

세틀뱅크

Machine Learning



# 목차

01

머신러닝의 특성

02

MTMS

03

개선할 점

04

논의

05

마무리

# 01

## 머신러닝의 특성

## 인공지능, 머신러닝 딥러닝

- 기계가 학습을 할 수 있도록 하는 연구 분야
- 인공지능 연구의 한 분야로서 최근들어 딥러닝을 통해서 빠르게 발전

### 인공지능

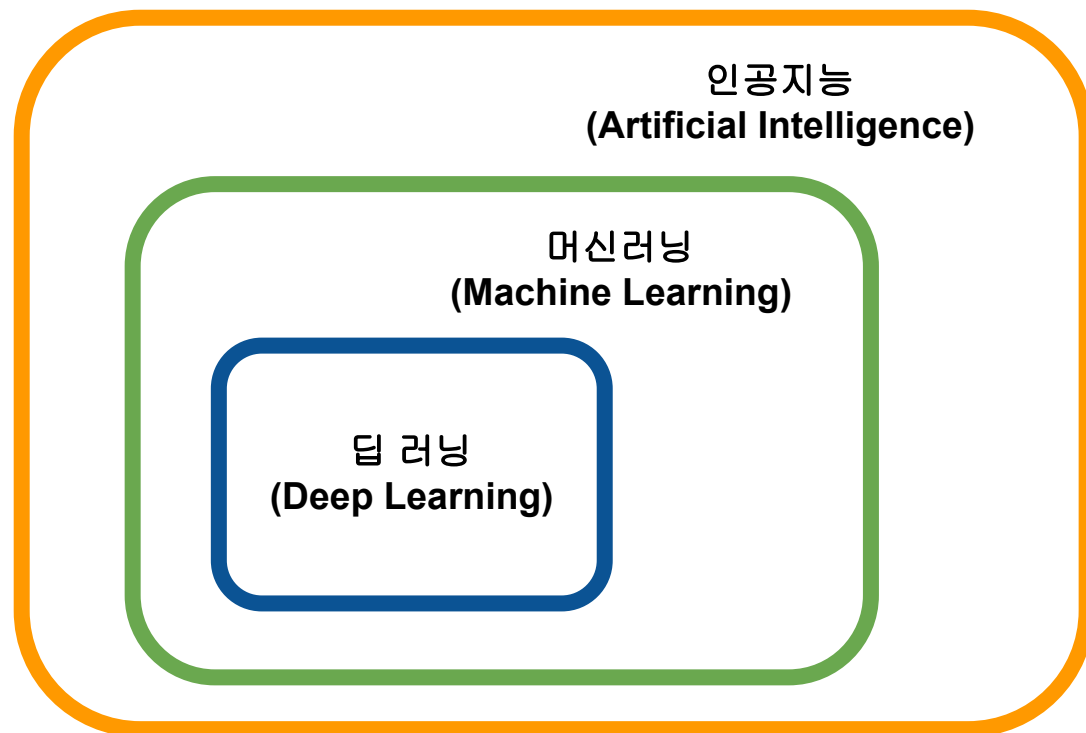
- 인간이 지닌 지적 능력의 일부 또는 전체, 혹은 그렇게 생각되는 능력을 인공적으로 구현한 것을 말한다

### 머신러닝

- 머신러닝은 경험적 데이터를 기반으로 학습을 하고 예측을 수행하고 스스로의 성능을 향상시키는 시스템과 이를 위한 알고리즘을 연구하고 구축하는 기술이다.

