Artificial Intelligence Model for the Prediction of Cardiac Arrests Using Time-Series Biometric Data

AI ZEN
AIIS Artificial Intelligence Institute Seoul National University

Jeongjun Heoh · Seonkyu Kim · Seong-gyu Woo

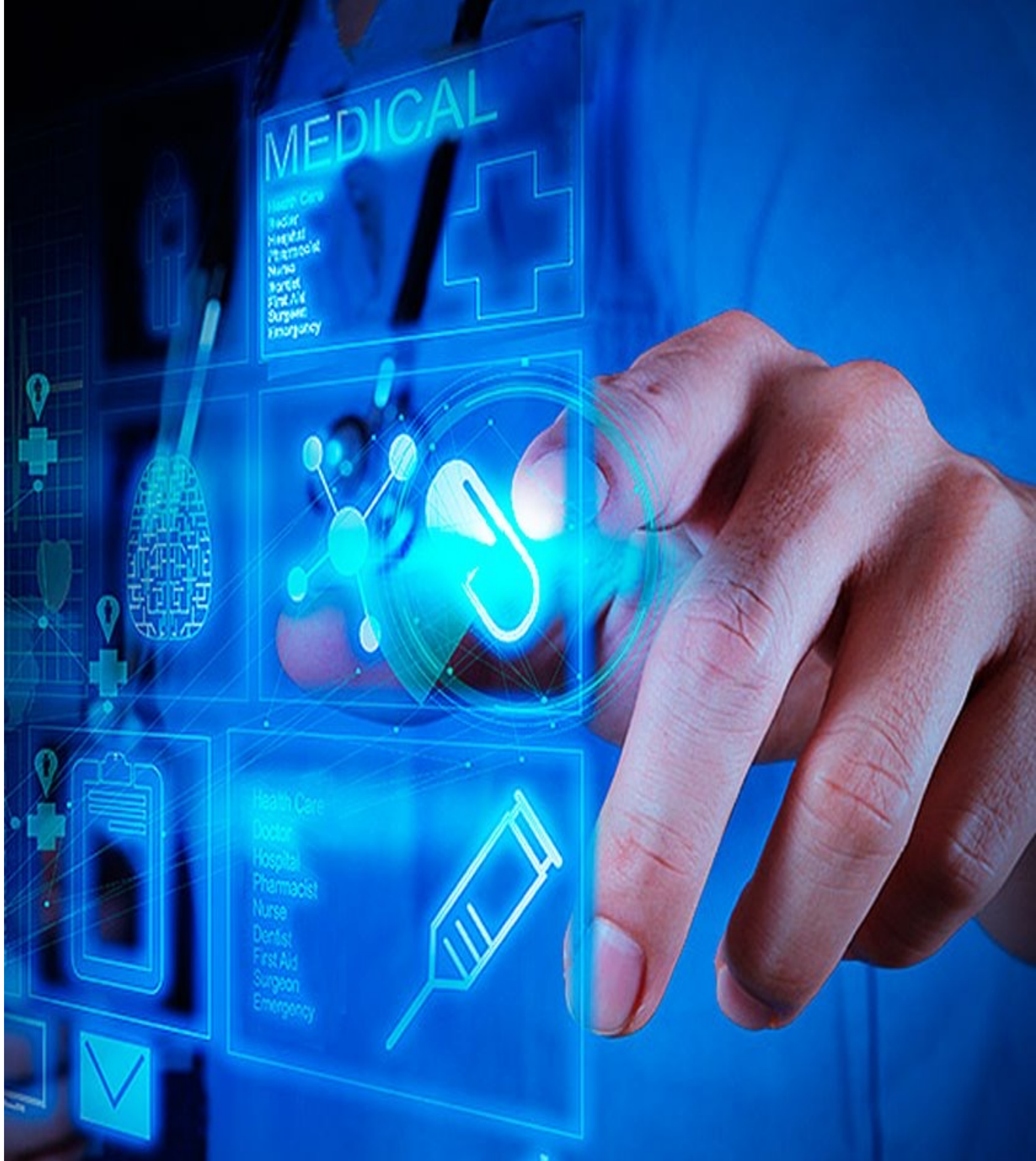
2020.7.3

Agenda

- Project Background
- Methods
- Model
- Results
- Implication

AI
ZEN

Project Background



Project Background

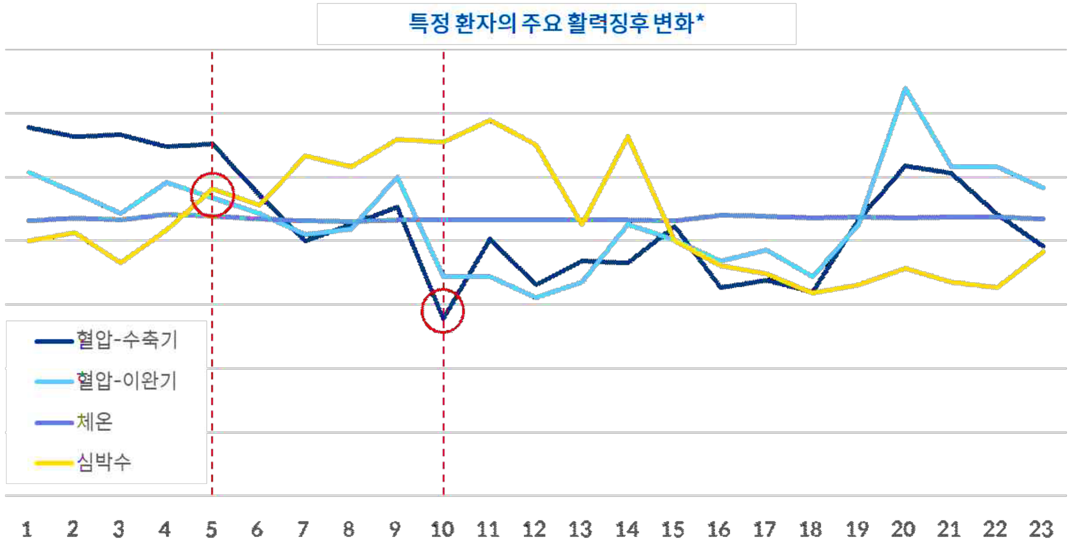
Project Outline

- Develop a python-based artificial intelligence model for predicting the risk points of critically ill patients.

- Integrate machine learning and deep learning algorithms into statistics to learn and apply changes in patterns, detecting patient risk.

- The purpose is to detect subtle abnormalities in biometrics data of critically ill patients so that medical staff can quickly correct the causative disease.

- Expected to secure survival rates and improve the quality of medical services by appropriate deployment of medical staff.

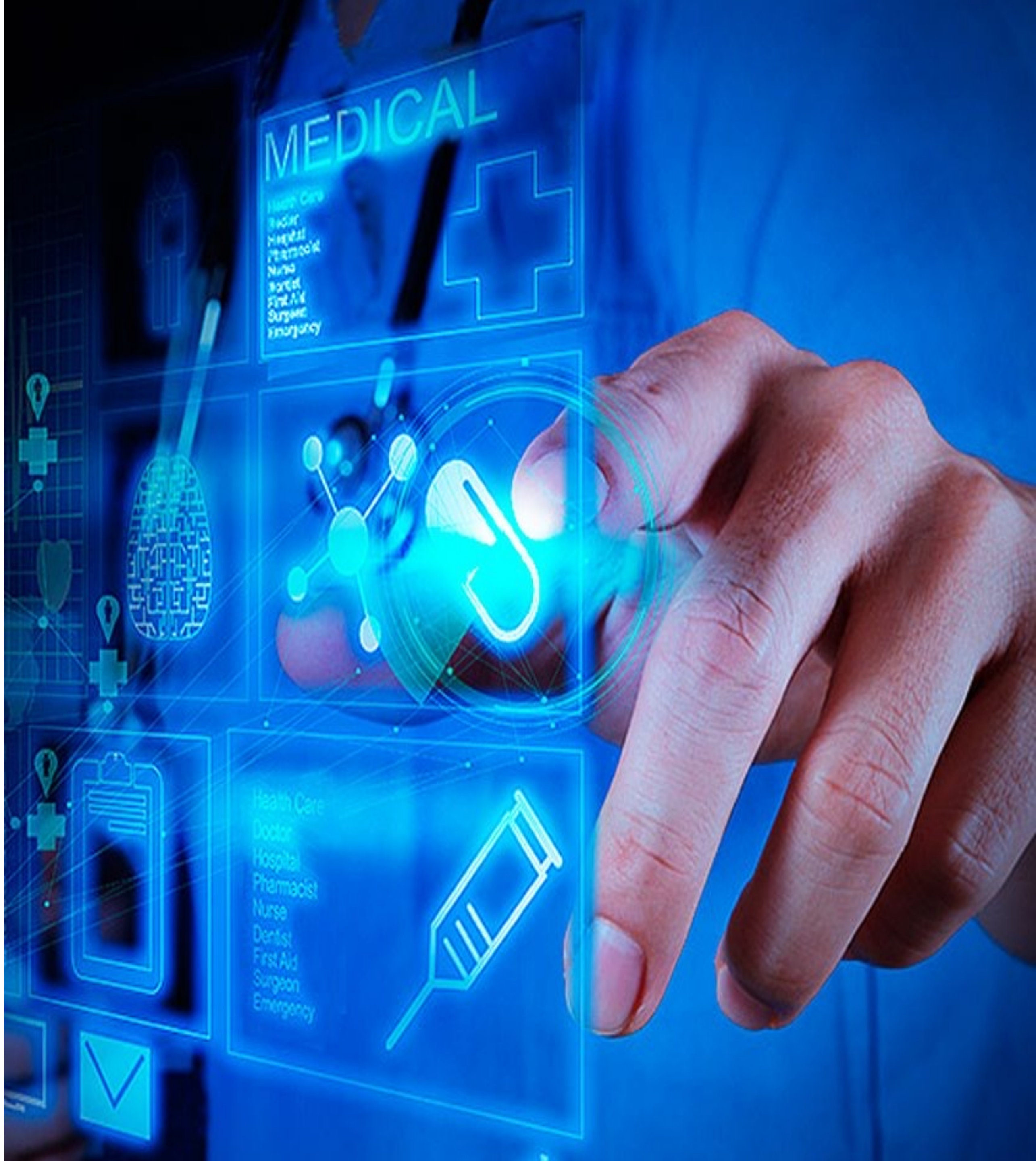


Project Background

Dataset

- Patient Data :
  - Sex, Age
- Vital Sign :
  - Systolic blood pressure, Body temperature, Pulse, Respiration, Oxygen saturation
- Blood Data :
  - WBC, Platelets, Hemoglobin, Creatinine, Glucose, Sodium
  - Potassium, Chlorine, Protein, Albumin, Bilirubin, Calcium, Lactic acid
  - CRP, ALP, AST, ALT, BUN

