Application of Machine Learning and IoT for Healthcare

Nano Information Technology Academy,
Dongguk University-Seoul

Ganjar Alfian
ganjar@dongguk.edu

dongguk UNIVERSITY

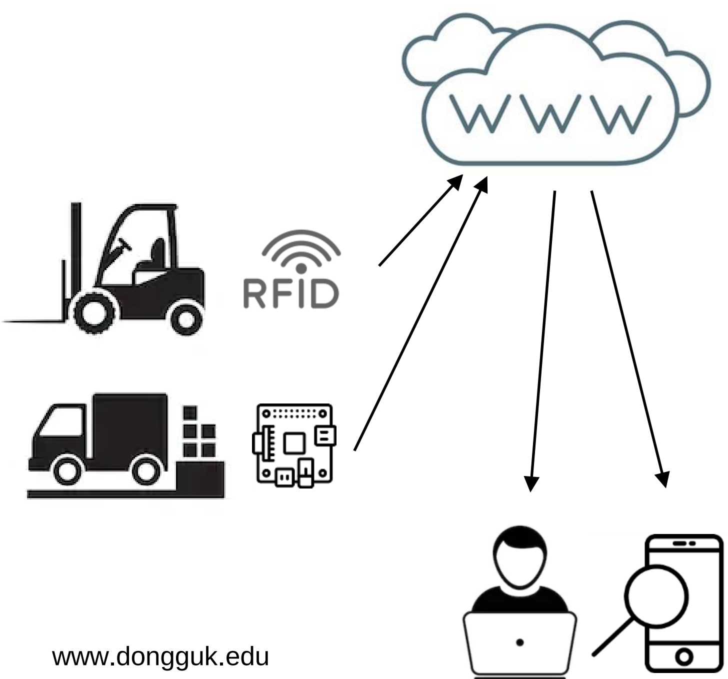
Contents

1. Introduction of IoT and Machine Learning
2. Application for Healthcare Monitoring System
3. Research collaboration

1. Introduction of IoT and Machine Learning

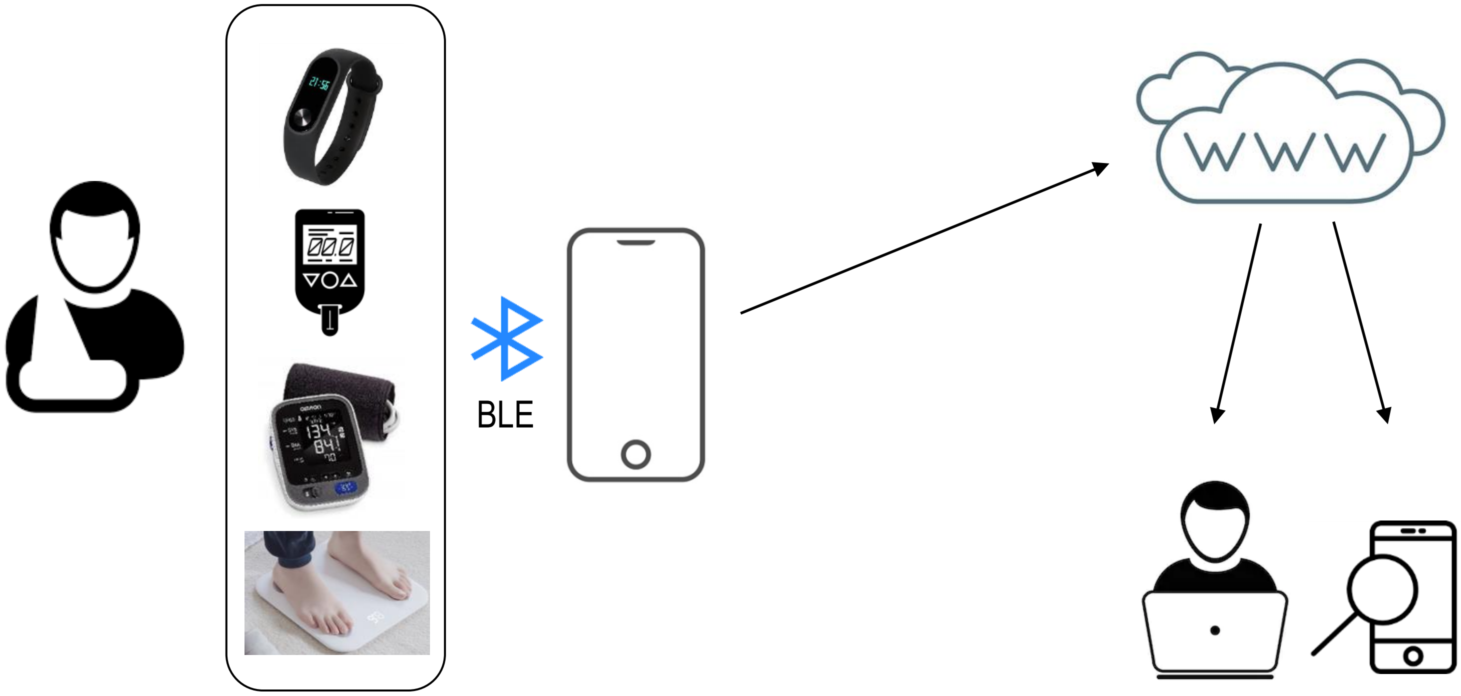
What is IoT

- Internet of Things (IoT) comprises things that have unique identities and are connected to the Internet.
- The focus on IoT is in the configuration, control and networking via the Internet of devices or “Things” that are traditionally not associated with the internet.