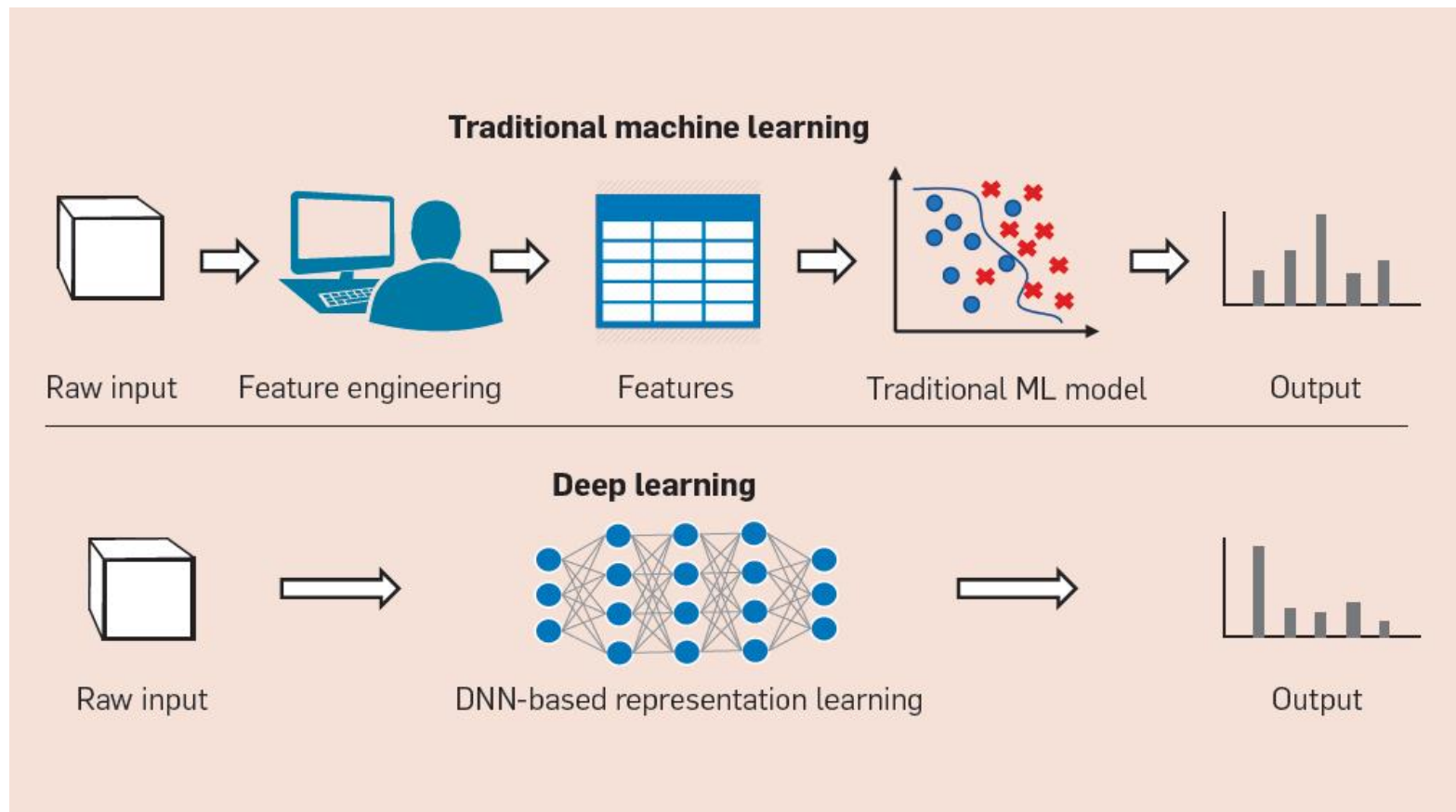
### 딥러닝

- 인공신경망을 이용하여 머신러닝을 수행하는 방법을 일컫는다.



**\*Feature**

- 데이터의 값을 잘 예측하기 위한 데이터의 특징들을 머신러닝/딥러닝에서는 “**Feature**”라고 부르며, 지도, 비지도, 강화학습에서 신뢰할만한 결과를 얻기 위해서는 적절한 **feature**를 잘 정의하는 것이 핵심입니다.



머신러닝의 학습 방법은 크게 3가지로 분류합니다.