Methods


MEDICAL
Health Care
Doctor
Hospital
Pharmacist
Nurse
Dentist
First Aid
Surgeon
Emergency

Methods

Timeline

	May 2nd Week	May 3rd Week	May 4th Week	Jun 1st Week	Jun 2nd Week	Jun 3rd Week	Jun 4th Week	July 1st Week
Plan	Outline, Literature Review							
Develop		Preprocessing, Model Selection, Integrating Feedback, Model Development						
Modify							Refinement and Correction	

Methods

Problem Definition

Existing System : Early Warning Score (EWS)

- A scoring system based on multiple vital signs to quickly determine the level of diseases of the patient
- Low Recall / High False Alarm Rate

※ Recall : Probability of judging a dangerous case as dangerous
False Alarm Rate : Probability of judging a non-dangerous case as dangerous (=False Positive Rate)

- Increased burden among medical staff

<Modified Early Warning Score (MEWS)>

Score	**3**	**2**	**1**	**0**	**1**	**2**	**3**
Respiratory rate		<9			9-14	15-20	
O_2 Saturation	<90						
Body temp.		<35.1	35.1-36.5	36.5-37.5	<37.5		
Blood pressure	<70	70-80	81-100	101-200			
Pulse		<40	40-50	51-100	101-110	111-130	>130
AVPU				A	V	P	U

<Kwon JM, Lee Y, Lee Y, Lee S, Park J. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest. J Am Heart Assoc 2018;7>

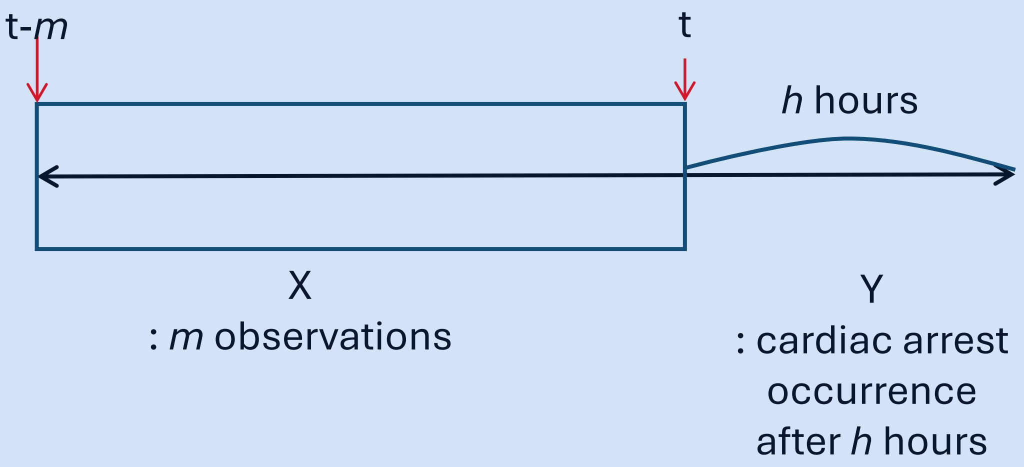
Methods

Problem Redefinition

Statistical Analysis

1) Basic stats
2) Regression to identify the relationship among variables
3) Extracting meaningful features

Goal Statement



Goal

1) Future risk detection
2) Weighted type 1 & 2 errors
3) High Recall / Low False Alarm Rate

Methods

Data Structure

Risk Group Patient Data

Time series vital sign data of patient receiving CPR or tracheal intubation

Event Data

CPR or tracheal intubation occurrence time data

Blood Data

Data on 1 to 4 batches of blood measured from patient

Control Group Patient Data

Time series vital sign data of patient without CPR or tracheal intubation

Blood Data

Data on 1 to 4 batches of blood measured from patient

Methods

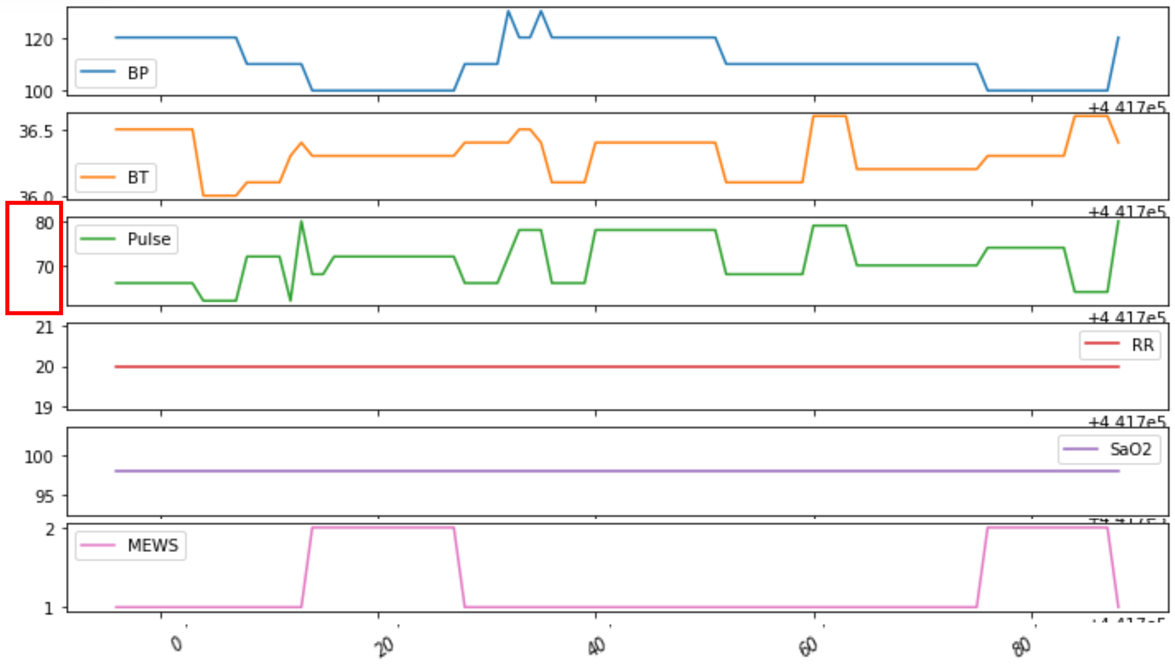
Data Structure

- Data in the train set and the test set are patient data from different hospitals.
- Since the data per patient is relatively short, approximately 2 to 5 days, it is appropriate to analyze data in hours rather than days.

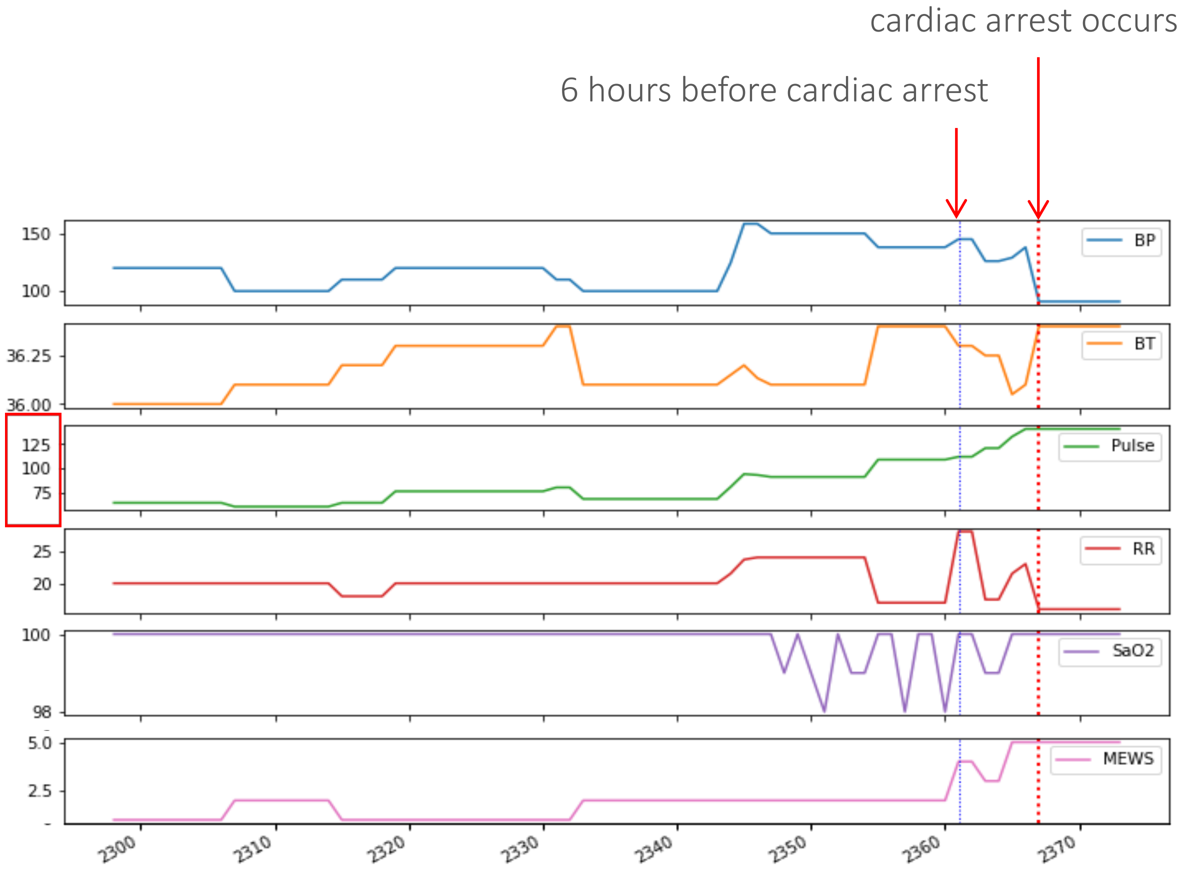
	Train (Hospital A)	**Test (Hospital B)**
Time	2017.01. ~ 2019.03.	2019.01. ~ 2019.05.
Number of Patients (Data size)	4,816 (161,710)	2,588 (103,565)
Proportion of Risk Group Patient	7.2%	2.4%
Measurement Period		
2 Days	3(91)	0(0)
3 Days	36(1527)	7(341)
4 Days	306(21266)	54(3653)
5 Days	4471(138826)	2527(99571)
Sex Ratio (M:F)	65.1% : 34.9%	56.7% : 43.3%
Age	65.7± 12.8	67.6±16.0
Min	2.2	16.2
Max	102.7	98.7

Methods

Exploratory Data Analysis (Chart)



Patient A in Control Group (54 years old, female)



Patient B in RIsk Group (54 years old, female)

- This is a case in which a 'significant increase in pulse to over 130' appears to be closely related to the occurrence of cardiac arrest.

