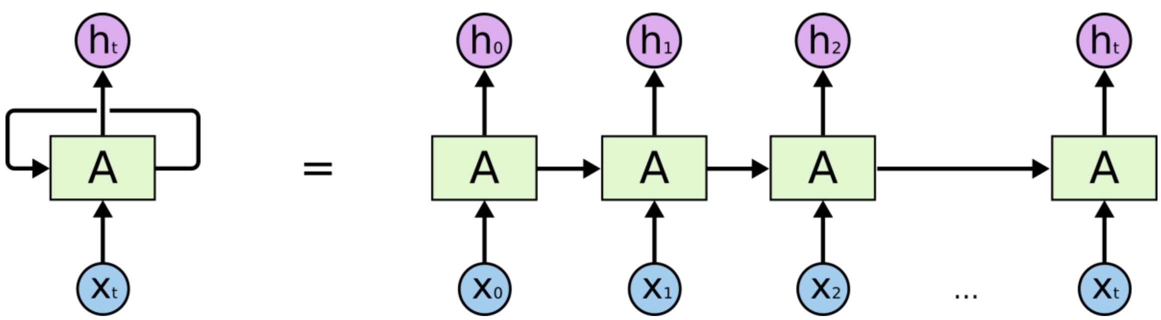

Fig. 10. Structure of Recurrent Neural Network (RNN) [24]

Long Short-Term Memory (LSTM)

A serious problem in training every deep neural network including RNN is vanishing gradient. Vanishing gradient happens when the gradient and, consequently, the network weight correction in the training process (e.g. backpropagation), tends to be very close to zero, especially when we move backward towards the front layers. It makes the training process very slow. This problem is even more serious for RNN with many timesteps.

The Long Short-Term Memory (LSTM) structure can help solve the vanishing gradient problem in traditional RNN [25]. An LSTM unit usually includes a cell, an input gate, an output gate and a forget gate. In this project, we will design and train RNNs with LSTM units. In this approach, the temporal correlation will be taken into account by the memory of the LSTM and the sequential nature of the input data.

As shown in fig. 11, the input would be current hour and the past 5 hours with extractions of the features, and the output would be our target hour 7.

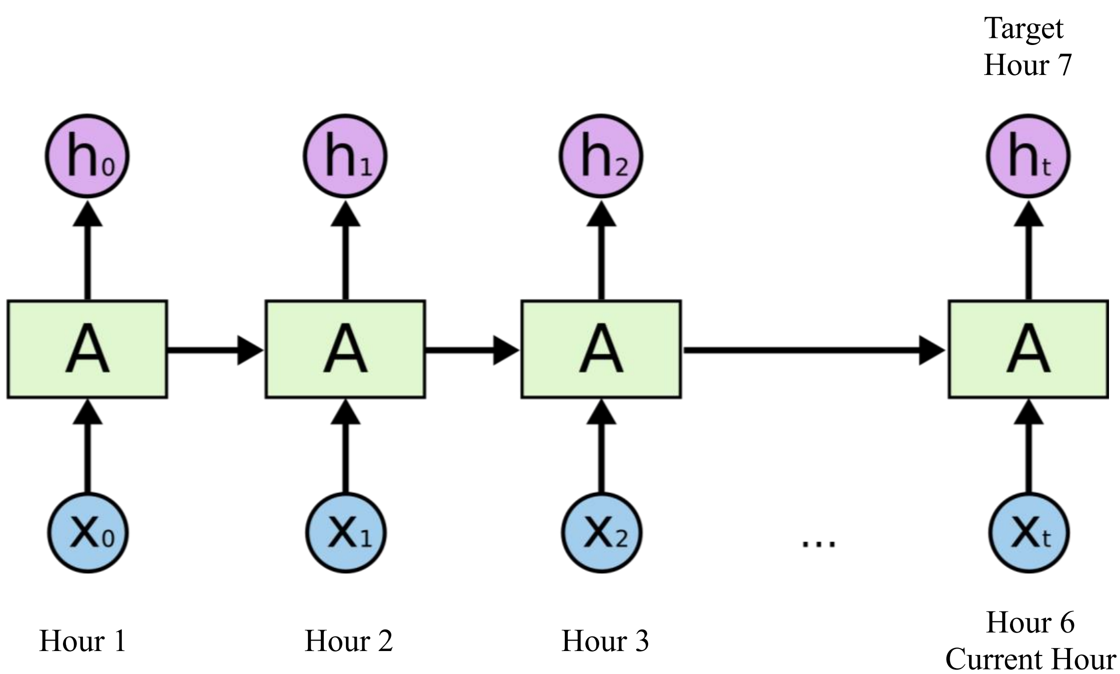

Fig. 11. Features and Target in LSTM model

CHAPTER 6

Conclusion

In this paper, we develop accurate predictive models using various machine learning algorithms including deep RNN with LSTM units to predict critical vital signs of ICU patients 1-hour beforehand. This system can significantly help doctors and nurses be aware of critical vital signs in advance and prevent the patient from deteriorating.