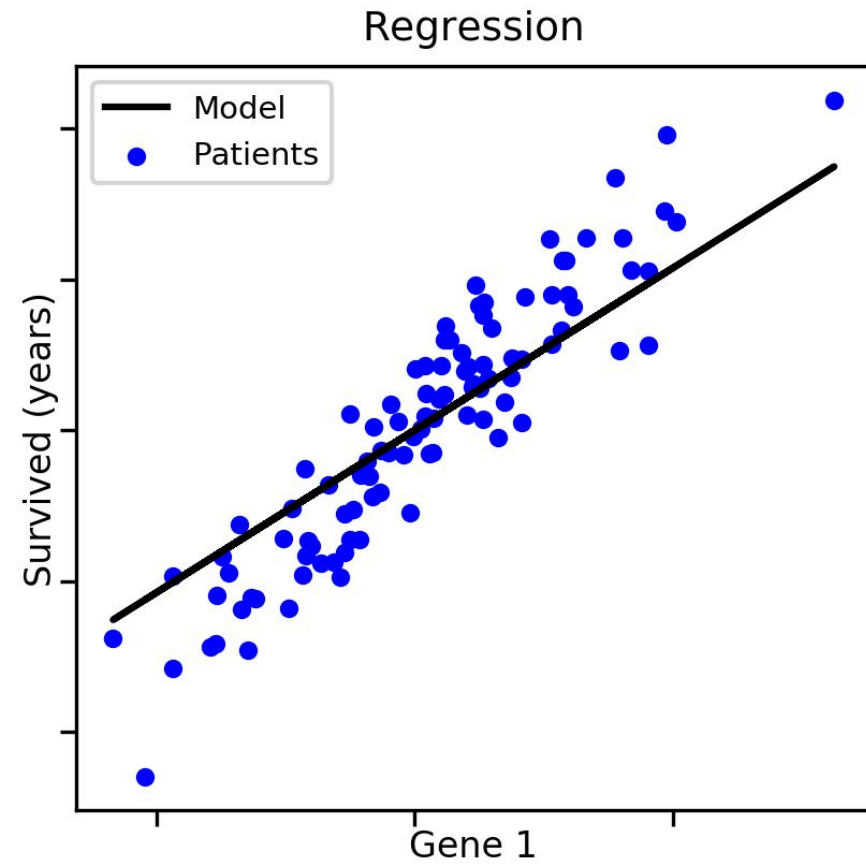
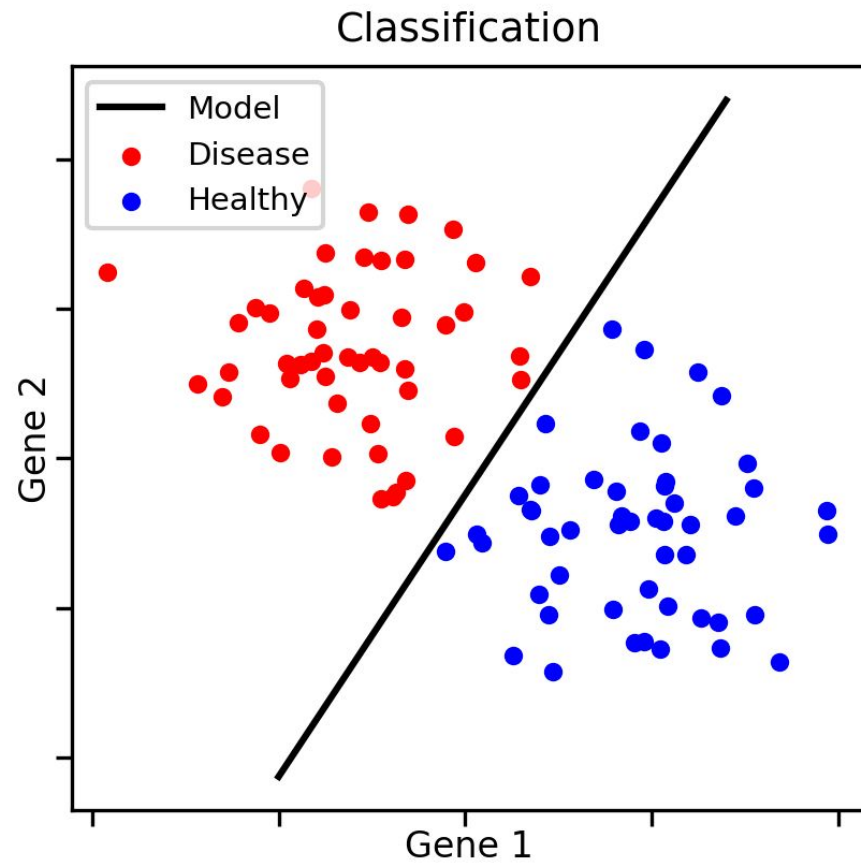
- 지도학습
- 비지도학습
- 강화학습

### 지도학습

- 지도 학습은 말 그대로 **정답이 있는 데이터**를 활용해 데이터를 학습시키는 것입니다. 입력 값(X data)이 주어지면 입력값에 대한 Label(Y data)를 주어 학습시키며 대표적으로 **분류, 회귀** 문제가 있습니다.

### 분류와 회귀

- **분류**는 주어진 데이터를 정해진 카테고리(라벨)에 따라 분류하는 문제를 말합니다. 분류는 맞다, 아니다 등의 이진 분류 문제 또는 사과다 바나나다 포도다 등의 2가지 이상으로 분류하는 다중 분류 문제가 있습니다. **회귀**는 어떤 데이터들의 피쳐(Feature)를 기준으로, 연속된 값(그래프를 예측하는 문제로 주로 어떤 패턴이나 트렌드, 경향을 예측할 때 사용됩니다. 즉 답이 분류 처럼 1,0이렇게 딱 떨어지는 것이 아니고 어떤 수나 실수로 예측될 수 있습니다.