Methods

Missing Values

- The main reasons for missing values are problems with EMR recording devices and differences in rounding methods on general wards.
- Body temperature, pulse, and respiration are often measured simultaneously, and oxygen saturation is often measured using a separate device.
- Rather than replacing or removing missing values, we want to convert them into dummy variables and utilize them.

	$\mathbf{d}$	$\mathbf{d} * \mathbf{X}$	$\Delta\mathbf{d} * \mathbf{X}$	$\Delta t1$	$\Delta t2$
Meaning	Dummy Variable	d * (Measured Value)	Difference between previous (d * X) values	(Current Time) – (Previous Existing Value Time)	(Current Time) – (Previous Data Time)
Value	Missing Value →0 Existing Value→1	Missing Value →0 Existing Value→X		$t_i - t_{i-n}$	$t_i - t_{i-1}$

Methods

Missing Value Conversion Example

				X				
20180120	1730	Sex	Age	BP	BT	Pul	Res	SaO2
20180120	1745	M	58.9	115	36.7	64	30	#N/A
20180120	1800	M	58.9	116	36.6	64	20	97
20180120	1900	M	58.9	#N/A	#N/A	#N/A	#N/A	98
20180120	2000	M	58.9	#N/A	#N/A	#N/A	#N/A	95
20180120	2100	M	58.9	#N/A	#N/A	#N/A	#N/A	95
20180120	2200	M	58.9	#N/A	#N/A	#N/A	#N/A	92
20180120	2215	M	58.9	122	36.6	62	30	#N/A
20180120	2230	M	58.9	125	36.7	62	28	#N/A
20180120	2300	M	58.9	#N/A	#N/A	#N/A	#N/A	97

← Before

↓ After

		d					d*X					Δd*X					Δt1					Δt2
Sex	Age	BP	BT	Pul	Res	SaO2	BP	BT	Pul	Res	SaO2	BP	BT	Pul	Res	SaO2	BP	BT	Pul	Res	SaO2	Time
1	58.9	1	1	1	1	0	115	36.7	64	30	0	0	0	0	0	0	0	0	0	0	0	0
1	58.9	1	1	1	1	1	116	36.6	64	20	97	1	-0.1	0	-10	0	0.25	0.25	0.25	0.25	0	0.25
1	58.9	0	0	0	0	1	0	0	0	0	98	0	0	0	0	1	0	0	0	0	1	1
1	58.9	0	0	0	0	1	0	0	0	0	95	0	0	0	0	-3	0	0	0	0	1	1
1	58.9	0	0	0	0	1	0	0	0	0	95	0	0	0	0	0	0	0	0	0	1	1
1	58.9	0	0	0	0	1	0	0	0	0	92	0	0	0	0	-3	0	0	0	0	1	1
1	58.9	1	1	1	1	0	122	36.6	62	30	0	6	0	-2	10	0	4.25	4.25	4.25	4.25	0	0.25
1	58.9	1	1	1	1	0	125	36.7	62	28	0	3	0.1	0	-2	0	0.25	0.25	0.25	0.25	0	0.25
1	58.9	0	0	0	0	1	0	0	0	0	97	0	0	0	0	5	0	0	0	0	1	0.5

Methods

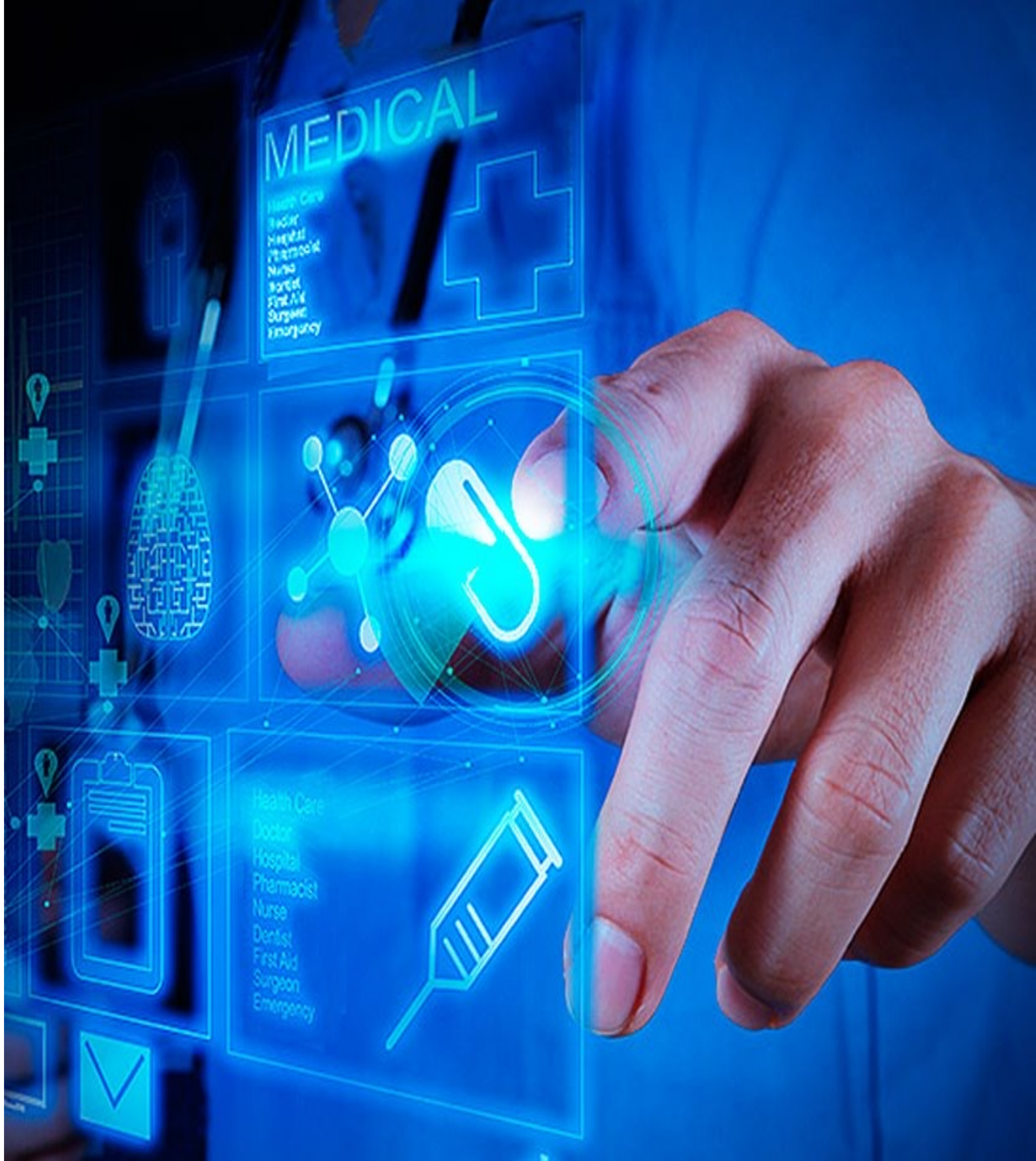
Regression

- The newly created variables show some valid values
- For example, pulse shows a significant relationship with the occurrence of cardiac arrest.

	coef	std err	t	P>|t|	[0.025	0.975]
Sex	0.0011	0.001	2.174	0.030	0.000	0.002
Age	0.0006	0.001	0.495	0.620	-0.002	0.003
Blood Pressure_exist	-0.0057	0.002	-3.057	0.002	-0.009	-0.002
Body Temp_exist	0.0007	0.002	0.409	0.682	-0.003	0.004
Pulse_exist	-0.0142	0.003	-4.901	0.000	-0.020	-0.008
Respiration_exist	-0.0320	0.003	-10.681	0.000	-0.038	-0.026
SaO2_exist	0.0647	0.004	16.909	0.000	0.057	0.072
Blood Pressure	0.0071	0.003	2.798	0.005	0.002	0.012
Body Temp	-0.0012	0.004	-0.292	0.770	-0.010	0.007
Pulse	0.0476	0.003	14.133	0.000	0.041	0.054
Respiration	0.0860	0.004	20.954	0.000	0.078	0.094
SaO2	-0.0731	0.004	-17.618	0.000	-0.081	-0.065
Blood Pressure_change	3.728e-05	0.001	0.032	0.974	-0.002	0.002
Body Temp_change	-0.0039	0.002	-2.176	0.030	-0.007	-0.000
Pulse_change	-0.0054	0.002	-2.933	0.003	-0.009	-0.002
Respiration_change	0.0069	0.001	5.817	0.000	0.005	0.009
SaO2_change	0.0055	0.002	3.514	0.000	0.002	0.009
TimeD_Blood Pressure	0.0384	0.010	3.690	0.000	0.018	0.059
TimeD_Body Temp	0.0554	0.014	3.922	0.000	0.028	0.083
TimeD_ Pulse	-0.0761	0.025	-3.028	0.002	-0.125	-0.027
TimeD_ Respiration	0.0372	0.023	1.618	0.106	-0.008	0.082
TimeD_ SaO2	-0.0248	0.002	-14.330	0.000	-0.028	-0.021
TimeDelta	-0.0602	0.004	-14.161	0.000	-0.068	-0.052

AI ZEN

Model

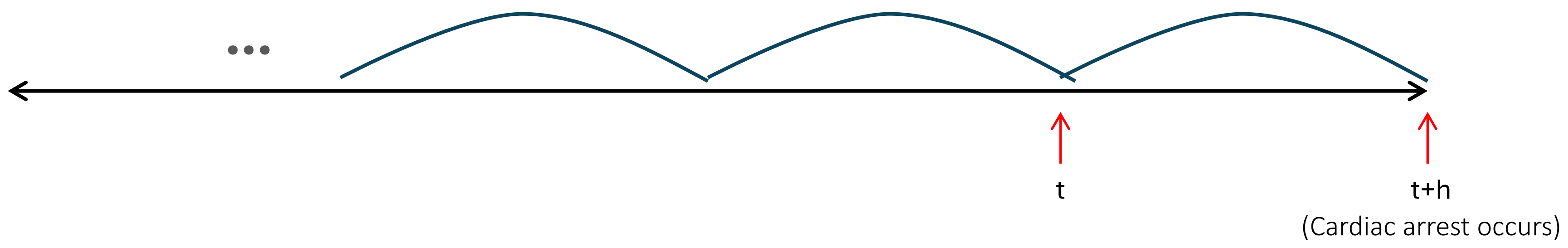

MEDICAL
Health Care
Doctor
Hospital
Pharmacist
Nurse
Dentist
First Aid
Surgeon
Emergency

Model

There is a trade-off in how many hours (h) we need to predict.