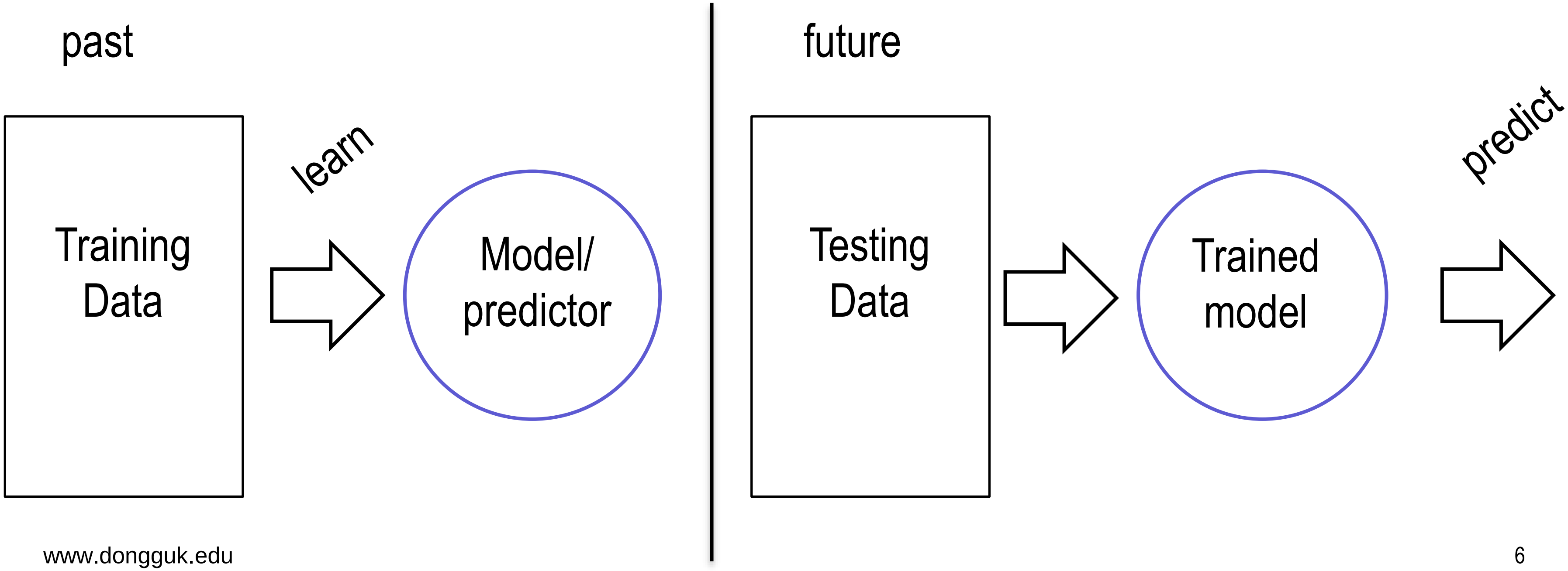
What is IoT

- The Scope of IoT is not limited to just connecting things (device, appliances, machines) to the Internet
- IoT allows these things to communicate and exchange data (control & information)
- Processing on these data will provide us various applications towards a common user.



What is machine learning?

- A branch of artificial intelligence (AI), concerned with the design and development of algorithms that allow computers to evolve behaviours based on empirical data.
- As intelligence requires knowledge, it is necessary for the computers to acquire knowledge.



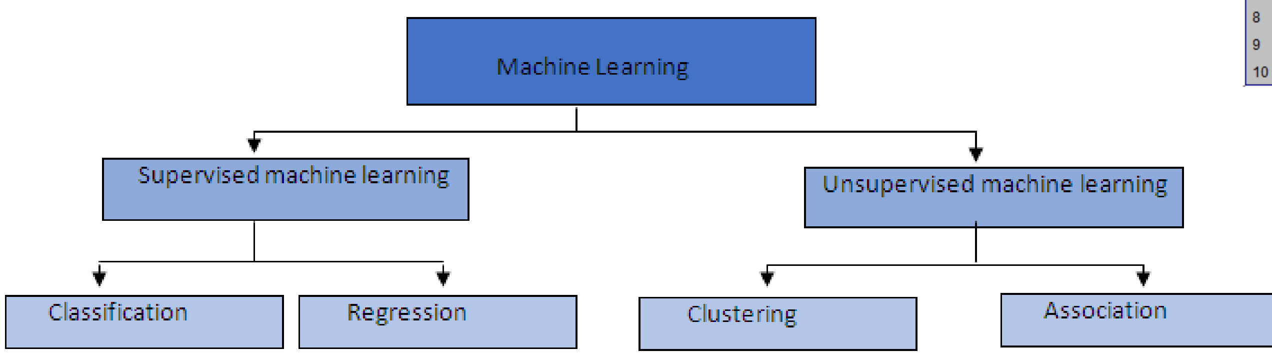
Types of Learning

- Supervised (inductive) learning
  - Training data includes desired outputs
  - Classification, regression/ prediction
- Unsupervised learning
  - Training data does not include desired outputs
  - Clustering
- Semi-supervised learning
  - Training data includes a few desired outputs
- Reinforcement learning
  - Rewards from sequence of actions