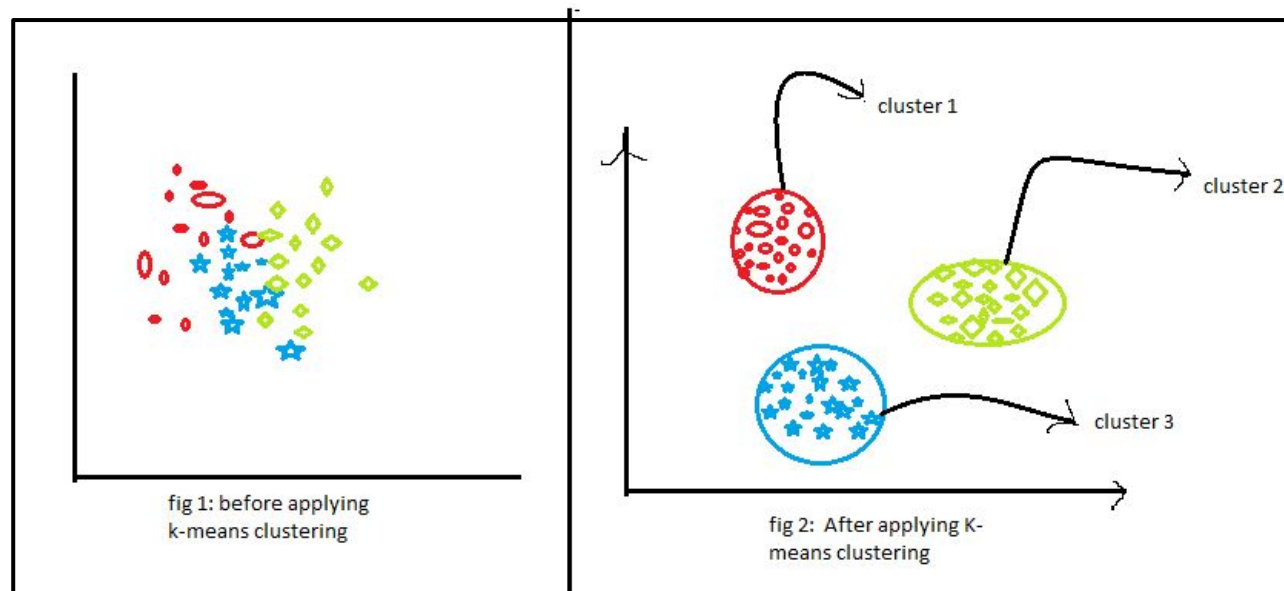
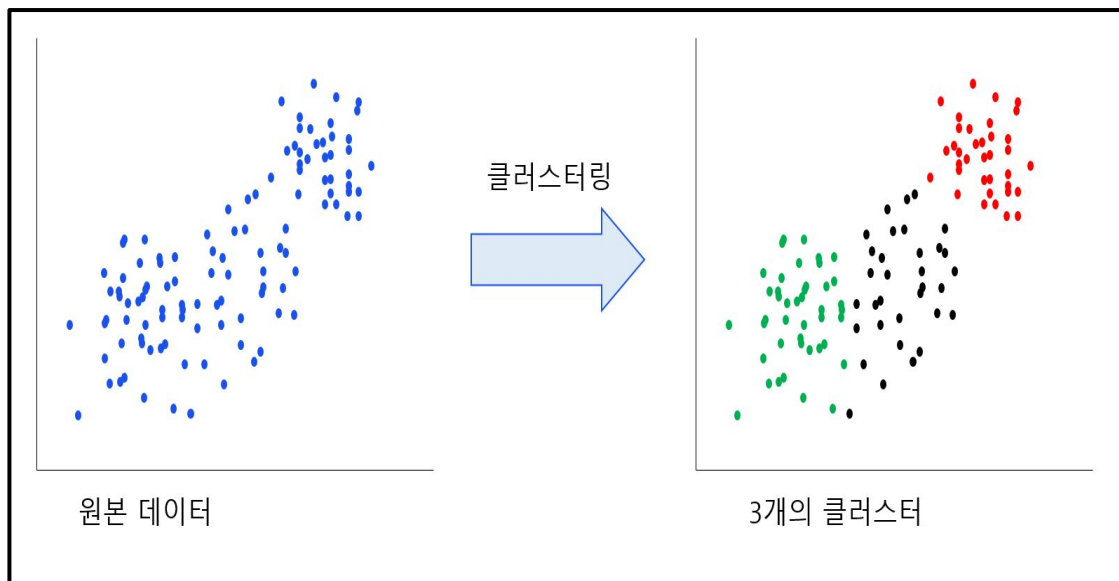


## 비지도학습

- 지도 학습과는 달리 정답 라벨이 없는 데이터를 비슷한 특징끼리 군집화 하여 새로운 데이터에 대한 결과를 예측하는 방법을 비지도학습 이라고 합니다. 라벨링 되어있지 않은 데이터로부터 패턴이나 형태를 찾아야 하기 때문에 지도학습보다는 조금 더 난이도가 있다고 할 수 있습니다.

