# Using Machine Learning Algorithms in Medication for Cardiac Arrest

# Early Warning System Construction and Forecasting

Hsiao-Ko Chang
Department of Computer Science
and Information Engineering,
National Taiwan University
Taipei, Taiwan, R.O.C
d06922027@ntu.edu.tw

Cheng-Tse Wu
Department of Computer Science
and Information Engineering,
National Taiwan University
Taipei, Taiwan, R.O.C
d02922011@ntu.edu.tw

Ji-Han Liu
Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei, Taiwan, R.O.C
d03944005@ntu.edu.tw

Jyh-Shing Roger Jang
Department of Computer Science
and Information Engineering,
National Taiwan University
Taipei, Taiwan, R.O.C
jang@csie.ntu.edu.tw

***Abstract***

**Target—**In this paper, we focus on using medicine for patients who have cardiac arrest then must have to do Cardiopulmonary Resuscitation (CPR). We want to know the medicine influence in predicting state of an illness deterioration. Therefore, we proposes a Medication for Cardiac Arrest Early Warning System (MCAEWS). It's not only assist physicians to early diagnose of an illness and immediately warning, but also increase sensitivity, decrease false positive rate and mortality rate. The most important role is greatly improve medical quality.

**Methods—**In this study, the data is from the emergency department of National Taiwan University Hospital (NTUH). It is from January 2014 to December 2015. The patients who stayed in the emergency detention area for more than six hours during this two years. The patients were included in the retrospective cohort study. To comparative measures for the machine learning models, we used such as the Area Under the Receiver Operating Characteristic Curve (AUROC) and the Area under the Precision-Recall Curve (AUPRC).

**Results—**The data were analyzed for CPR and non-CPR groups respectively. Furthermore, we evaluated sensitivity and specificity. The Random Forest Algorithm (AUC: 0.98; AUP: 0.23) compare with others such as Logistic Regression Algorithm (AUC: 0.94; AUP: 0.13), Decision Tree (AUC: 0.97; AUP: 0.05), and Extreme Random Tree (AUC: 0.91; AUP: 0.08), it was significantly high performance.

**Conclusion—**Increasing the drug factors in vital signs, that it effectively improved the accuracy of predicting cardiac arrest. The results of this study, it's help for emergency clinical Physicians and hospital quality management will validly solve clinical medical resource allocation issues and improve medical quality through decision support systems.

**Keywords—** Deep Learning; Machine Learning; Cardiac Arrest; Cardiopulmonary Resuscitation; Electronic Health Record; Artificial Intelligence

## I. Introduction

**Research Background and Motivation**

Nowadays, in the information explosion period, the innovation and intelligence of various technology will provide medical personnel with more accurate medical diagnosis and treatment. It is a common clinical issue and dilemma to focus on the personalized disease prevention, assessment, diagnosis, treatment and rehabilitation care programs.

Therefore, we use the clinical practice situation of the National Taiwan University Hospital's Medical System as a research field, and develops a comprehensive early warning system for cardiac arrest in the emergency medical system. Using more precision and artificial intelligence (AI), to create a clinical decision support system. It will help the physicians to determine the exacerbation and immediate early warning, then effectively help patients to get assist, treat and achieve effective treatment opportunities. This study is expected to improve the quality of medical care, reduce medical costs and develop the auxiliary tools that are the basis for physicians diagonsis.

**Research Issues and Objectives**

The major medical institutions in Taiwan, whether it is a teaching hospital or a community hospital, the emergency department is the primary key process for emergency medical treatment of patients. The triage is in order to strengthen the treatment of patients with different priorities and related interventions. However, its reduction in the allocation and utilization of medical resources, that seems to be insufficient and waste in medical institutions.

The original Early Warning Score system (EWS) was designed to use major related vital signs as a key indicator of early detection abnormalities, before it worsens to severe disease. In the later stage, a Modified Early Warning Score system (MEWS) which mainly modified the relevant physiological parameters for scoring, and has been proved to be a tool for effectively predicting state of an illness deterioration.

In order to increase the sensitivity, decrease false positive rate and mortality. It was to extract characteristic factors of major vital signs (such as blood pressure, heart rate, respiration, body temperature, etc.) and add drug factors. It was developed a machine learning algorithm based on Medication for Cardiac Arrest Early Warning System (MCAEWS). The experimental results effectively increase predictive performance.