Tid	Attrib1	Attrib2	Attrib3	Class
1	Yes	Large	125K	No
2	No	Medium	100K	No
3	No	Small	70K	No
4	Yes	Medium	120K	No
5	No	Large	95K	Yes
6	No	Medium	60K	No
7	Yes	Large	220K	No
8	No	Small	85K	Yes
9	No	Medium	75K	No
10	No	Small	90K	Yes

Tid	Attrib1	Attrib2	Attrib3
1	Yes	Large	125K
2	No	Medium	100K
3	No	Small	70K
4	Yes	Medium	120K
5	No	Large	95K
6	No	Medium	60K
7	Yes	Large	220K
8	No	Small	85K
9	No	Medium	75K
10	No	Small	90K

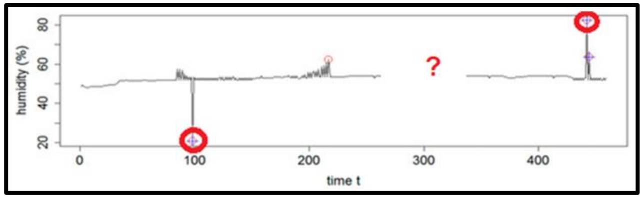
Tid	Attrib1	Attrib2	Attrib3	Class
1	Yes	Large	125K	No
2	No	Medium	100K	
3	No	Small	70K	
4	Yes	Medium	120K	
5	No	Large	95K	
6	No	Medium	60K	No
7	Yes	Large	220K	No
8	No	Small	85K	Yes
9	No	Medium	75K	
10	No	Small	90K	Yes



Stage in Machine Learning

- Data pre-processing
  - Data cleaning, filling missing value, remove outlier
- Train models
  - Select the algorithm
  - Feature selection and extraction
- Evaluate models
  - Assess performance
  - Model comparison
- Deploy model
  - Apply model to new data
  - Real-time demonstration

Sex	Age	BMI	DM type	DM duration	FBS	Sys BP	Dias BP	Retinopathy
Male	65	25	II	20	129	130	80	Yes
Male	42	27	II	300	210	140	90	No
Female	31	21	I	11	164	145	80	Yes
Male	70	32	II	29	208	160	100	Yes
Female	54	34	II	6	183	155	95	No
	46	29	II	7	198	160	100	No
Female	16	24	I	-1	250	135	80	No
Male	67	30	II	12	243	165	90	Yes
Female	51	28	II	7	163	130	85	No
Girl	70	36	II	20	250	150	90	Yes
Female	63	35	II	14	203	160	110	No
Male	44	39	II	3	149	140	90	No
Boy	51	24	II	9	160	155	80	No
Male	27	19	I	5	170	140	90	No



Dataset preparation and pre-processing → Train models → Evaluate models → Model deployment

1. Introduction of IoT and Machine Learning

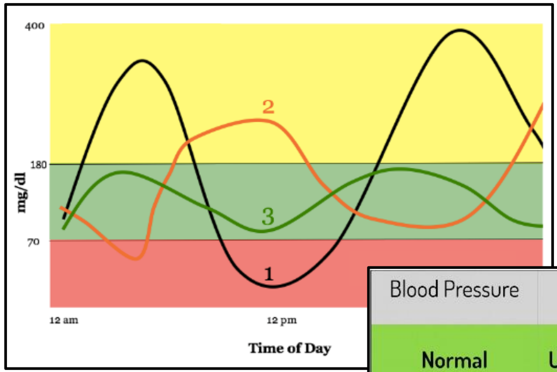
2. Application for Healthcare Monitoring System

Introduction: Diabetes and Hypertension

- Diabetes is a long-term metabolic disorder in which the blood glucose (BG) level varies and is caused by
  - insufficient insulin production in the body (T1D)
  - the body's inability to utilize its produced insulin (T2D)
  - gestational (GDM), during pregnancy
- Ineffectively supervised diabetes
  - cardiovascular diseases (CVD), including hypertension and stroke
- Hyperglycemia (BG > 180 mg/dL)
  - long-term complications, e.g., retinopathy, nephropathy, and CVD,
- Hypoglycaemia (BG < 70 mg/dL)
  - short-term adverse conditions that can cause coma or even death (ADA, 2018)
- Regular monitoring of the BG level
  - reducing and avoiding the complications of diabetes
- Hypertension is condition where
  - systolic blood pressure ≥140 mmHg
  - diastolic blood pressure ≥90 mmHg.