## 클러스터링(군집화)

- 클러스터링은 데이터가 주어졌을 때, 데이터들을 몇개의 부분 그룹으로 나누는 과정을 의미합니다.



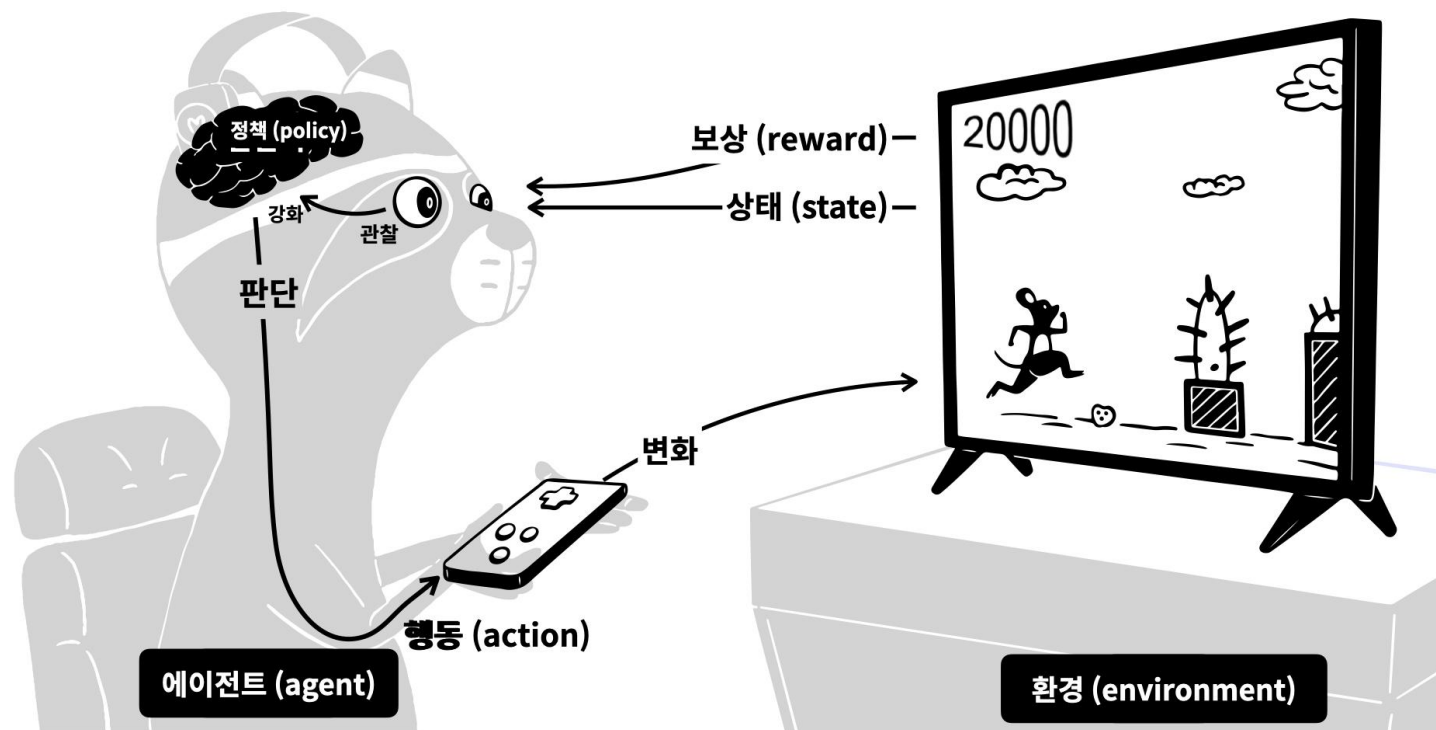


## 강화학습

- 분류할 수 있는 데이터가 존재하는 것도 아니고 데이터가 있어도 정답이 따로 정해져 있지 않으며 자신이 한 행동에 대해 보상(reward)를 받으며 학습하는 것을 말합니다.