**Research Methods and Processes**

This research method is mainly focus on machine learning algorithms. It is the most popular in the field of artificial intelligence and computer science. In order to compare the traditional statistical methods, we experimented by four methods as comparison benchmarks: Decision Tree, Logistic Regression, Random Forests, Extremely Random Trees. In the future, we will use deep learning algorithms in this article. The research process is shown in Figure 1 below:

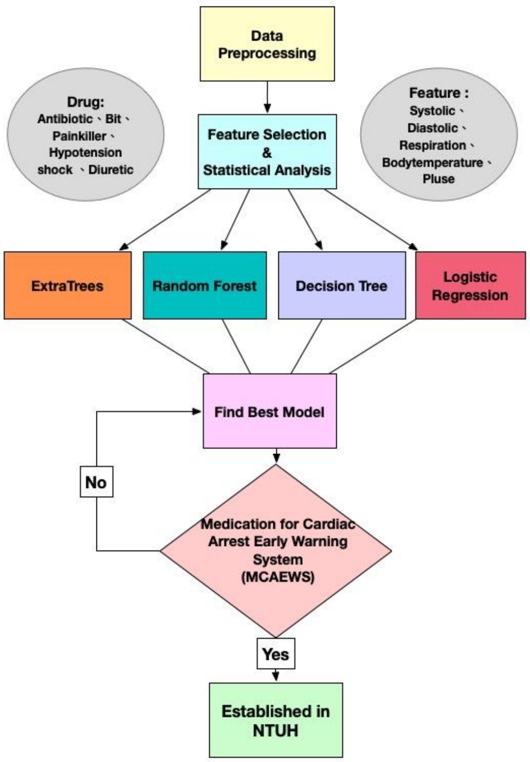


Figure 1: Experimental research flow chart

## II. LITERATURE REVIEW

This study mainly discusses about the early warning score system in the past of medical emergency. In order to understand the differences between the methods and prediction accuracy proposed by all scholars then we start to analyze and compare this method and data. At this part, we discuss two major relevant literatures: The Early Warning Score System and the Machine/Deep Learning Application. The development direction of these theoretical foundations then through the integration of literature at different levels. It will stimulate the main axis of the whole research.

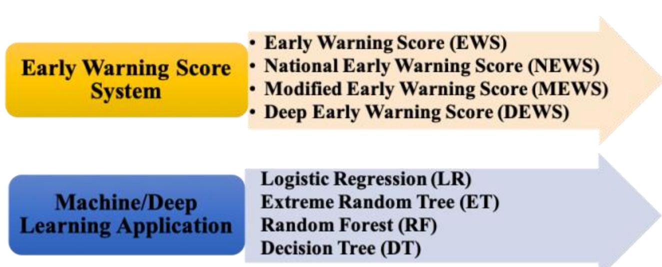


Figure 2: Schematic diagram of the literature

### A. *Early Warning Score System*

The medical Early Warning Score System aims to analyze and early warning nurses and physicians through the results of patient's vital signs detection and residence time in the hospital. If the patient has a sudden situation, they can early detection and treatment then avoid irreparable results. There is an auxiliary system to achieve effective resource utilization and distribution, then it can early intervention in the patient's exacerbation, eliminating unnecessary waste and time. In the past research was about predicting sudden Cardiac Arrest (CA), the accuracy of prediction has been stagnant. If it can through machine learning and deep learning algorithms, will effectively improve the accuracy of prediction and early prevention.

The hospitalized patients were analyzed and predicted according to Vital Signs(VS). In addition to cardiac arrest, death and transfer to the Intensive Care Unit (ICU) were included in the observational conditions. Most observations find that retrieval vital signs information can effectively increase the forecast rate.

Churpek et al. [3] used the vital signs of patient measurement results to investigate increasing the accuracy of using vital sign factors. The data is from five hospitals during the five-years period to compare the accuracy of different methods using Area Under the Receiver Operating Characteristic Curve (AUC). This research has important implications for development of early warning systems. Green et al. [4] used BTF to adjust the MEWS, NEWS, and e-CART scoring criteria and found that e-CART scores generated using electronic medical records were more accurate than manual paper observation tools. It can be used to predict cardiac arrest, ICU metastasis, and death within 24 hours.