Criteria for Diagnosis

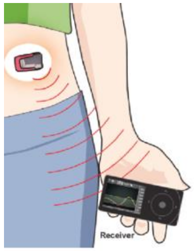
2-h plasma glucose	≥ 200 mg/dL
random / casual	≥ 200 mg/dL
fasting plasma glucose	≥ 126 mg/dL.
A1C	≥ 6.5%



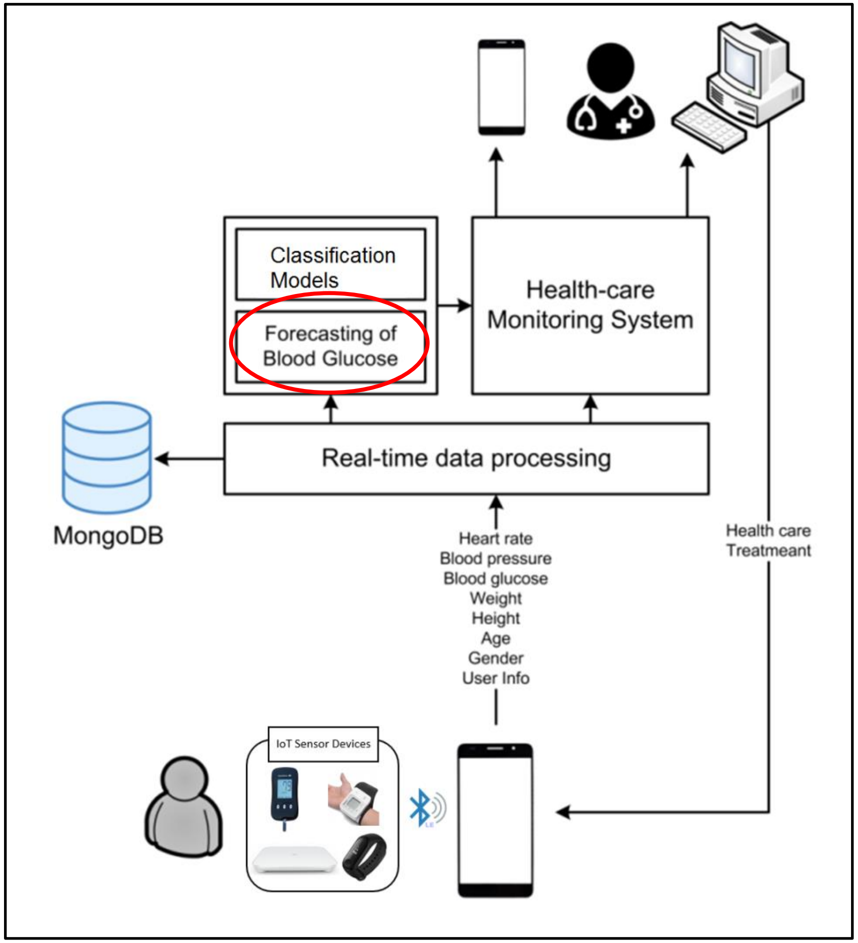
Blood Pressure	Systolic (Upper#)	Diastolic (Lower#)
Normal	Under 120	Under 80
Prehypertension	120-139	80-89
Hypertension (Stage 1)	140-159	90-99
Hypertension (Stage 2)	Above 160	Above 100
Hypertensive Crisis	Above 180	Above 110

Introduction: Monitoring Device

- Glucose monitoring
  - To optimize patient treatment strategies (effect of medications, exercise, and/or diet).
- Self-monitoring of blood glucose (SMBG) portable device
  - Sampling blood from a finger via pricking and use glucose meter.
  - Example : CareSens and Accu-Chek
- Continuous Glucose Monitoring (CGM) sensors
  - Provide real-time measurements (every 1–5 min) for several consecutive days (Example Dexcom).
  - The CGM devices utilize a wire-based sensor, usually placed in the subcutaneous tissue.
- Blood Pressure Monitor (Wrist, Upper Arm)
  - Wrist monitors are often smaller, lighter, and more portable.



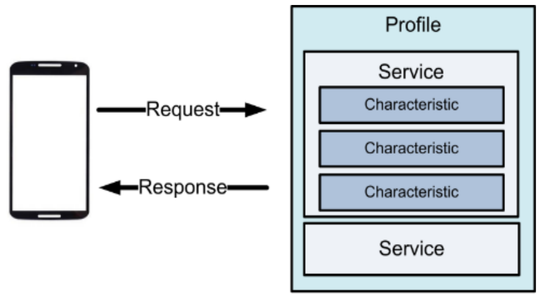
Healthcare Monitoring System



- BLE-based Sensor are utilized
  - Glucometer, BP Monitor, Smartwatch, weight scale, CGM.
- The generated sensor data:
  - HR, BP, Weight, BG, Personal input data (Gender, Height, Age).
- BG forecasting and classification models to predict future condition.
- Medical doctor can monitor patient condition.
- Healthcare recommendation is sent to the patient.