### 강화학습의 개념

- 에이전트(Agent)
- 환경(Environment)
- 상태(state)
- 행동(Action)
- 보상(Reward)



# 02

## Machine Learning Traffic Monitoring System



mTMS 봇

[mTMS:거래없음 정상화 - n6re2i]

- ◎OK알람일시: 07/03 12:43
- ◎장애알람일시: 07/03. 12:38
- ◎가맹점: 굿핀[goodp]
- ◎서비스: PG(차세대)
- ◎거래종류: 휴대폰결제 인증[MPA0]
- ◎최근 3분간 거래건수: 4

알람시 13분간 거래 발생 안함

◎실시간 상황 URL 링크: <http://210.112.82.188:8080/t.jsp?key=ybnk2b>



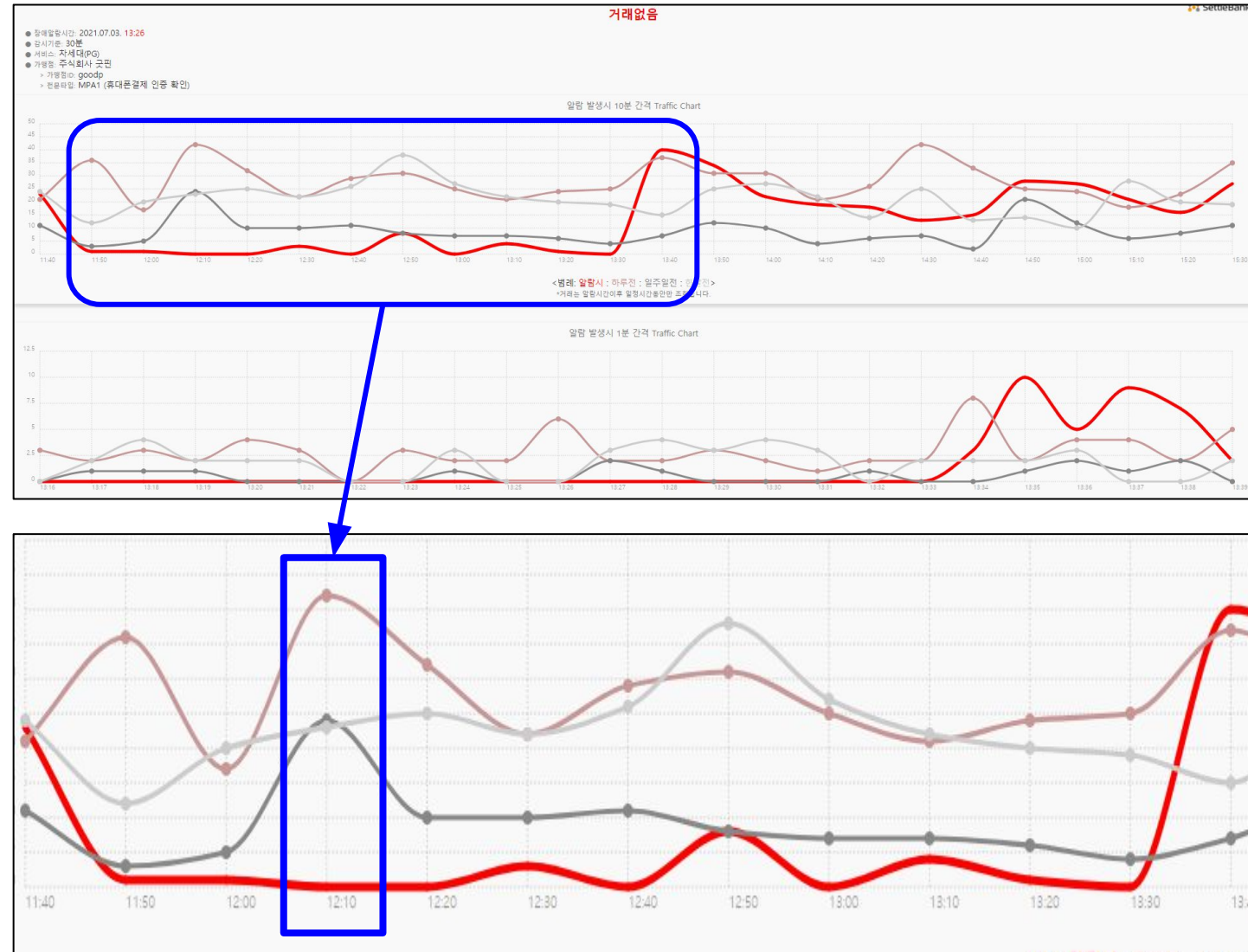
mTMS 봇

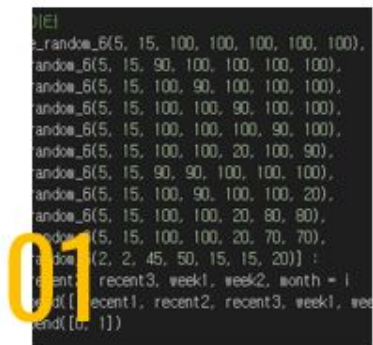
[mTMS:거래없음 - 8pab3b]