- Predictions for the near future → Performance ↑, Practicality ↓
- Predictions for the distant future → Practicality ↑ , Performance ↓

- Therefore, it is necessary to set the goal of how many hours (h) later to predict risk.



- To predict the future after *h* hours, *m* previous data is needed. At this time, *m* is searched as a parameter.

Model

Model Suggestion

- Two models are proposed depending on the purpose.

Predictive Model in 6 hours

- On average, doctors predict risk situations six hours in advance. Therefore, comparing with the current standard could be a good standard.

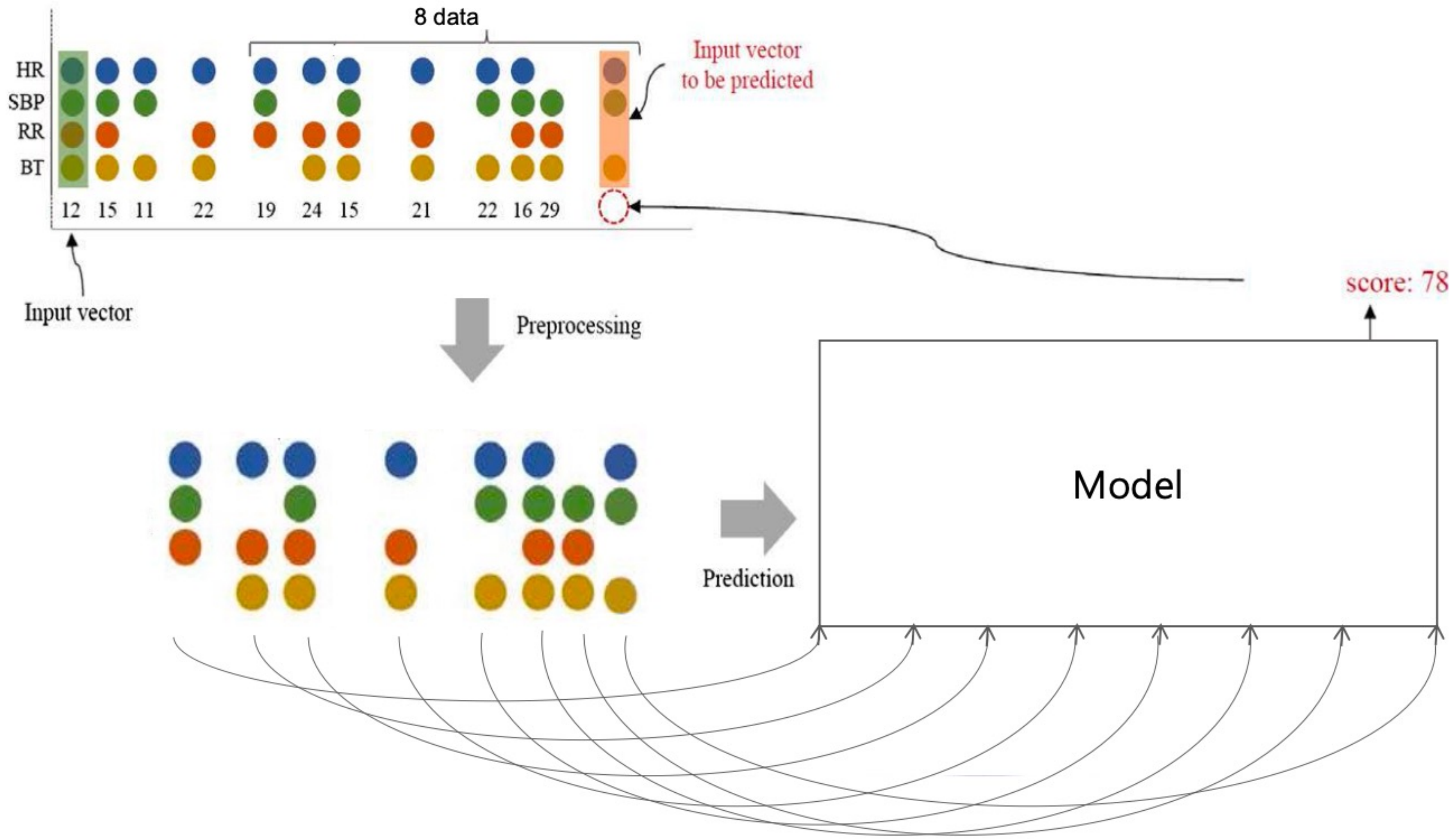
Predictive Model in 1 hour

- The short-term prediction will reduce the waiting time of emergency medical staff. The model can clearly show the difference between long-term and short-term forecasts.

- We can expect long-term and short-term forecasting effects, respectively.