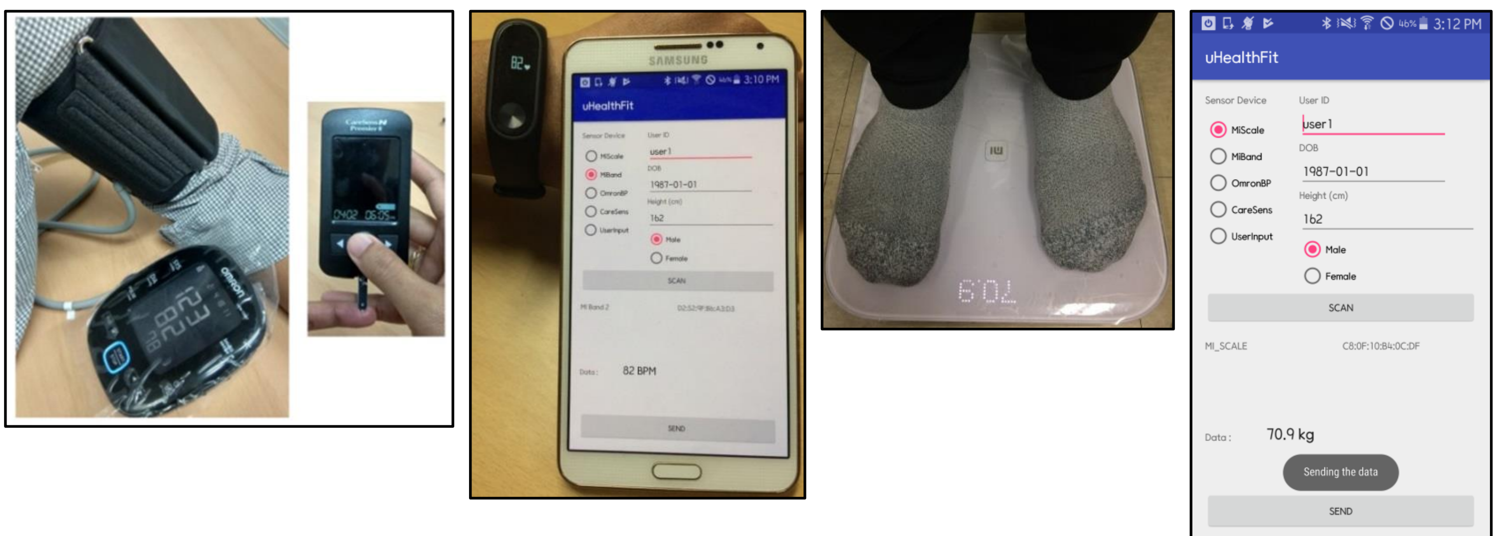
Source: Alfian et al. (2018). A personalized healthcare monitoring system for diabetic patients by utilizing BLE-based sensors and real-time data processing. Sensors. http://doi.org/10.3390/s18072183.

IoT sensor device



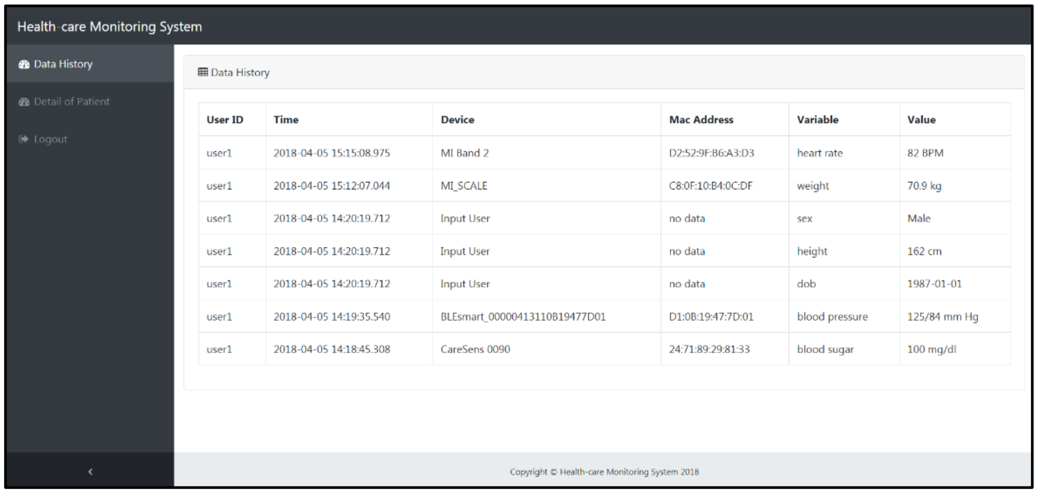
GATT client GATT Server

- Bluetooth Low Energy (BLE)
  - Short range
  - Low power consumption rate
- Generic Attributes (GATT)
  - To communicate between BLE peripheral and gateway (central device).
- We developed a prototype of an android app
  - To retrieve the vital signs data from sensor to gateway.
  - Retransmitting it to the secure remote server wirelessly