- 가맹점: 굿핀[goodp]
- 서비스: PG(차세대)
- 거래종류: 휴대폰결제 인증 확인[MPA1]
- 최근 15분간 거래없음!
- (평상시 거래없음 간격: 4분)

●알람일시: 07/03 13:26

●실시간 상황 URL 링크: <http://210.112.82.188:8080/t.jsp?key=8pab3b>





## 임시 데이터 모델

임시 데이터로 LSTM 모델  
활용



## 실제 데이터 모델

임시 데이터로 학습된  
모델에서 예측된 실제  
데이터의 정상  
비정상 %를 보고 라벨링  
작업 후 LSTM 모델 적용.



## 데이터 분석

실제 데이터 모델에서도  
정확도는 90%이상이지만  
예측률이 좋지않아 데이터  
분석 시작. 문제점 파악 후  
분석 및 라벨링 알고리즘  
개발.



## 최종 모델

정확도 99%, 예측률  
90%이상을 보여주며  
개발을 마침.

수정전

전체	3370	정상	3297	비정상	73
----	------	----	------	-----	----

		실제	
		정상	비정상
예측	정상	3294	25
	비정상	3	48

수정후

전체	3370	정상	3302	비정상	68
----	------	----	------	-----	----

		실제	
		정상	비정상
예측	정상	3302	0
	비정상	0	68



## 수정전

loss(손실값)	0.03	Accuracy(정확도)	0.9800
-----------	------	---------------	--------

[[49, 82, 94, 111, 94, 105]] [0, 1]	1	변태: [[21, 105, 105, 90, 105, 0]] [[0.34325865 0.6567414 ]]
[[39, 77, 105, 111, 105, 94]] [0, 1]	2	변태: [[12, 52, 105, 62, 34, 50]] [[0.00197509 0.99802494]]
[[32, 72, 83, 105, 105, 111]] [0, 1]	3	변태: [[0, 105, 105, 70, 70, 50]] [[0.00130555 0.9986945 ]]
[[36, 105, 100, 107, 105, 84]] [0, 1]	4	
[[39, 105, 86, 113, 113, 93]] [0, 1]	5	
[[49, 105, 75, 116, 105, 83]] [0, 1]	6	
[[30, 106, 105, 112, 105, 75]] [0, 1]	7	
[[14, 60, 105, 70, 70, 70]] [0, 1]	8	
[[14, 56, 105, 78, 70, 70]] [0, 1]	9	
[[0, 105, 105, 70, 70, 50]] [0, 1]	10	
[[0, 105, 114, 70, 70, 70]] [0, 1]	11	변태: [[0, 105, 114, 70, 70, 70]] [[0.00130555 0.9986945 ]]
[[62, 105, 117, 82, 117, 117]] [0, 1]	12	
[[21, 109, 109, 90, 105, 105]] [0, 1]	13	변태: [[21, 109, 109, 90, 105, 105]] [[0.00135355 0.99864644]]
[[21, 100, 81, 105, 105, 63]] [0, 1]	14	변태: [[21, 100, 81, 105, 105, 63]] [[0.2211322 0.7788678 ]]
[[78, 100, 110, 100, 105, 105]] [0, 1]	15	
[[64, 109, 105, 90, 118, 109]] [0, 1]	16	
[[6, 105, 105, 105, 105, 0]] [0, 1]	17	변태: [[6, 105, 105, 105, 105, 0]] [[0.00130146 0.9986985 ]]
[[60, 84, 103, 105, 105, 80]] [0, 1]	18	변태: [[27, 84, 115, 76, 105, 100]] [[2.7393363e-04 9.9972600e-01]]
[[45, 105, 84, 84, 107, 105]] [0, 1]	19	변태: [[0, 105, 105, 0, 105, 105]] [[0.00130482 0.9986952 ]]
[[42, 105, 78, 92, 92, 105]] [0, 1]	20	
[[39, 113, 73, 80, 113, 105]] [0, 1]	21	
[[73, 95, 109, 76, 114, 100]] [0, 1]	22	
[[64, 90, 105, 77, 95, 105]] [0, 1]	23	
[[51, 104, 91, 108, 105, 100]] [0, 1]	24	
[[73, 88, 107, 97, 114, 105]] [0, 1]	25	
[[75, 92, 110, 94, 107, 105]] [0, 1]	26	
[[56, 105, 105, 105, 94, 105]] [0, 1]	27	
[[38, 97, 105, 73, 102, 119]] [0, 1]	28	
[[76, 91, 117, 68, 120, 105]] [0, 1]	29	
[[63, 105, 100, 105, 105, 113]] [0, 1]	30	
[[82, 116, 105, 63, 105, 119]] [0, 1]	31	
[[0, 105, 105, 0, 105, 105]] [0, 1]	32	변태: [[0, 105, 105, 33, 105, 105]] [[0.00130417 0.9986959 ]]
[[0, 105, 105, 33, 105, 105]] [0, 1]	33	
[[0, 83, 108, 75, 105, 105]] [0, 1]	34	변태: [[0, 83, 108, 75, 105, 105]] [[3.4799294e-05 9.9996519e-01]]
[[60, 100, 92, 114, 85, 114]] [0, 1]	35	변태: [[27, 115, 105, 61, 105, 115]] [[0.00325333 0.99674666]]
[[81, 100, 93, 105, 112, 105]] [0, 1]	36	
[[75, 103, 92, 75, 105, 111]] [0, 1]	37	
[[40, 82, 93, 105, 105, 48]] [0, 1]	38	

