

긴급대응 시스템을 위한 심층 해석 가능 학습

우엔 쯡 니아, 보탄형, 고보건, 이귀상, 양형정, 김수형
 전남대학교병원 호흡기내과
 전남대학교 인공지능융합학과
 {trongnghia7171, thanhhungqb, imdrkbg}@gmail.com, {gslee, hjyang, shkim}@jnu.ac.kr

Deep Interpretable Learning for a Rapid Response System

Trong-Nghia Nguyen, Thanh-Hung Vo, Bo-Gun Kho, Guee-Sang Lee,
 Hyung-Jeong Yang, Soo-Hyung Kim
 Pulmonology and Critical Care Medicine, Chonnam National University Hospital
 Department of Artificial Intelligence Convergence, Chonnam National University
 {trongnghia7171, thanhhungqb, imdrkbg}@gmail.com, {gslee, hjyang, shkim}@jnu.ac.kr

Abstract

In-hospital cardiac arrest is a significant problem for medical systems. Although the traditional early warning systems have been widely applied, they still contain many drawbacks, such as the high false warning rate and low sensitivity. This paper proposed a strategy that involves a deep learning approach based on a novel interpretable deep tabular data learning architecture, named TabNet, for the Rapid Response System. This study has been processed and validated on a dataset collected from two hospitals of Chonnam National University, Korea, in over 10 years. The learning metrics used for the experiment are the area under the receiver operating characteristic curve score (AUROC) and the area under the precision-recall curve score (AUPRC). The experiment on a large real-time dataset shows that our method improves compared to other machine learning-based approaches.

1. Introduction

In the context of a serious epidemic worldwide, reducing the burden on the medical system is an urgent requirement for many countries. In recent years, alongside the development of Early Warning Score (MEWS) [1] which have been widely used by hospitals for Rapid Response System (RRS) to detect the patient's abnormal status (cardiac arrest and mortality), the proposed of deep learning approaches (DEWS) are also got the attention of the research community. The application of novel deep learning methods achieved significant enhancement over the traditional clinical score and machine learning method. Recurrent neural network (RNN) and Long-Term Short-Term Memory (LSTM) have demonstrated outstanding performance when considering various benchmarks [2], [7]. In 2017, the introduction of attention learning et al. [3] opened a promising future about the integration of a state-of-the-art deep learning method for Rapid Response Systems. In this paper, we proposed a warning system applied TabNet [4] - attentive interpretable tabular learning method for early detection of patients in hospital abnormal status. The system was studied and validated on highly challenging patient dataset. Our method partly solves the inherent problems of imbalanced data and achieves specific efficiency in detecting abnormal states.

2. Proposed Method

We considered detecting a person's in-hospital abnormal status as a binary classification problem that determines the status of patients in the point of time t based on the basic medical measurement values within the previous N hour of

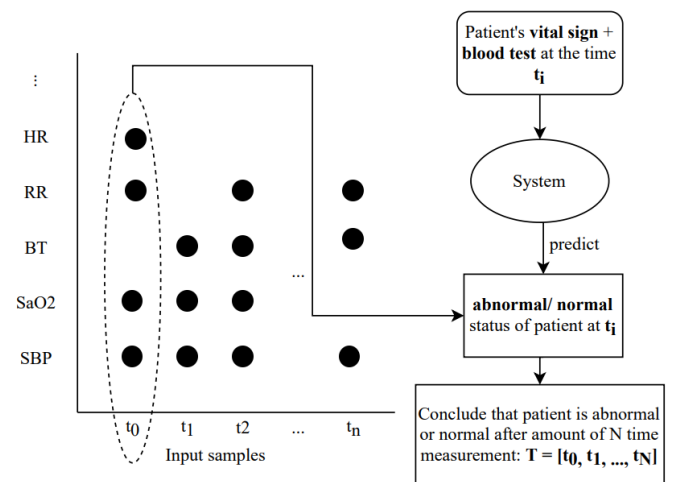


Figure 1: The overall system

that patient. We choose $N = 8$ in this study. The patient's measurement indexes were evaluated every hour. The overall of the system is described as Figure 1, the input of the system are the measurement indexes includes the vital-sign and the blood test values of a patient at a timestamp, output is the