The predictive models were trained and tested on a big dataset including the vital sign measurements, laboratory test results, procedures, medications collected from over 40,000 patients. We randomly split the data samples into training (70%) and testing (30%) datasets.

As shown in Chart I and Table II, for all the vital signs, the RNN-LSTM has the best score of Area Under the Curve (AUC) based on the True Positive Rate (TPR) and the False Positive Rate (FPR) plane. Fig. 2 demonstrates ROC curves for RNN-LSTM model.

TABLE 3. BEST RESULTS FOR VITAL SIGNS

Vital signs	**AUC Score**
Heart Rate (HR)	91.3%
Blood Oxygen Level (SpO_2)	80.1%
Mean Arterial Pressure (MAP)	75.0%
Respiratory Rate (RR)	82.2%
Systolic Blood Pressure (SBP)	78.7%

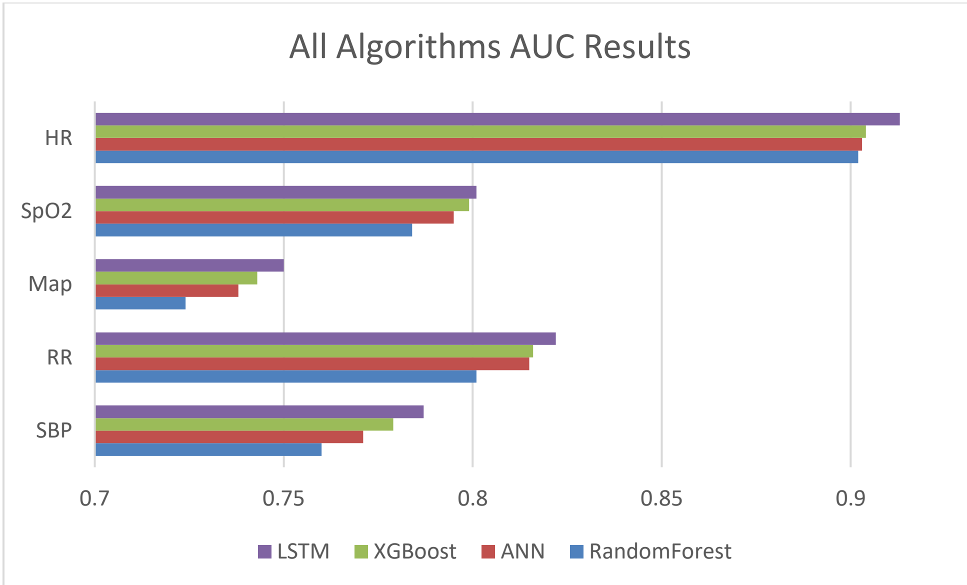

Fig. 12. AUC Results from all Algorithms

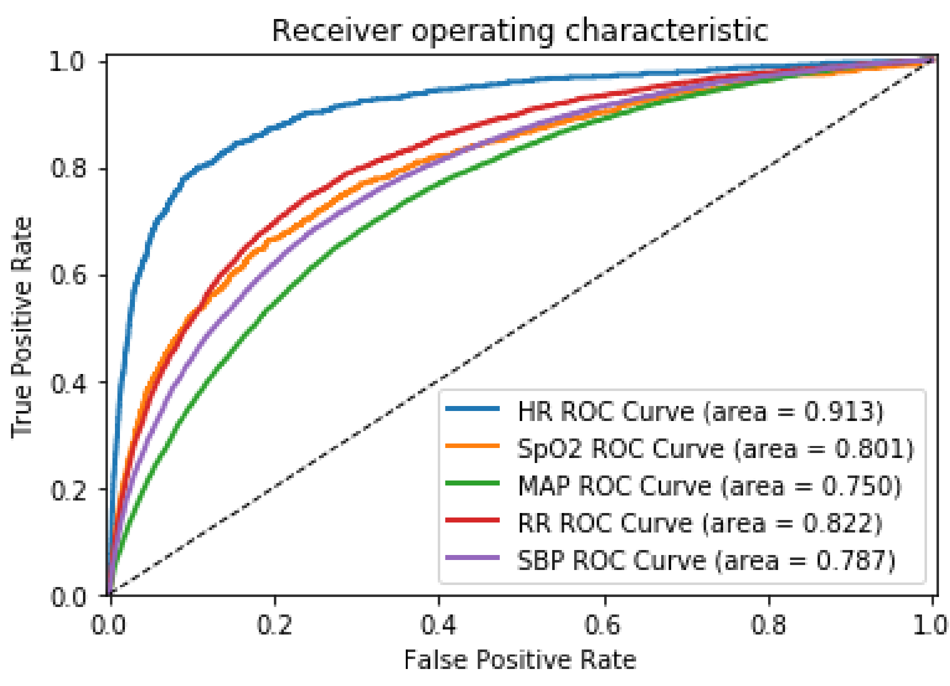

Fig. 13. ROC curves for RNN-LSTM model

REFERENCES

[1] Daniel Chang, David Chang, Mohammad Pourhomayoun, “Risk Prediction of Critical Vital Signs for ICU Patients Using Recurrent Neural Network,” The 2019 International Conference on Computational Science and Computational Intelligence (CSCI'19), Las Vegas, Dec 2019.