Model

Input Method Example

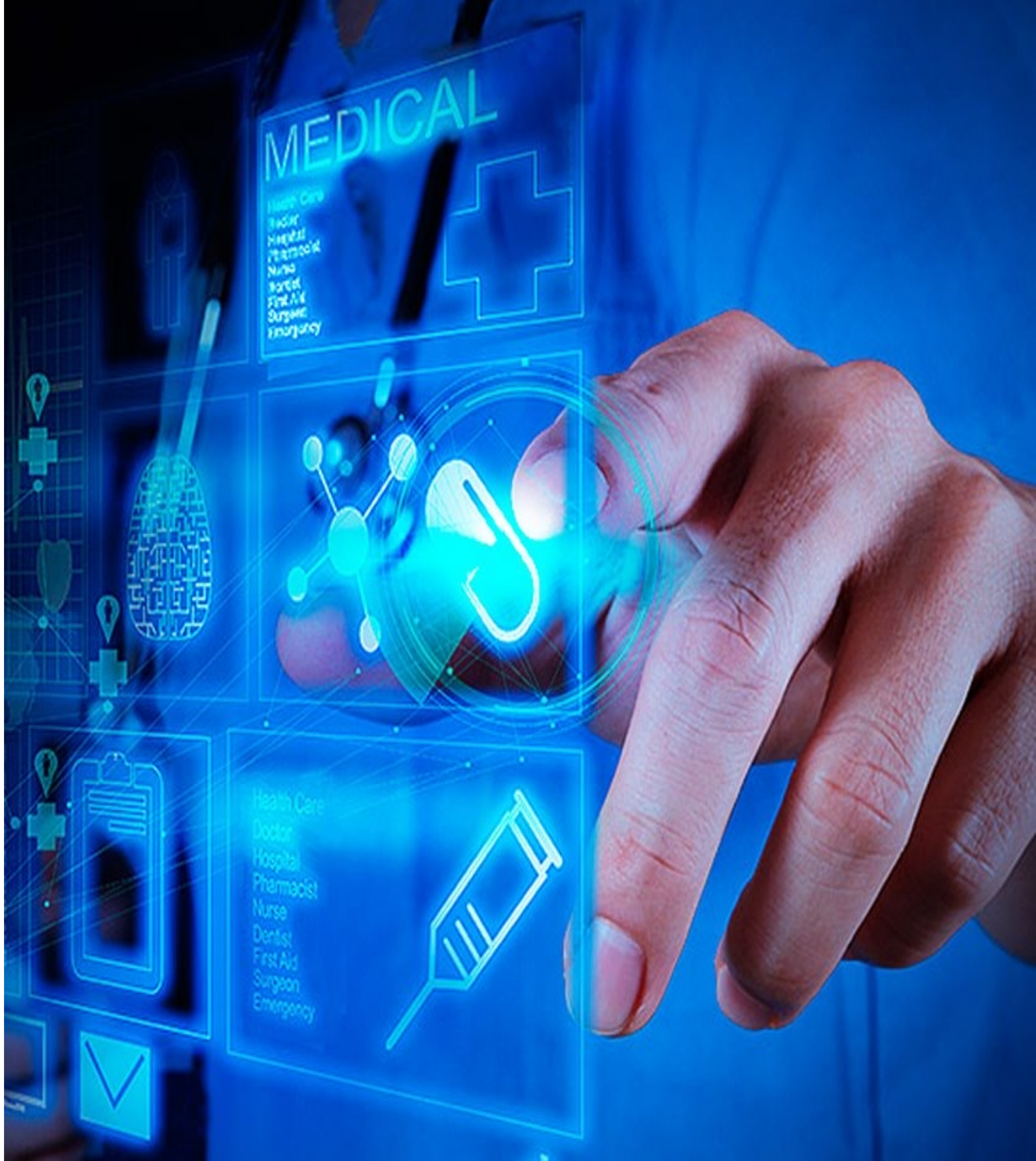


Model

Model Selection

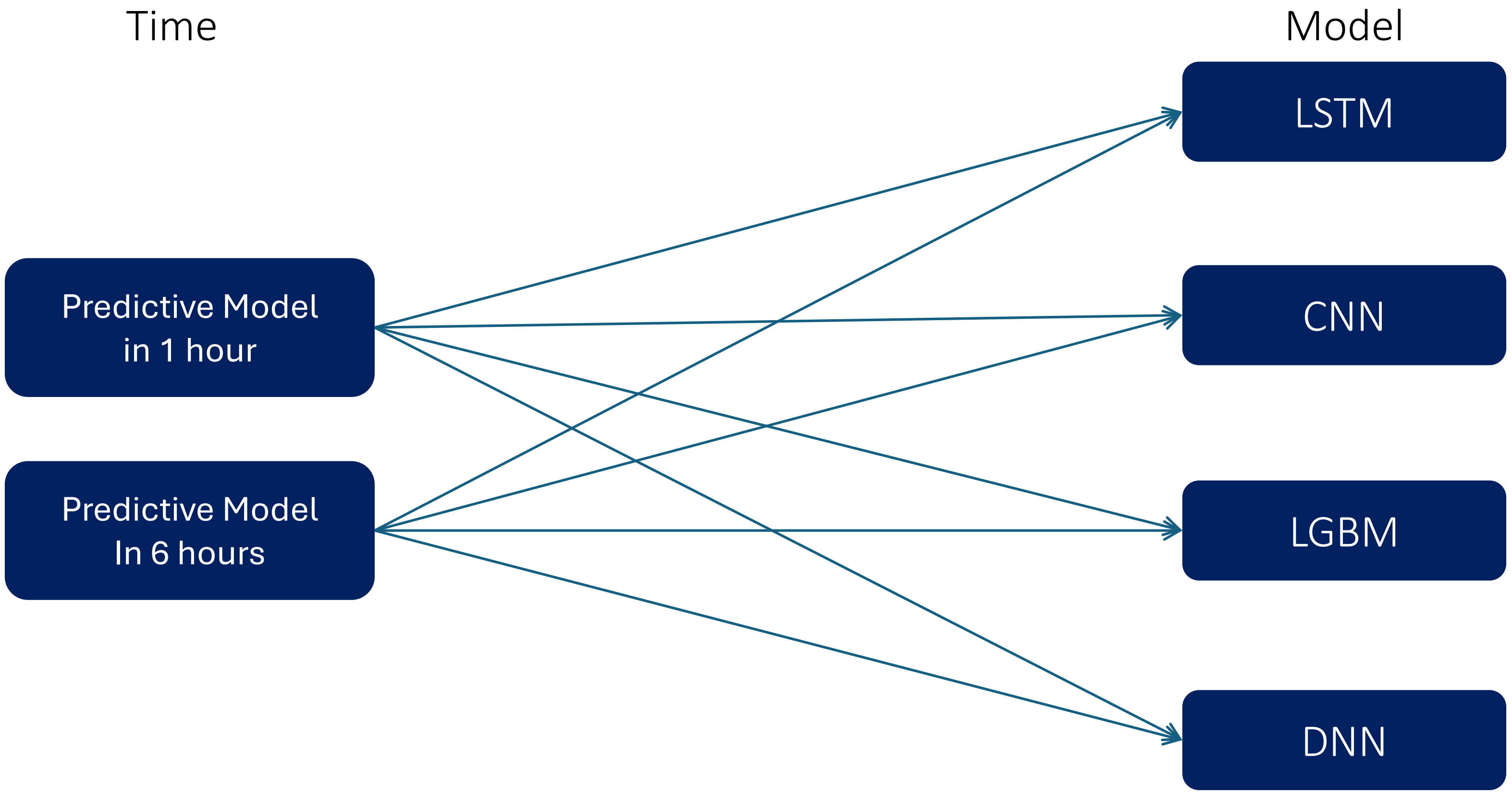
- CNN
  - CNN was designed to accommodate data as a time series using 1D convolution.
- LSTM
  - LSTM is ideal for entering the patient’s past eight consecutive data because it recognizes time series data well.
- DNN
  - DNN was included in the comparison as a basic deep-learning model, providing a benchmark for the other models' performance.
- LGBM
  - We constructed models using machine learning techniques and compared their performance.

Results


MEDICAL
Health Care
Doctor
Hospital
Pharmacist
Nurse
Dentist
First Aid
Surgeon
Emergency

Results

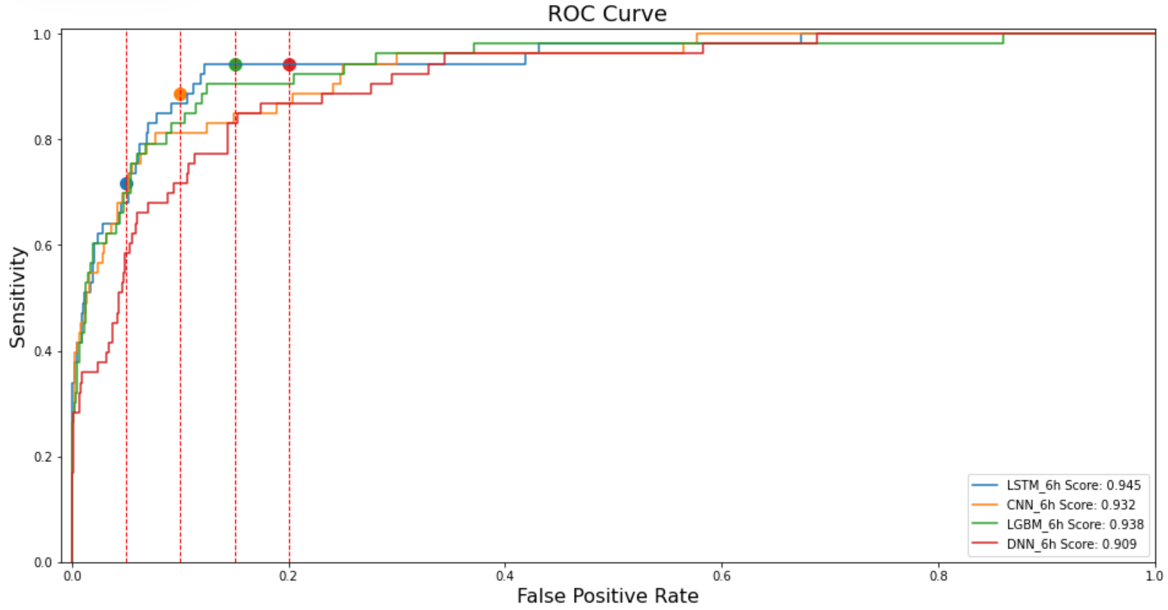
Number of Cases



Results

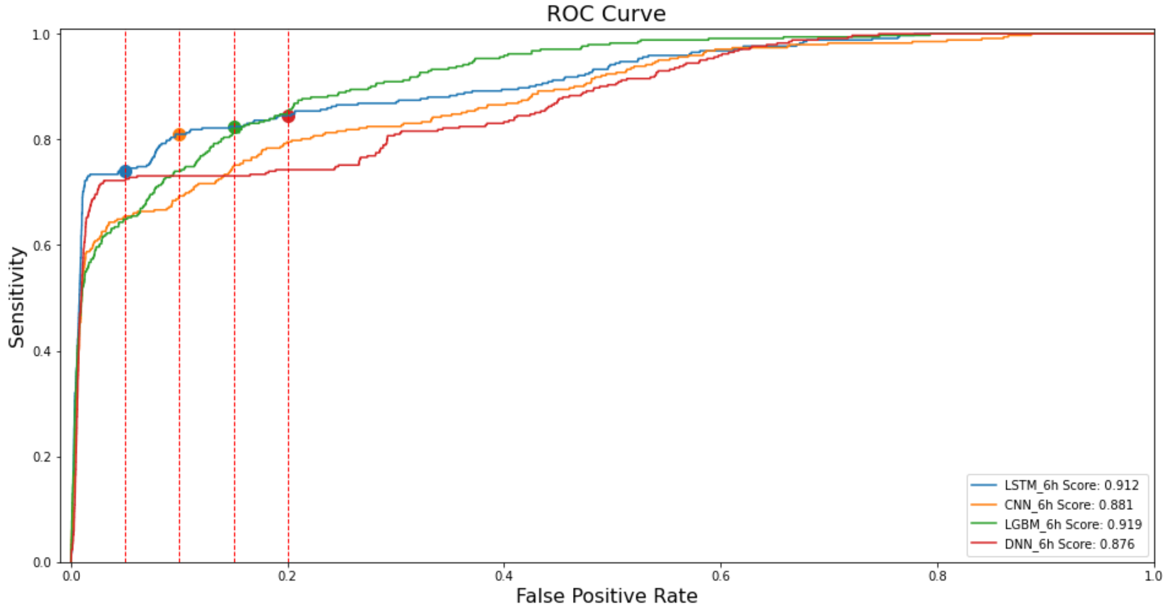
Predictive Model in 6 hours

Hospital A (Validation)



	0.05	0.10	0.15
LSTM	0.679	0.868	0.925
CNN	0.698	0.792	0.830
LGBM	0.660	0.830	0.887
DNN	0.547	0.698	0.830

Hospital B (Test)