prediction of abnormal/normal status of that patient at the same time. An abnormal patient is the people who performed cardiac arrest and in-hospital airway intubation among hospitalized patients, the normal case is the patient admitted to the internal medicine department and did not undergo cardiopulmonary resuscitation and airway intubation during hospitalization. The final output that we aim for is the result of predicting patient's status after a measurement period $X_i = [X_0, X_1, \dots, X_N]$ with N is the number of measurement hours. As the proposed system was presented in detail as Figure 2, our main idea is firstly converting the raw input data (time series) into tabular data. In the next step, we generate the statistic features from the processed data to increase the training process's information dimension. Finally, the TabNet model is implemented for classifying.

A. Features processing:

The pre-processing stage is the transfer progress from time series to tabular data. The original data have total 23 measurement features by time, includes 5 vital-sign values (Heart Rate (HR), Body Temperature (BT), Oxygen

Raw time series data (n=23)

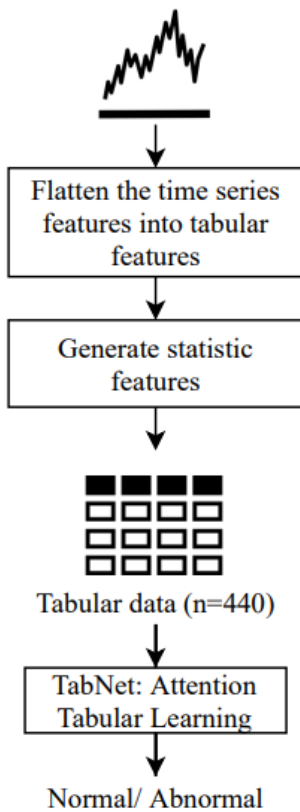


Figure 2: The proposed method

Saturation (SaO₂), Respiration Rate (RR) and Symbolic Blood Pressure (SBP)) and 18 blood test values (Albumin, Lactate, Total Protein, Platelet, Total Bilirubin, Calcium, BUN, CRP, Creatinine, Hgb, WBC count, AST, Chloride, Glucose, Alkaline Phosphatase, Sodium, Potassium, Alt). We

flattened these features based on the window time $N = 8$ we mentioned before to change the data dimension from time series to pure 2D tabular. Therefore, for each feature X in the time-series dimension, we would have 8 values of X_i in the 2D dimension with $X_i = [X_0, X_1, \dots, X_7]$. Besides, the statistical features are also generated from X at each timestamp. These features are the min, max, mean and standard deviation of each set of feature values. After processed, we have total 440 features ($n=440$).

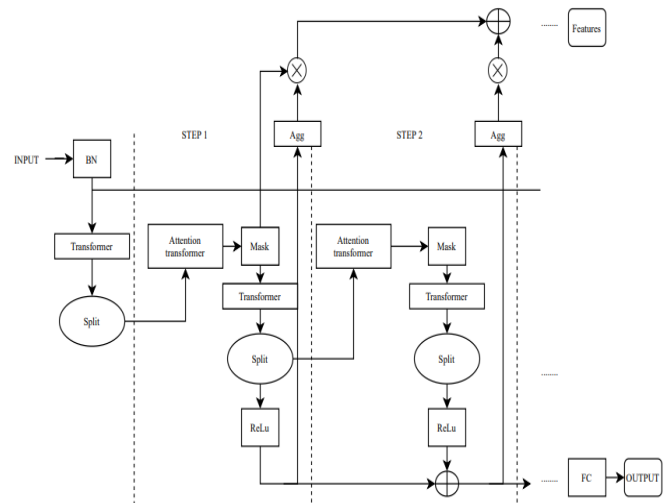


Figure 3: TabNet encoder architecture [4]

B. TabNet model:

TabNet [4] is trained basing on gradient descent-based optimization, allowing for flexible integration into the end-to-end learning process. Sequence attention is applied to the instance-wise features selection at each decision step for better interpretability and learning as the learning capacity is used for the most helpful features. In addition, a single deep learning engine is utilized for feature selection and reasoning. The architecture of Tabnet is presented at Figure 3. The encoder stage of the model is built up by the feature transformer, feature masking, and the attentive transformer. The feature selection masks that contained the model's interpretable information are concatenated for obtaining the important features. In the encoder stage, feature transformer blocks are used at each step for feature reconstructing.