### B. *Machine / Deep Learning Application*

Most medical information application research uses machine learning technology to make predictions, which has a substantial improvement in accuracy, but few methods are used to explore related models. In order to achieve the absolute effect of precision medicine, it creates a new clinical early warning scoring system to achieve better prediction accuracy. Clifton et al. [2] have proposed using Gaussian Process Regression (GPR) methods to improve existing EWS. Churpek et al. [1] compared with various machine learning methods to improve the accuracy of clinical deterioration in the detection ward for large and multi-center databases. The random forest algorithm is the best accuracy.

The newest research, Kwon et al. [5] proposes an early warning scoring system based on deep learning. Using the Recurrent Neural Network(RNN) for training to effectively improve sensitivity and reduce false positives rate. It can be used for the detection of patients with cardiac arrest. Chen Lin et al. [7] proposed to use LSTM and capture the time dependence of time series data.The CNN is used to extract important feature and predict the early diagnosis of septic shock.

## III. METHODS

### A. *Clinical data collection*

This study uses the National Taiwan University Hospital's medical system as the research situation field. The data source is from the emergency department of National Taiwan University Hospital. The data collection period is from 2014/01/01 to 2015/12/31, and the patient stays in the emergency department for more than 6 hours. This research database includes emergency codebook, medication, inspection, emergency patients information, CPR patients information, physicians diagnostic code information, vital signs measurement information, medical orders and so on.

In this study, the patients will have CPR to carry out the classification of the first stage. After pre-processing, including abnormal data cleaning, filling in missing values, etc. The post-organized data will be trained and analyzed in the second stage. Four machine learning algorithms are used for training validation and subsequent analysis. The current selection methods: Decision Tree, Logistic Regression, Random Forest, and Extreme Random Tree.

### *B. Data preprocessing*

The early period data were all manually input, so it has many missing and errors. First of all, there is a duplicate problem for same patients during the same registration time, and it will be retrieved a piece of data. If the age data has abnormal values, it will be replaced by the average value. If there are missing values for multiple heights and weights, the average value will be used to fill it. Then same patients, same types, same items, same recording times, and same values were selected a piece of data. For measurements without a clear measurement interval, then we used within 30 minutes.

For the abnormal values of various vital signs, human error correction will be performed. For example, the blood pressure value should not exceed 300, and the respiratory value should be between 10 to 30 (the data should be between 10 to 60 in non-CPR). The pulse value should be within 300 and the body temperature should be between 28 to 42. After sorting the CPR and non-CPR categories, the patients who has immediately CPR within 1 hour were deleted. The total sample data was 43,445 with non-CPR patients, and 124 with CPR patients.

### *C. Feature selection method*

This study select 15 important factors: sex, age, fever, height, weight, systolic, diastolic, pulse, respiration, body-temperature, etc. In order to include drug factors, consider the five drugs that affect the change in vital signs: bit, painkiller, antibiotic, diuretic, Hypotension shock, and vital signs difference value, a total of 45 features were selected.

Because of the best features are selected from the original features, it will let the recognition rate can reach the highest value. These features with better discriminative ability can simplify the classifier calculation and understand the causal relationship of the classification problem. Therefore, this study we used heuristic methods, in which the Sequential forward selection (SFS) proposed by Whitney in 1971 was used as the method of feature selection. The steps are as follows:

1. Use the nearest neighbor rule (KNNC) and the leave-one-out (LOO) recognition rate prediction method.
2. The first selected feature must be the feature with the highest recognition rate.
3. The next selected feature must be the one with the highest recognition rate after merging with the originally selected features.
4. Repeat step 3 until all features are selected.

Because the hypotension shock medication is necessary for patients with CPR, so the drug and its difference value are deleted, then after the above steps, a total of 23 important features are obtained.

### *D. Drug difference value processing*

We focused on the drugs extraction, and aims to process drug with the five vital signs of systolic blood pressure, diastolic blood pressure, respiration, body temperature and pulse. According to the database, the patient stayed more than 6 hours, so the measurement of vital signs will also have a cyclical measurement. Therefore, the drug difference value = (the first measurement after the time of the drug — the first measurement after admission).