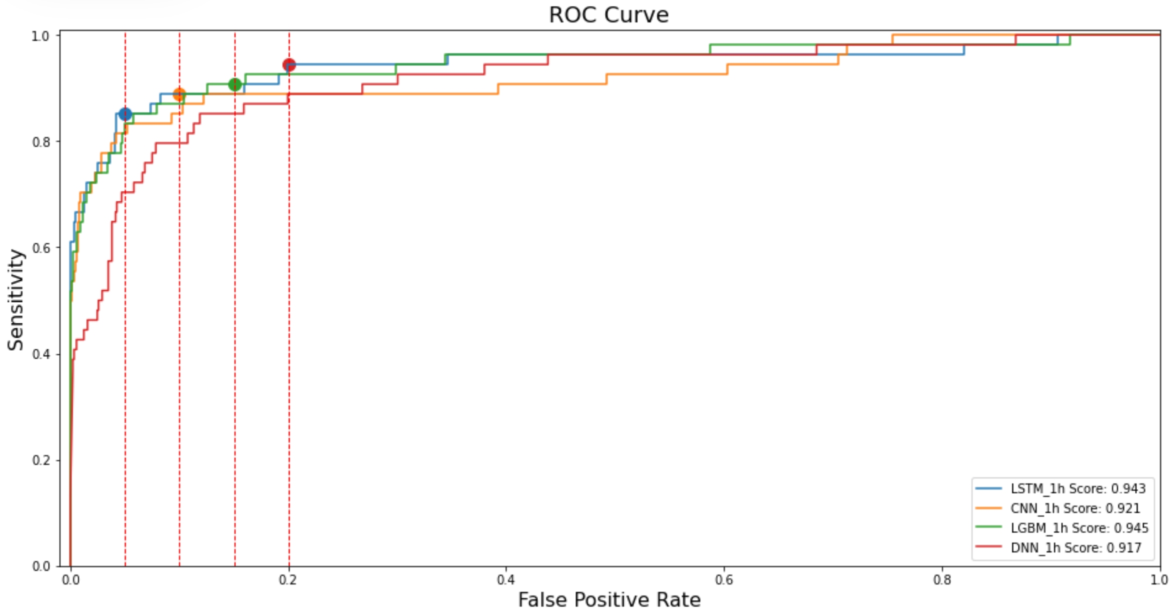


	0.05	0.10	0.15
LSTM	0.740	0.807	0.825
CNN	0.652	0.690	0.749
LGBM	0.646	0.740	0.810
DNN	0.722	0.731	0.731

Results

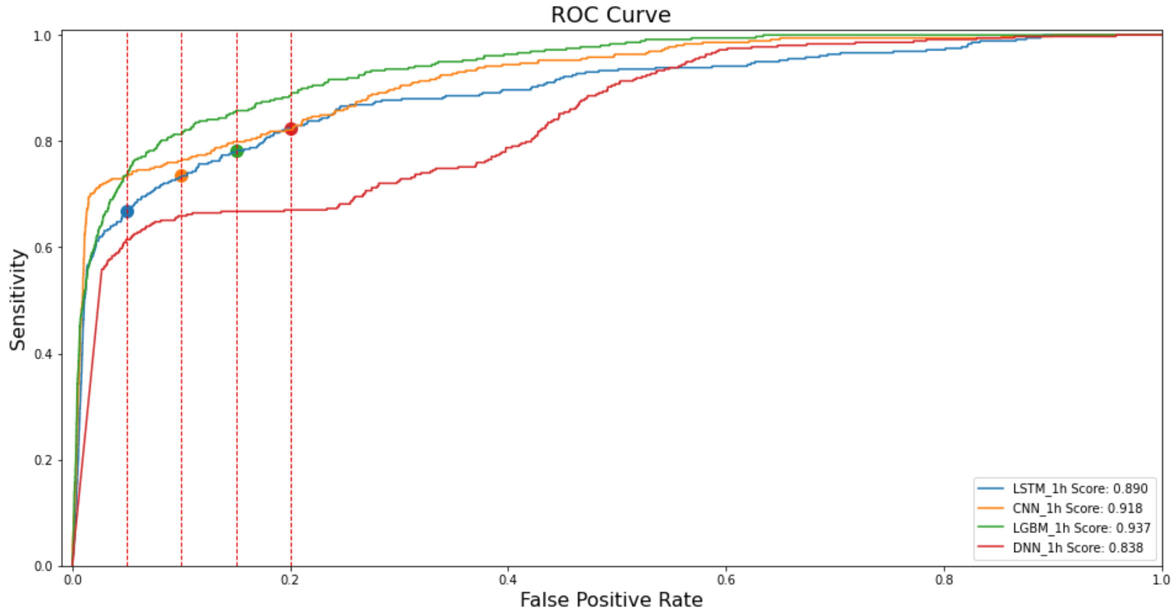
Predictive Model in 1 hour

Hospital A (Validation)



	0.05	0.10	0.15
LSTM	0.815	0.870	0.889
CNN	0.815	0.852	0.870
LGBM	0.815	0.870	0.907
DNN	0.685	0.796	0.852

Hospital B (Test)

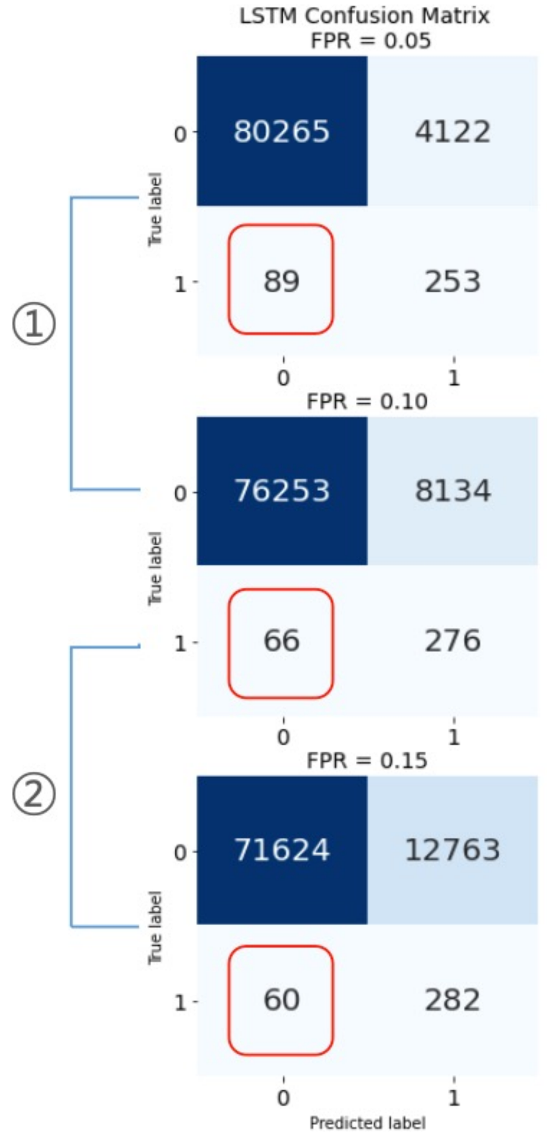


	0.05	0.10	0.15
LSTM	0.665	0.732	0.779
CNN	0.735	0.763	0.799
LGBM	0.735	0.813	0.855
DNN	0.615	0.659	0.668

Results

Confusion Matrix of Predictive Model in 6 hours

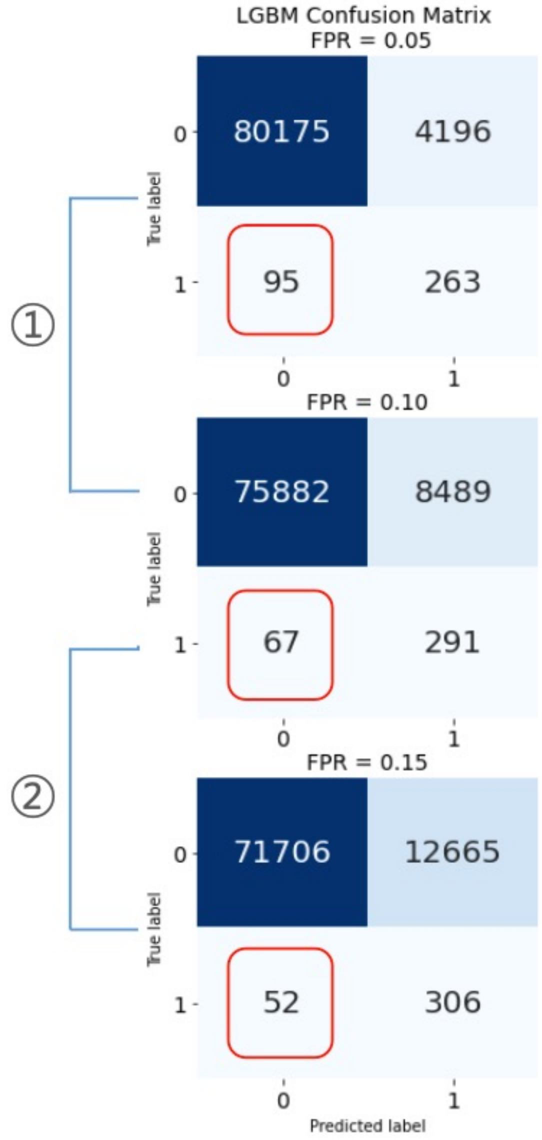
- LSTM performs best among predictive models in 6 hours
- Type 2 error decrease : Type 1 error increase
    - ① 23 : 4012 ≒ 1 : 174
    - ② 6 : 4629 ≒ 1 : 771



Results

Confusion Matrix of Predictive Model in 1 hour

- LGBM performs best among predictive models in 1 hour
- Type 2 error decrease : Type 1 error increase
  - ① 28 : 4293 ≒ 1 : 153
  - ② 15 : 4176 ≒ 1 : 278



Results

Comparison with Existing System

	LSTM	DEWS	MEWS
Recall			
FPR = 0.05	0.740	0.493	-
FPR = 0.10	0.807	0.607	0.373
FPR = 0.15	0.825	0.630	0.493
AUROC	0.912	0.837	0.765
AUPRC	0.158	0.239	0.028

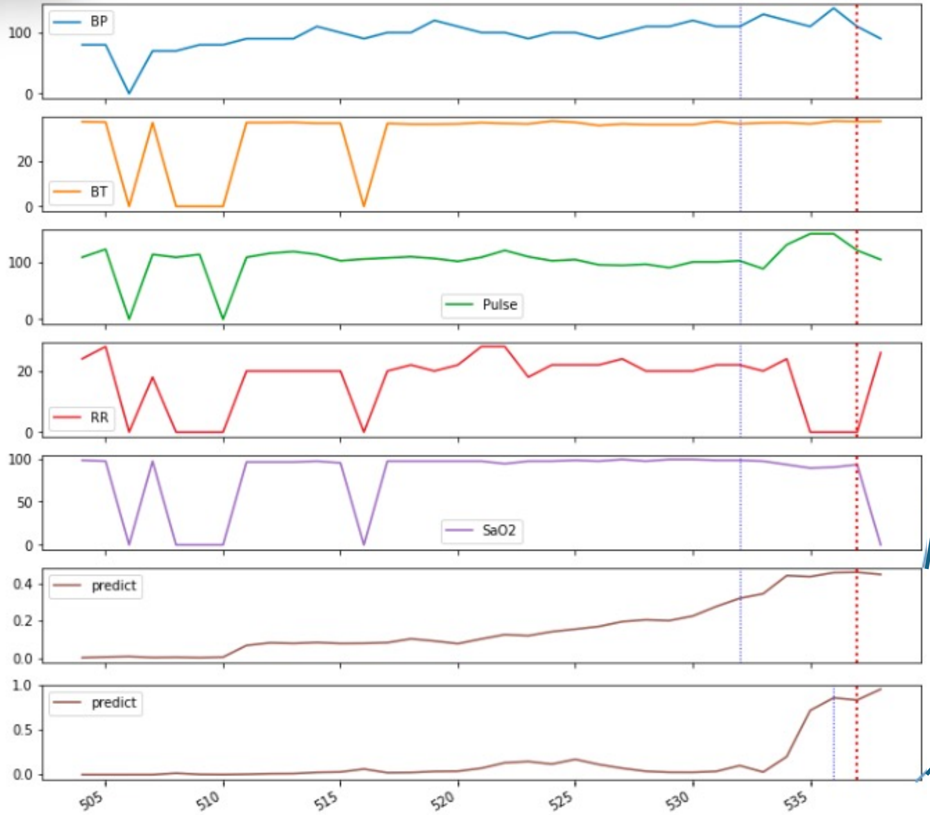
Predictive Model in 6 hours

	LGBM	DEWS	MEWS
Recall			
FPR = 0.05	0.735	0.493	-
FPR = 0.10	0.813	0.607	0.373
FPR = 0.15	0.855	0.630	0.493
AUROC	0.937	0.837	0.765
AUPRC	0.144	0.239	0.028