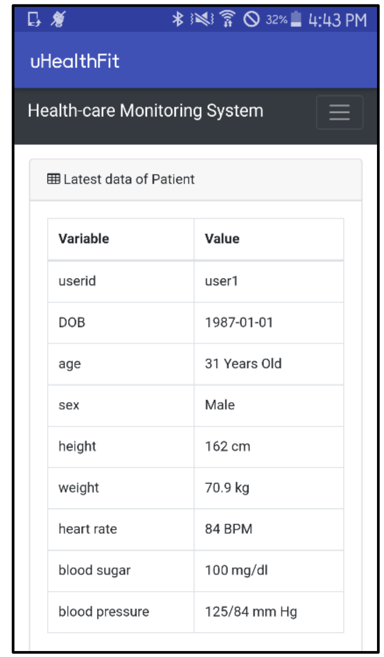


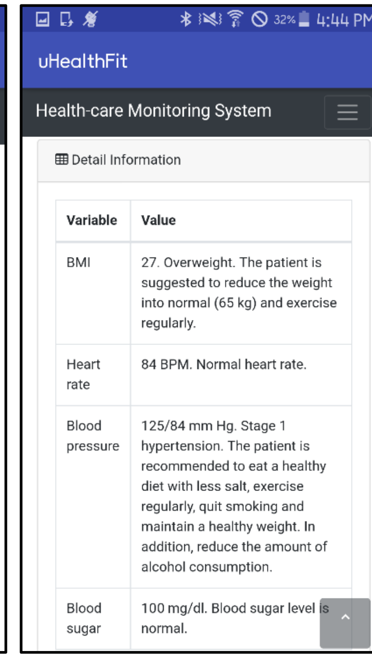
Prototype

Personalized Healthcare Monitoring System



Web-based monitoring





Personalized app

- Web-based
  - Storing vital signs and presenting in web page
- Android app
  - Presenting vital signs and recommendation

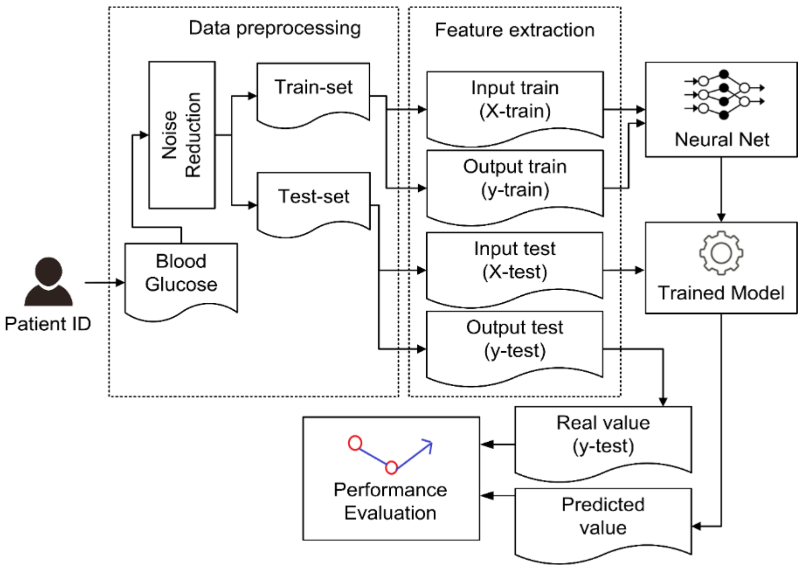
Forecasting of Blood Glucose (BG)

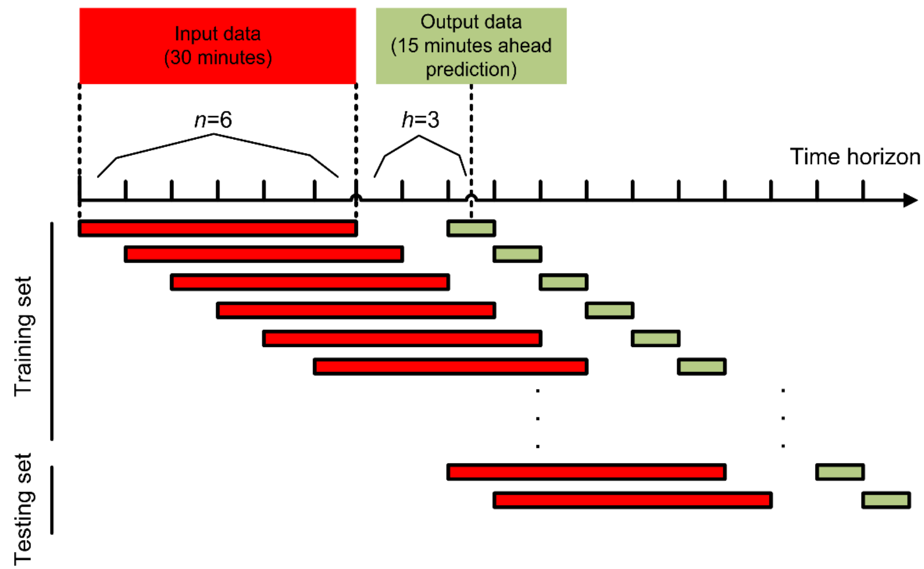
Study	Method	Subject
Perez-Gandia et al. (2010)	ANN with last 20 mins of BG.	15 T1D
Wang et al. (2013)	Adaptive-weighted-average framework based on AR, ELM and SVR.	10 T1D
Hamdi et al. (2018)	SVR based on DE algorithm.	12 T1D
Ben Ali et al. (2018)	ANN with optimized input for each patient.	12 T1D
Martinsson et al. (2020)	LSTM with 60 mins of glucose level history.	6 T1D
Proposed model	ANN last 30 mins with additional time-domain features.	12 T1D

