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

Daniel R. Chang
Computer Science Department
California State University, Los Angeles
Los Angeles, USA
dchang33@calstatela.edu

David R. Chang
Division of Nephrology, Department of Internal Medicine, China Medical University Hospital and College of Medicine, China Medical University,
Taichung, Taiwan
D22064@mail.cmuh.org.tw

Mohammad Pourhomayoun
Computer Science Department
California State University, Los Angeles
Los Angeles, USA
mpourho@calstatela.edu

***Abstract*— 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.

CCS CONCEPTS

• Applied computing → Life and medical sciences → Health informatics

KEYWORDS

Predictive Analytics, Vital Signs, ICU, Machine Learning

I. INTRODUCTION AND MOTIVATION

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 [1],[2].

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 [3].

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 [7]. These systems have shown potential effectiveness in decreasing healthcare costs and reducing morbidity and mortality [7]-[21].

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

II. METHOD

A. Data and Target Classifications

The data of this paper, MIMIC-III (Medical Information Mart for Intensive Care III) [4], includes patient general information, vital sign measurements, laboratory test results, procedures, medications, and mortality collected from over 40,000 patients between 2001 and 2012. Table I shows the list of primary vital sign measurements and its normal range based on the National Early Warning Score (NEWS) and AHA ACLS guidelines [5],[6].

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 I. SOME VITAL SIGN MEASUREMENTS

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)

B. *Data Preprocessing*

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. Next, we selected the vital sign measurements as features, and inner join all measurements based on the timestamps and ICU Id. Next, we remove some outliers in the dataset, such as SpO_2 is higher than 100% and RR is suddenly dropped to 0. In total, we collected 1,323,067 (hours) data samples from all patients.

C. *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 currenly in normal situation will change to abnormal range in the next hour or still stays in the normal range).

Secondly, we balanced the three 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. 1).

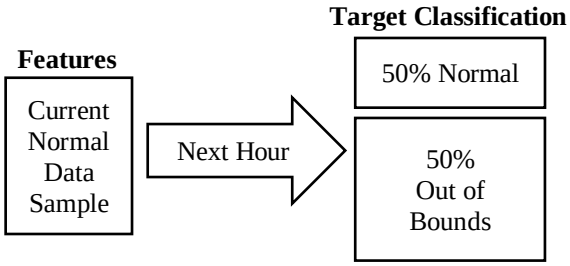


Fig. 1. Balanced Training Dataset

D. *Feature Extraction*

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.

E. *Predictive Model and Machine Learning Algorithms*

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 RNN-LSTM.

One of the most effective algorithms for this application is deep RNN due to the strong temporal correlation in the data. RNN has shown strong potential on forecasting outcomes of time-series or sequential data [22]. 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 [22].

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 [23]. 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.

III. Results and Conclusion

In this paper, we develop accurate predictive models using various machin learning algorithms incluing 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 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 traning (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.

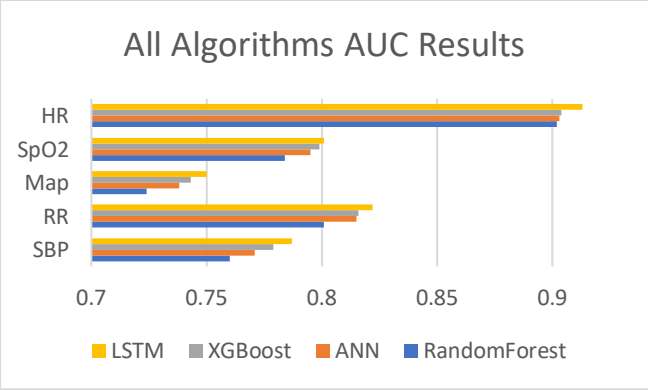
IV. Future Work

Generative Adversarial Networks (GAN) can generate new data by learning the statistic from the training set. With this technology, we can build a semi-supervised learning machine to predict the vital signs. In our future work, we will use GAN and LSTM to improve the performance of the prediction.

TABLE II. 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%

Chart I. AUC Results from all Algorithms



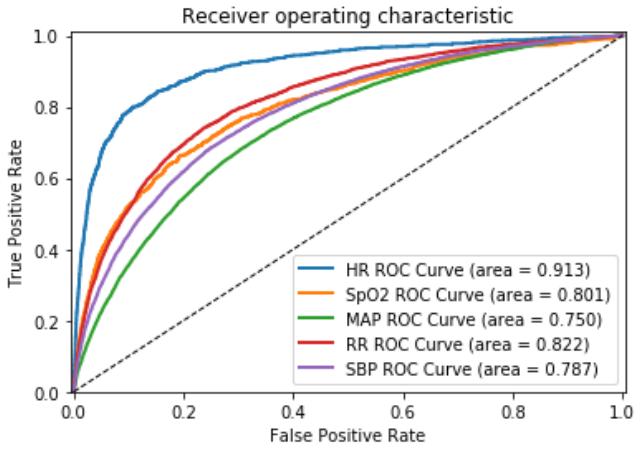


Fig. 2. ROC curves for RNN-LSTM model

V. Acknowledgment

The authors would like to appreciate MIT Laboratory for Computational Physiology and collaborating research groups for providing valuable data.

References

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

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

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

[4] 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. Available from: http://www.nature.com/articles/sdata201635

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

[6] Circulation. 2015;132:S444–S464

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

[12] 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.

[13] 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.
[14] 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.
[15] 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..
[16] C. Sideris, et al., Using electronic health records to predict severity of condition for congestive heart failure patients. SmartHealthSys, 2015.
[17] 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.
[18] E. Jovanov, et al., "Patient Monitoring Using Personal Area Networks of Wireless Intelligent Sensors," Biomed Sci Instrum, 2001.
[19] 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).
[20] N. Alshurafa, et al., "A Framework for Predicting Adherence in Remote Health Monitoring Systems," Wireless Health Conf. (WH2014), 2014.
[21] 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.
[22] J. Schmidhuber, Deep learning in neural networks: An overview, Neural networks , vol.61, pp. 85-117, 2015.
[23] 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.