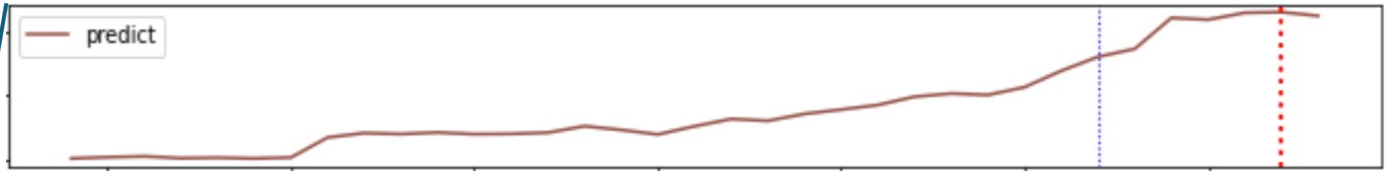
Predictive Model in 1 hour

Results

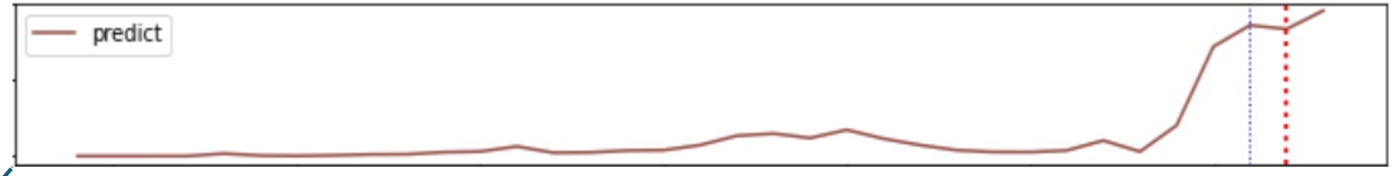
Real-Time Risk Prediction Chart



Patient in Hospital B
(35 years old, female)



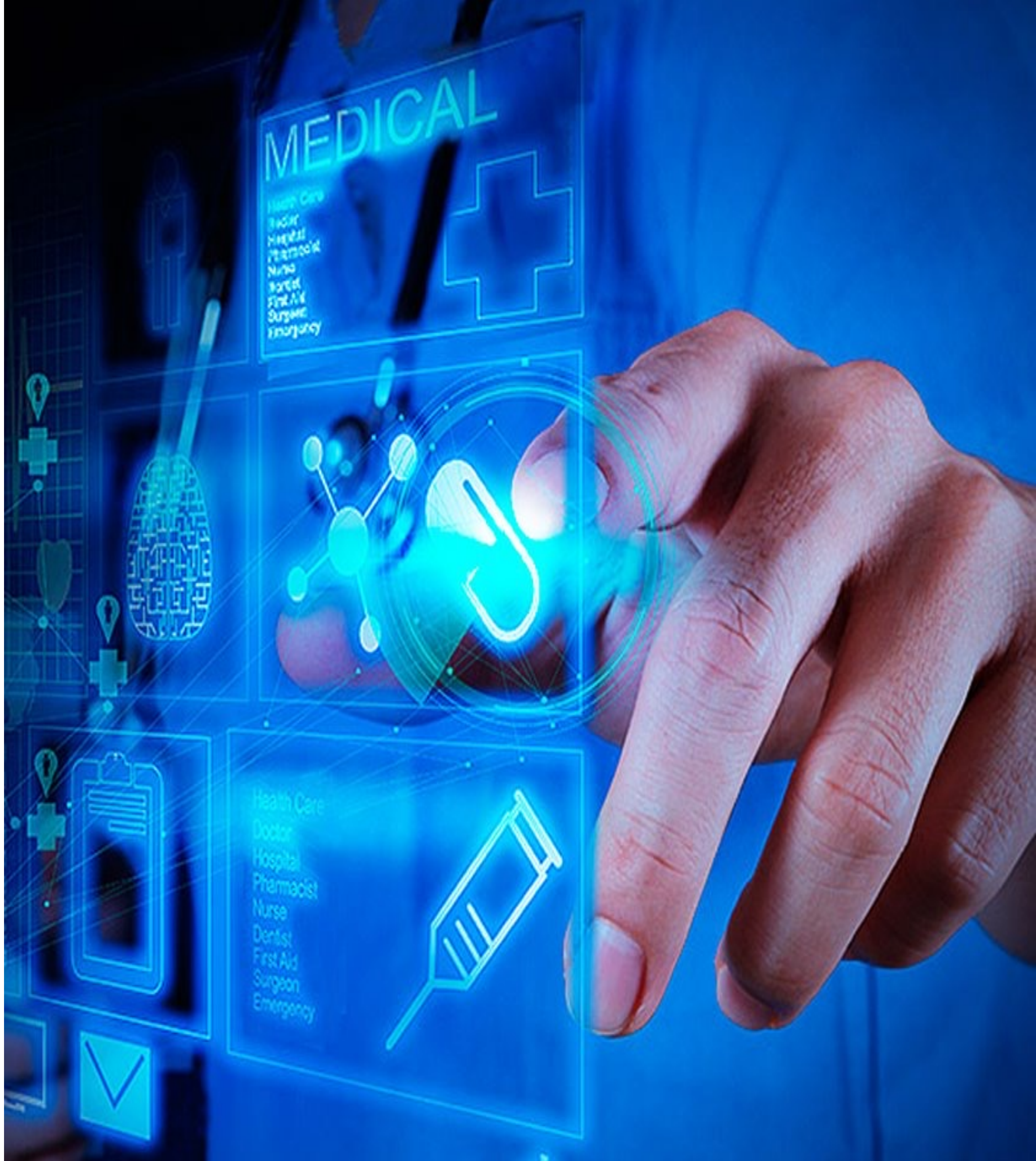
Predictive Model in 6 hours



Predictive Model in 1 hour

AI ZEN

Implication



Implication

- Limitation
  - Lack of data on risk group
  - Approach by grouping variables by measurement period
  - Explanatory power

- Implication
  - Rather than viewing data as simple numbers and processing it mechanically, we learned the need for a statistical approach based on a sufficient understanding of the data.

Thank you

Appendix

Reference

- An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest

<Kwon JM, Lee Y, Lee Y, Lee S, Park J. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest. J Am Heart Assoc 2018;7>

- A deep learning model for real-time mortality prediction in critically ill children

<Kim SY, Kim S, Cho J, et al. A deep learning model for real-time mortalityprediction in critically ill children. Crit Care. 2019;23:279>

- Recurrent Neural Networks for Multivariate Time Series with Missing Values

<Che, Z. et al. Recurrent neural networks for multivariate time series with missing values. Rep. 8, 1–12 (2018)>

- Learning representations for the early detection of sepsis with deep neural networks

<Kam, H. J. & Kim, H. Y. Learning representations for the early detection of sepsis with deep neural networks. Computers in biology and medicine 89, 248–255 (2017)>

Appendix

Distribution of Risk Group and Control Group by Train and Test Dataset

	Train, Hospital A	
	Risk Group	**Control Group**
Time	2017.01~2019.02	2017.03~2019.03
Patient Number (Data Size)	345(22884)	4471(138826)
Measuring Period		
2 days	3(91)	-
3 days	36(1527)	-
4 days	306(21266)	-
5 days	-	4471(138826)
Sex Ratio (M:F)	62.6% : 37.4%	65.3% : 34.7%
Age	63.7± 15.9	71.4±10.9
Min	2.2	40.8
Max	102.7	87.2

	Test, Hospital B	
	Risk Group	**Control Group**
Time	2019.03 ~ 2019.05	2019.01~2019.05
Patient Number (Data Size)	61(3994)	2527(99571)
Measuring Period		
2 days	0(0)	-
3 days	7(341)	-
4 days	54(3653)	-
5 days	-	2527(99571)
Sex Ratio (M:F)	50.8% : 49.2%	56.8% : 43.2%
Age	70.2± 14.9	67.5±16.1
Min	23	16.2
Max	95	98.7

Appendix

Data Details

Patient Data

Patient_ID	Hospitalization date	Hospitalization days	Measurement time
Sex	4 days after hospitalization	Original data number	Body temperature
Birthdate	Prescription date	Measurement date	Pulse
Age	4 days before prescription	Systolic blood pressure	Respiration
Deathdate	Discharge date	Diastolic blood pressure	Sa02