- We focused on prediction model with single variable (CGM) only.
- Our proposed Neural Network to predict BG for the next 15, 30, 45 and 60 minutes.
- Proposed model could :
  - Predict BG, therefore critical conditions can be avoided.

Source : Alfian et al. (2020). Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features. Biocybernetics and Biomedical Engineering. https://doi.org/10.1016/j.bbe.2020.10.004

BG Prediction Model



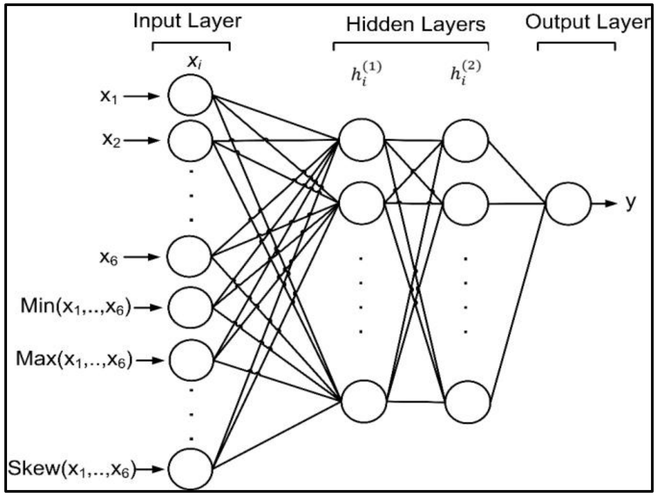


Feature extraction for multi-step forecasting using direct method

- **Dataset**
  - 12 Children with T1D
  - Open dataset = DirecNet (https://public.jaeb.org/direcnet/stdy/167)
  - Recorded by CGM (Guardian RT), 5 mins interval, approximately 7 days.

Proposed model

Attribute	Description
min	Minimum value of BG
max	Maximum value of BG
mean	Average value of BG
std	Standard deviation value of BG
diff	Difference between highest and lowest
median	Middle value of BG in a window.
kurtosis	Describing tails of BG distribution
skew	Distribution asymmetry



Given the list of BG data G, $n$previous values (or the window size) and theforecasting horizon, h, the input X can be derived by creating an$[(N-n-h+1)\times n]$

$$X = \begin{bmatrix} g_1 & \cdots & g_{n-1} & g_n & min(g_1,\ldots,g_n) & \cdots & skew(g_1,\ldots,g_n) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{N-n-h} & \cdots & g_{N-h-2} & g_{N-h-1} & \cdots & \cdots & skew(g_{N-n-h},\ldots,g_{N-h-1}) \\ g_{N-n-h+1} & \cdots & g_{N-h-1} & g_{N-h} & \cdots & \cdots & skew(g_{N-n-h+1},\ldots,g_{N-h}) \end{bmatrix}$$

and the $[(N-n-h+1)\times 1]$ $$Y = \begin{bmatrix} g_{n+h} \\ \vdots \\ g_{N-1} \\ g_N \end{bmatrix}$$

- Hold-out with 80:20
- Optimized networks :
  - 2 hidden layers, 100 nodes each
  - ReLU, Adam optimizer