3. Experiments and results

This study was conducted on the dataset established for patients at Chonnam National University Hospital in Hakdong and Hwasun from 2009 to 2019. The data includes two groups of patients: abnormal group and normal group. Considering the measurement indexes of patients at each point of time are a sample, our study uses total 315,946 samples from 3,279 patients. The train set has 2,615 patients with 291,364 samples, and the test set includes 24,582 samples from 664 patients. The train set has 938 abnormal samples, while the test set has 57 abnormal samples. This

distribution shows a significantly imbalance rate between the minority class and majority class. All these data were directly labeled by the doctors.

Table 1: Experiment result on test set

Model	AUROC (%)	AUPRC (%)
DEWS (TabNet)	66.8	29.1
RNN	65.0	13.1
MEWS	50.0	6.2
LGB	54.2	7.1
XGBoost	53.7	8.9

With a highly imbalanced dataset, we used Python 3.7 with scikit-learn Gridsearch library to hyper-tuning parameters for our model. Because of the imbalance problem, we apply the AUROC and AUPRC score for metric learning. We compare our system with Modify early warning system (MEWS), Light gradient boosting (LGB) [5], Extreme gradient boosting (XGBoost) [6] and simple Recurrent Neural Network (RNN) model. Table 1 shows the results from our experiments. We could see that our proposed approach outperformed the whole machine learning based methods, achieved 66.8% AUROC score and 13.1% AUPRC score. Comparing to the RNN model, our method does not show the significantly improvement with AUROC score. However, the comparison base on AUPRC score, which makes it possible to assess the performance of a classifier on the minority class, shows a remarkable improvement with 16% higher.

4. Conclusion

In this paper, we presented a novel method implemented for Rapid Response System that detects the in-hospital abnormal status of patients based on basic vital signs and blood test values. The application of the attention tabular learning approach shows the potential improvement of the whole benchmark machine learning systems and early warning score. Based on that, our future goal is to extend the system by implementing other deep learning models to achieve better results.

Acknowledgment

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) & funded by the Korean government (MSIT) (NRF-2019M3E5D1A02067961), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT). (NRF-2020R1A4A1019191) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1I1A3A04036408).

References

- [1] M. Chirag, F. Wu, and V. Karen, "Modified early warning system improves patient safety and clinical outcomes in an academic community hospital," *Journal of Community Hospital Internal Medicine Perspectives*, vol. 5, 2015.
- [2] J. Kwon, Y. Lee, Y. Lee, J. Park, and S. Lee, "An algorithm based on deep learning for predicting in-

hospital cardiac arrest," *Journal of the American Heart Association*, vol. 7, no. 13, pp. 1–11, 2018.

- [3] V. Ashish, S. Noam, and P. Nikki, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 2017-December, pp. 5999–6009, 2017.
- [4] S. O. Arik and T. Pfister, "Tabnet: Attentive interpretable tabular learning," *CoRR*, vol. abs/1908.07442, 2019. [Online]. Available: <http://arxiv.org/abs/1908.07442>
- [5] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. - Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017.
- [6] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *CoRR*, vol. abs/1603.02754, 2016. [Online]. Available: <http://arxiv.org/abs/1603.02754>
- [7] Shamout FE, Zhu T, Sharma P, Watkinson PJ, Clifton DA. "Deep Interpretable Early Warning System for the Detection of Clinical Deterioration." *IEEE J Biomed Health Inform.* 2020 Feb;24(2):437-446. doi: 10.1109/JBHI.2019.2937803. Epub 2019 Sep 19. PMID: 31545746.