Event Data

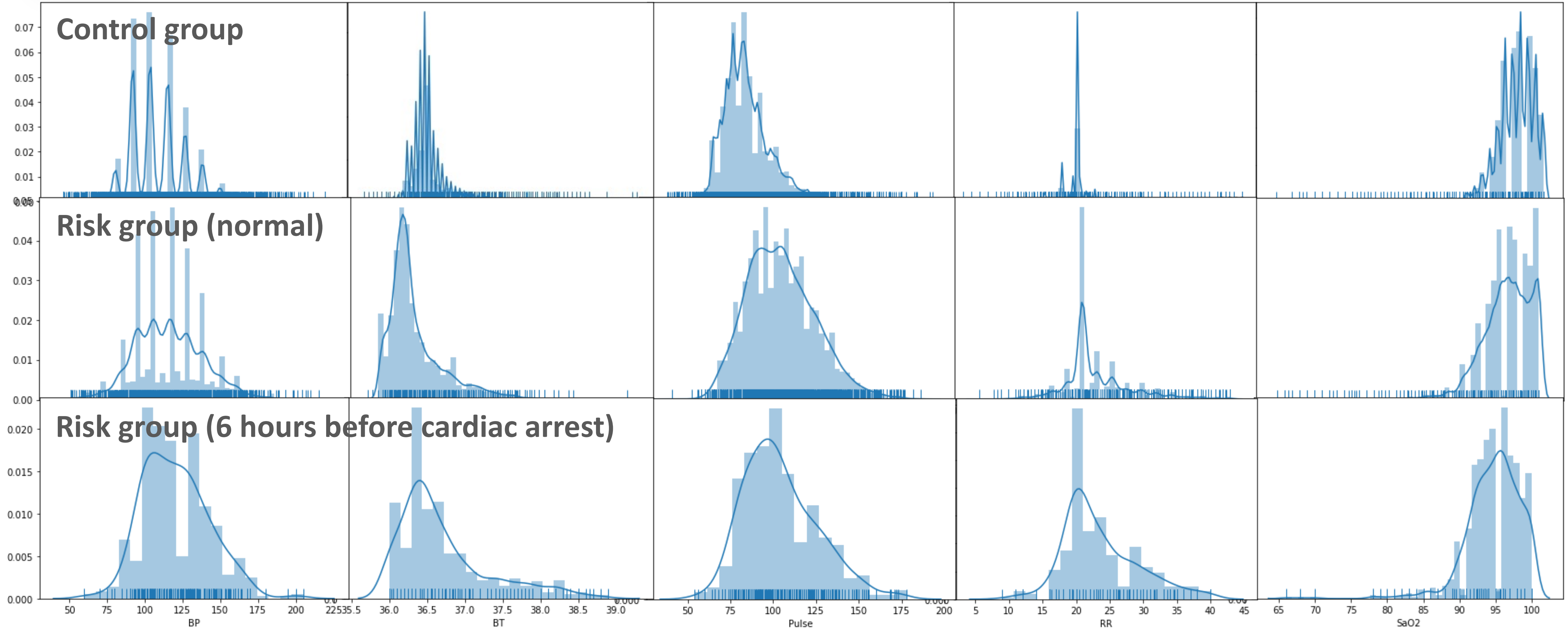
Event_ID	Deathdate (R)	Encounter date	Detection_date
Sex	Deathdate	Event_date	Detection_time
Birthdate	Age	Event_time	

Blood Data

WBC count	platelet	Hgb	BUN
Creatinin	Glucose	Sodium	Potassium
Chloride	Total protein	Total bilirubin	Albumin
CRP	Total calcium	Lactate	Alkaline phosphatase
AST	ALT		

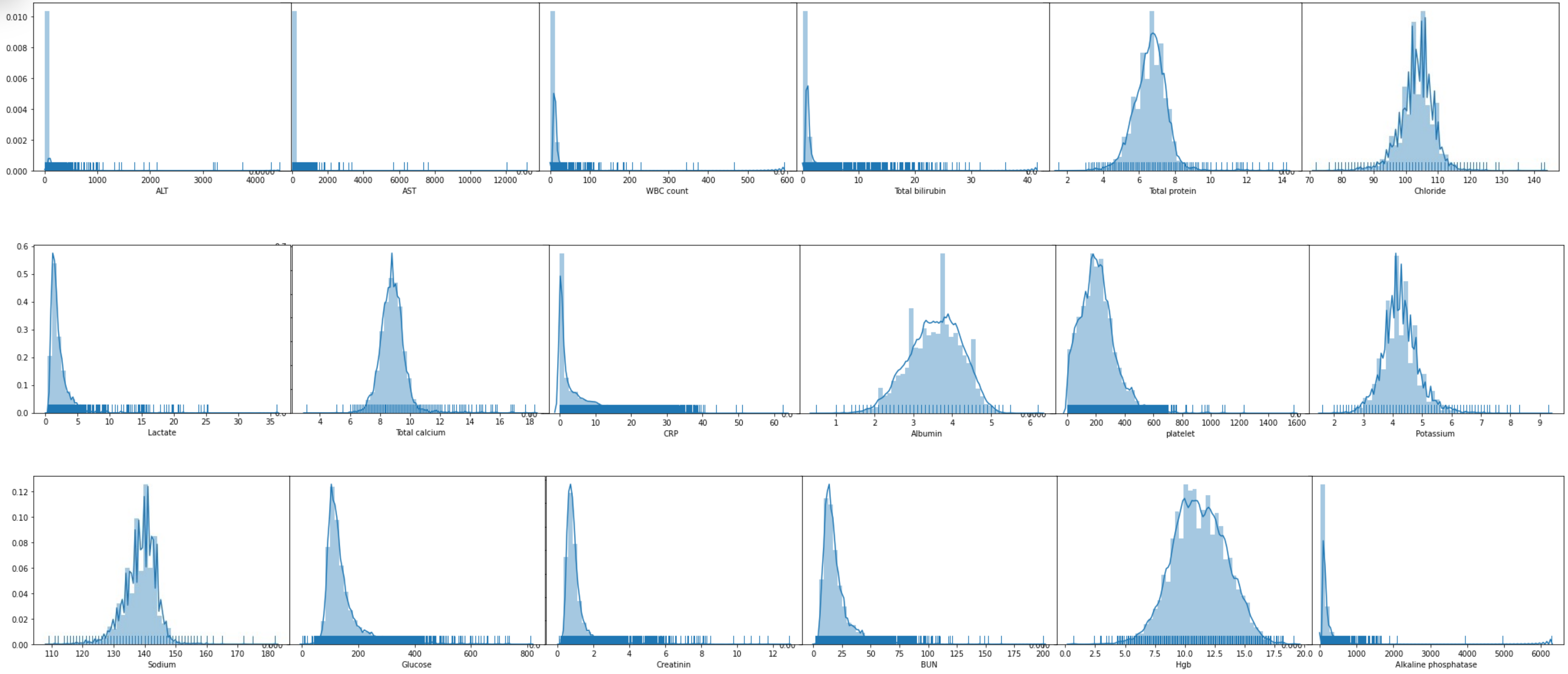
Appendix

Vital Sign Distribution



Appendix

Blood Data Distribution



Appendix

Proportion of Missing Values

	Blood Pressure	Body Temperature	Pulse	Respiration	SaO2
Risk Group	54.7%	52.0%	54.5%	55.5%	23.1%
Control Group	21.1%	17.9%	22.3%	22.8%	53.4%

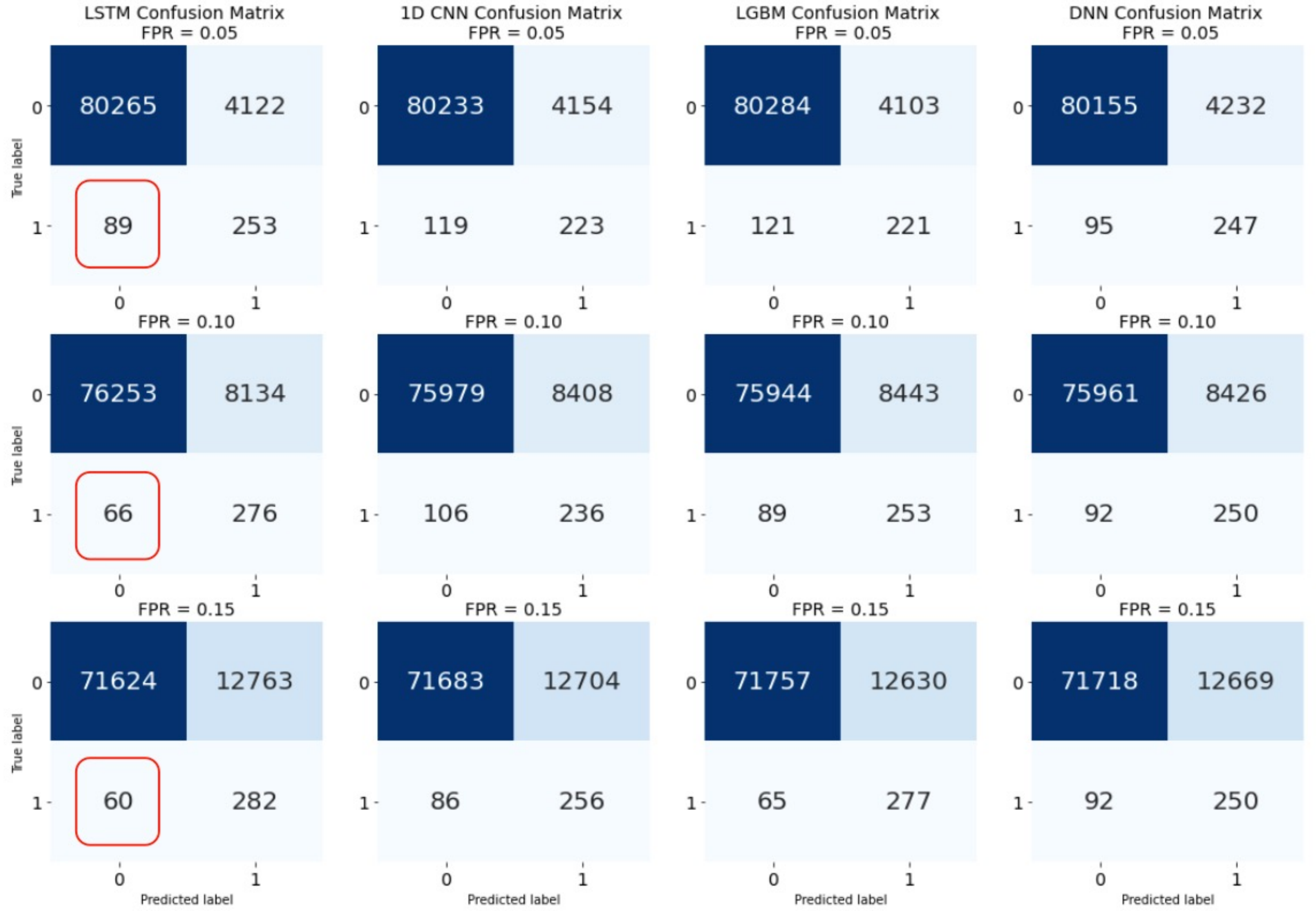
Missing Value Ratio by Variable

	Blood Pressure	Body Temperature	Pulse	Respiration	SaO2
Risk Group	72.3%	70.9%	72.7%	73.1%	45.0%
Control Group	78.2%	77.1%	78.5%	78.6%	86.7%

Missing Value Ratio with 1-hour equalization

Appendix

Confusion Matrix of All Models (6 hour prediction)



Appendix

Confusion Matrix of All Models (1 hour prediction)