According to whether the patient applies these four drugs to become the characteristic value of "Boolean number", and then take the measurement value of the vital signs after the first time of each drug to perform the subsequent difference value processing. The analysis results of the drug difference value for CPR patients and non-CPR patients are shown in Fig. 3 and Fig. 4 below. The four drugs that are applied to CPR patients have the most obvious vital signs: systolic blood pressure, diastolic blood pressure, and pulse. It is consistent with the features proposed in the past about the Early Warning Scoring System (EWS) literature. It also has its significance in clinical practice.

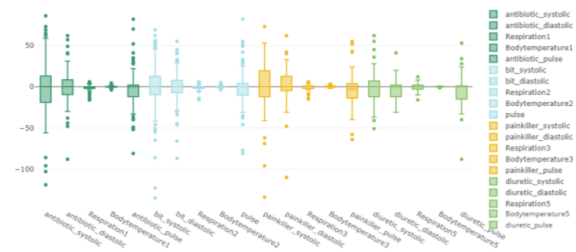

Figure 3. Difference analysis of drug vital signs in patients with CPR

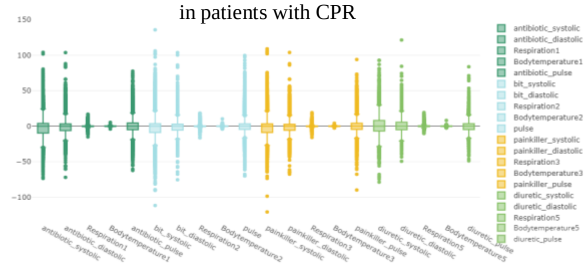

Figure 4. Difference analysis of drug vital signs in patients with non-CPR

### *E. Machine learning methods*

This study will use four machine learning methods to train and validate the model's fit respectively: logistic regression, decision tree, random forest, and extremely random tree. Machine learning has its definition: machine learning is a science of artificial intelligence, mainly researching objects as artificial intelligence, and how to improve the performance of specific algorithms in empirical learning.

There are several categories of machine learning: supervised learning, semi-supervised learning, and unsupervised learning. The four methods in this study are supervised learning classification methods. A model is trained for a given training data set, and new data is used to predict the results based on this model. The target outcome of the training data set is marked by human. Therefore, this study is based on the emergency medical information of National Taiwan University Hospital, and it's target are given by doctors based on clinical experience, then it use the computer to performs subsequent training and calculate the results.

## IV. RESULTS

According to the emergency database of National Taiwan University Hospital, 15 characteristics were selected, and the difference between the five influence drugs and their vital sign characteristics was added by the clinician. After the feature selection method, 23 features were selected and total sample data: 43,445 non-CPR patients and 124 CPR patients. 70% of which were used as training data sets and 30% of data were used as test data sets. Among of them, 10% were randomly selected from 70% training data as a validation data set. And draw area under the receiver operating characteristic curve (AUROC) and the area under the precision recall curve (AUPRC) as the benchmark for comparison. It shown in Figure 5 and Figure 6 below:

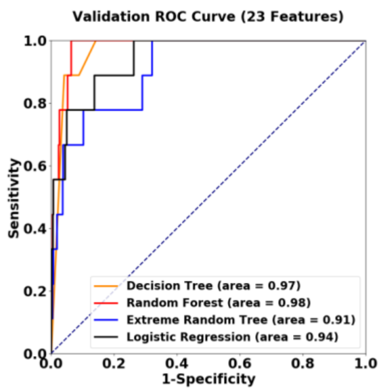


Figure 5. ROC Curve for Four Methods

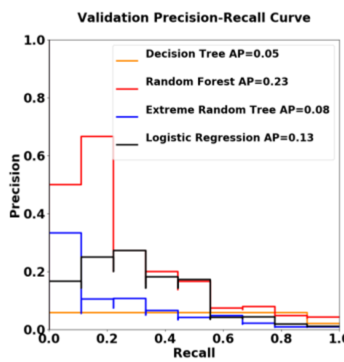


Figure 6. PR Curve for Four Methods

As seen from the above Figures 5 and 6, the AUC area of the logistic regression is 0.94, the AUP area is 0.13; the AUC area of the decision tree is 0.97, the AUP area is 0.05; the AUC area of the random forest is 0.98, and the AUP area is 0.23; The AUC area of the extreme tree is 0.91 and the AUP area is 0.08. The F1 Score of the four methods calculated according to the experimental results is shown in Table 3 below. The results can be seen from Table 3 that the random forest algorithm has the best model verification results.