[2] Curr Opin Crit Care. 2011 Apr. 17(2):153-9

[3] JAMA. 2016 Feb 23. 315 (8):801-10

[4] ATLS Advanced Trauma Life Support Program for Doctors (7th Ed.). Chicago: American College of Surgeons, 2014.

[5] MIMIC-III, a freely accessible critical care database. Johnson AEW, Pollard TJ, Shen L, Lehman L, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, and Mark RG. Scientific Data (2016). DOI: 10.1038/sdata.2016.35

[6] Resuscitation. 2013Apr;84(4):465-70. DOI: 10.1016/j.resuscitation.2012.12.016. Epub 2013 Jan 4.

[7] Circulation. 2015;132:S444–S464

[8] A. Kalatzis, B. Mortazavi, M. Pourhomayoun, “Interactive Dimensionality Reduction for Improving Patient Adherence in Remote Health Monitoring,” The 2018 International Conference on Computational Science and Computational Intelligence (CSCI'18), Las Vegas, Dec 2018.

[9] M. Pourhomayoun, et al., Multiple model analytics for adverse event prediction in remote health monitoring systems. Proc IEEE EMBS Conf. Healthcare Innovation & Point-of-Care Technologies.

[10] S. Yoo , A. Kalatzis, N. Amini, M. Pourhomayoun, “Interactive Predictive Analytics for Enhancing Patient Adherence in Remote Health Monitoring”, The 8th ACM MobiHoc2018 Workshop on Pervasive Wireless Healthcare, 2018.

[11] M. Pourhomayoun , Nemati E, Mortazavi B, Sarrafzadeh M, Context-aware data analytics for activity recognition. Data Analytics 2015.

[12] M. Pourhomayoun, et al.,“A Robust Remote Health Monitoring and Data Processing System for Rural Area with Limited Internet Access,” International Conference on Body Area Networks, BodyNets 2016, Italy, Dec 2016.

[13] M. R. Vahedi, K. B. MacBride, W. Wunsik, Y. Kim, C. Fong, A. J. Padilla, A. Zhong, S. Kulkarni, B. Jiang, S. Arunachalam, M. Pourhomayoun, “Predicting Glucose Levels in Patients with Type1 Diabetes Based on Physiological and Activity Data”, The 8th ACM MobiHoc2018 Workshop on Pervasive Wireless Healthcare, 2018.

[14] N Alshurafa, C Sideris, M. Pourhomayoun, H Kalantarian, M Sarrafzadeh, JA Eastwood, “Remote Health Monitoring Outcome Success Prediction using Baseline and First Month Intervention Data,” IEEE journal of biomedical and health informatics, 2016.

[15] N. Alshurafa, J. Eastwood, J. Liu, W. Xu, M. Pourhomayoun, M. Sarrafzadeh, "Improving Compliance in Remote Health Monitoring System using Battery Optimization," IEEE Journal of Biomedical and Health Informatics (JBHI), 2014.

[16] H. Kalantarian, N. Alshurafa, M. Pourhomayoun, S. Sarin, M. Sarrafzadeh, " Non-Invasive Detection and Classification of Food Intake using Audio Spectrograms," IEEE Sensors Journal, 2014..

[17] C. Sideris, et al., Using electronic health records to predict severity of condition for congestive heart failure patients. SmartHealthSys, 2015.

[18] M. Pourhomayoun, et al., “Why Do We Need a Remote Health Monitoring System? A Study on Predictive Analytics for Heart Failure Patients,” International Conf. on Body Area Networks, BodyNets 2016.

[19] E. Jovanov, et al., "Patient Monitoring Using Personal Area Networks of Wireless Intelligent Sensors," Biomed Sci Instrum, 2001.

[20] C. Sideris, et al., " Validating Severity of Condition Prediction Model built from Electronic Health Records on Remote Health Monitoring Patients", Elsevier Pervasive and Mobile Computing Journal (PMC).

[21] N. Alshurafa, et al., "A Framework for Predicting Adherence in Remote Health Monitoring Systems," Wireless Health Conf. (WH2014), 2014.

[22] N. Alshurafa, et al., "Remote Health Monitoring: Predicting Outcome Success based on Contextual Features," 36th IEEE Conference of the Engineering in Medicine and Biology (EMBC2014), 2014.

[23] J. Schmidhuber, Deep learning in neural networks: An overview, Neural networks , vol.61, pp. 85-117, 2015.

[24] Christopher Olah, “Colah’s Blog: Understanding LSTM Networks,” August 27, 2015

[25] H. Sak, A. Senior, and F. Beaufays, Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition, 2014, arXiv:1402.1128.

[26] PJ Wang, “Using Keras Build LSTM Model, Stock Prediction as Example,” April 13, 2018