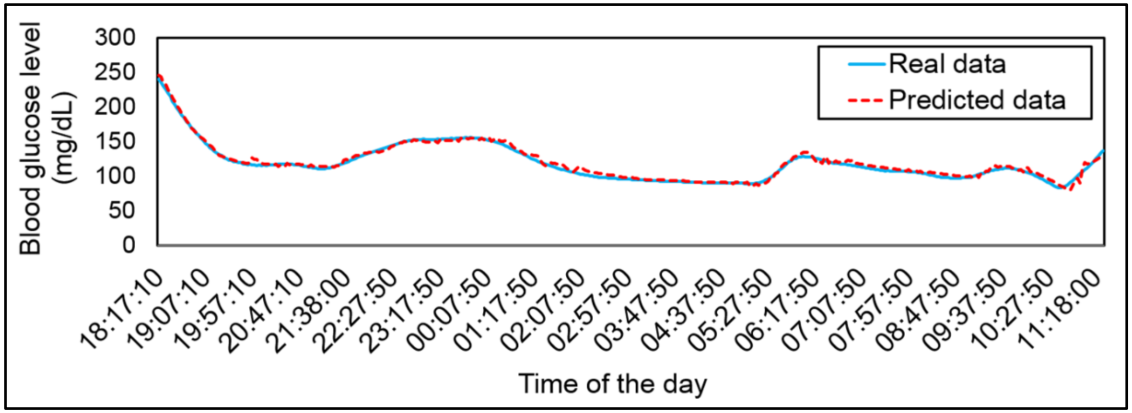
Result and Discussion

Model	Feature type		RMSE (mg/dL)			
	Last 30 mins	Time-domain	15 mins	30 mins	45 mins	60 mins
SVR	√		13.13	17.86	22.68	27.08
	√	√	15.47	19.73	23.74	27.68
KNN	√		6.03	11.81	18.86	26.32
	√	√	7.89	12.65	18.66	25.25
DT	√		7.87	15.04	23.35	32.86
	√	√	7.82	15.33	22.33	32.26
RF	√		5.06	10.68	17.72	24.65
	√	√	4.86	9.85	16.38	23.21
AdaBoost	√		7.39	14.56	22.44	28.78
	√	√	7.44	13.56	20.24	26.76
XGBoost	√		5.42	10.79	17.20	23.49
	√	√	5.02	9.92	15.89	21.17
MLP	√		3.53	7.72	12.60	18.72
	√	√	2.82	6.31	10.65	15.33

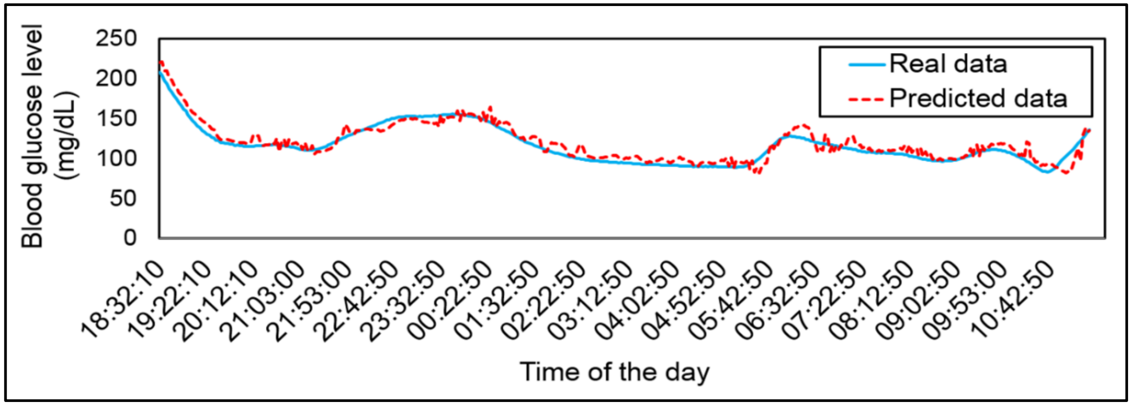
- The proposed model outperformed other models
  - RMSE = 6.31 mg/dL (PH 30) and 15.33 mg/dL (PH 60)
  - Time domain features improved most of the ML models

Prediction output : Patient ID 21

PH 30 min



PH 60 min



- Increasing PH
  - increase RMSE, MAPE, gMSE
  - decrease *R2*

Comparison with some other previous studies

Study	Method	PH (min)	RMSE (mg/dL)
Perez-Gandia et al. (2010)	ANN with last 20 minutes of BG.	15	9.70
		30	17.50
		45	27.10
Wang et al. (2013)	Adaptive-weighted-average framework based on AR, ELM and SVR	15	±10
		30	±19
		45	±28
Hamdi et al. (2018)	SVR based on DE algorithm.	15	9.44
		30	10.78
		45	11.82
		60	12.95
Ben Ali et al. (2018)	ANN with optimized input for each patient.	15	6.43
		30	7.45
		45	8.13
		60	9.03
Martinsson et al. (2020)	LSTM with 60 min of glucose level history.	30	18.87
		60	31.40
Proposed model	ANN with additional time-domain features.	15	2.82
		30	6.31
		45	10.65
		60	15.33

- Better for short-term forecasting
- Future works
  - Need to consider different strategy, such as recursive
  - Need optimized input for each patient

1. Introduction of IoT and Machine Learning
2. Application for Food Traceability System
3. Research collaboration

Research Collaboration

- Research Area
  - Artificial Intelligence, Machine Learning, Deep Learning, Simulation, Health Informatics, Food Supply Chain, Intelligent Store, IoT, RFID
- Partners
  - Joint Project and Publication : VSB-TUO (Czech)
  - Joint Publication : UGM, UII, Telkom University, UIN Sunan Kalijaga, Unjani Yogyakarta, UAD (Indonesia), GIKI Pakistan, UBD Brunei
- Hardware
  - UHF RFID Readers, Handheld Readers, Antennas, Passive Tags
  - IoT devices (Blood glucose, blood pressure, weight scale, smartwatch, raspberry pi, and other sensors)
  - Server computers (linux, window-based)
- Software
  - Python, Java, PHP, JSP, R, C++, MySQL, Oracle, MongoDB

End of Presentation

- Thank You for the attention

- Q & A

- For further discussion/ question, please send to email (ganjar@dongguk.edu)