Table 3 F1 Score of the four methods

| $F_1$ Score | | | |
|---|---|---|---|
| **Decision Tree** | **0.9108** | **Extreme Random Tree** | **0.8124** |
| **Random Forest** | **0.9672** | **Logistic regression** | **0.8036** |

## V. Conclusion and Discussion

The results of this study show that adding the drug factors as feature can effectively improve the prediction accuracy, and the four methods results that the best model is random forest. However, there are two important research restrictions in this study. On the one hand, in addition to the application for IRB, the collection of data from National Taiwan University Hospital is not easy due to the privacy protection issue. On the other hand, it is the emergency data for predicting cardiac arrest, the predictive indicator of its deterioration, and the imbalance of data is also a serious situation, which is a major limitation in medical research. In order to solve the above problems, the clinician will be requested to apply for more data sets to verify the appropriateness of the early warning score system, and to amplify the data collection with a relatively fewer of data. In order to meet the medical clinical common sense, and to achieve the balance of the data.

Training and prediction will be done with newly algorithms, especially the time series will be incorporated into the training objectives, and the deep learning technology convolutional neural network will be used with the long and short-term memory models to expect more effective improvement of the prediction results.

## Acknowledgments


The authors would like to express their sincere gratitude to the anonymous reviewers for their constructive comments on the manuscript of this paper. The work presented in this paper was supported in part by the Ministry of Science and Technology, Taiwan, R.O.C., under grant number (MOST 107-2634-F-002-015).


## References


[1] Matthew M Churpek, Trevor C Yuen, Christopher Winslow, David O Meltzer, Michael W Kattan, and Dana P Edelson; Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards; Crit Care Med. 2016 February ; 44(2): 368–374.

[2] Clifton L, Clifton DA, Pimentel MA, Watkinson PJ, Tarassenko L.;Gaussian Process Regression in Vital-Sign Early Warning Systems; 34th Annual International Conference of the IEEE EMBS San Diego, California USA, 28 August - 1 September, 2012.

[3] Matthew M. Churpek, Richa Adhikari, Dana P. Edelson; The value of vital sign trends for detecting clinical deterioration on the wards; Resuscitation Volume 102, May 2016, Pages 1-5.

[4] Malcolm Green, Harvey Lander, Ashley Snyder, Paul Hudson, Matthew Churpek, Dana Edelson; Comparison of the Between the Flags calling criteria to the MEWS,NEWS and the electronic Cardiac Arrest Risk Triage (eCART) score forthe identification of deteriorating ward patients; Resuscitation Volume 123, February 2018, Pages 86-91.

[5] Joon-myoung Kwon, Youngnam Lee, Yeha Lee, Seungwoo Lee, Jinsik Park; An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest; Journal of the American Heart Association. Volume 7,Issue 13, July3, 2018.

[6] Isao Nishijima, Shouhei Oyadomari, Shuuto Maedomari, Risa Toma, Chisato Igei, Shinya Kobata, Jyun Koyama, Ryuichiro Tomori, Natsuki Kawamitsu, Yoshiki Yamamoto, Masafumi Tsuchida, Yoshihiro Tokeshi, Ryo Ikemura, Kazufumi Miyagi, Koichi Okiyama and Kiyoshi Iha; Use of a modified early warning score system to reduce the rate of in-hospital cardiac arrest; Journal of Intensive Care (2016) 4:12.

[7] Chen Lin, Yuan Zhang, Julie Ivy, Muge Capan, Ryan Arnold, Jeanne M. Huddleston, and Min Chi; Early Diagnosis and Prediction of Sepsis Shock by Combining Static and Dynamic Information using Convolutional-LSTM. IEEE International Conference on Healthcare Informatics, 2018. pp.219-228.