## 수정후

loss(손실값)	0.001	Accuracy(정확도)	0.9997
-----------	-------	---------------	--------

```

# [[39, 100, 100, 80, 113, 120]] [0, 1]
# [[30, 91, 105, 108, 105, 91]] [0, 1]
# [[37, 112, 50, 100, 105, 118]] [0, 1]

```

```

# [[0, 40, 105, 79, 33, 70]] [0, 1]
# [[32, 81, 45, 105, 105, 118]] [0, 1]
# [[27, 100, 53, 105, 105, 115]] [0, 1]
# [[45, 107, 115, 105, 107, 76]] [0, 1]
# [[39, 83, 105, 105, 83, 91]] [0, 1]
# [[21, 100, 45, 70, 70, 70]] [0, 1]
# [[19, 91, 41, 70, 70, 70]] [0, 1]

```

```

# [[26, 100, 90, 86, 95, 105]] [0, 1]
# [[31, 110, 105, 84, 84, 105]] [0, 1]
# [[31, 110, 105, 100, 63, 105]] [0, 1]
# [[30, 105, 100, 105, 95, 100]] [0, 1]

```

```

# [[14, 102, 105, 101, 53, 105]] [0, 1]
# [[0, 45, 100, 27, 105, 105]] [0, 1]
# [[0, 93, 105, 105, 105, 46]] [0, 1]
# [[0, 82, 105, 17, 105, 105]] [0, 1]
# [[0, 107, 105, 105, 115, 46]] [0, 1]

```

```

# [[39, 105, 73, 100, 105, 105]] [0, 1]
# [[32, 105, 105, 81, 90, 105]] [0, 1]
# [[30, 108, 116, 91, 116, 105]] [0, 1]

```

```

9 # 9 9: [[39, 100, 100, 80, 113, 120]] [[0.04978701 0.950213 ]]
10 # 10 10: [[30, 91, 105, 108, 105, 91]] [[1.0196192e-04 9.9989808e-01]]
11 # 11 11: [[37, 112, 50, 100, 105, 118]] [[1.02035345e-04 9.99897957e-01]]
12 # 12 12: [[0, 78, 105, 70, 70, 0]] [[0.00536502 0.994635 ]]
13 # 13 13: [[0, 92, 105, 70, 70, 0]] [[0.06135615 0.93864393]]
14 # 14 14: [[0, 68, 105, 70, 70, 0]] [[1.0248068e-04 9.9989748e-01]]
15 # 15 15: [[0, 40, 105, 79, 33, 70]] [[1.0195414e-04 9.9989808e-01]]
16 # 16 16: [[32, 81, 45, 105, 105, 118]] [[0.00424816 0.99575186]]
17 # 17 17: [[27, 100, 53, 105, 105, 115]] [[1.0195424e-04 9.9989808e-01]]
18 # 18 18: [[45, 107, 115, 105, 107, 76]] [[0.00690449 0.99309546]]
19 # 19 19: [[39, 83, 105, 105, 83, 91]] [[3.4787305e-04 9.9965215e-01]]
20 # 20 20: [[21, 100, 45, 70, 70, 70]] [[0.00363879 0.9963612 ]]
21 # 21 21: [[19, 91, 41, 70, 70, 70]] [[1.1316184e-04 9.9988687e-01]]
22 # 22 22: [[0, 105, 105, 70, 26, 70]] [[0.00730239 0.99269754]]
23 # 23 23: [[26, 100, 90, 86, 95, 105]] [[1.01966485e-04 9.99898076e-01]]
24 # 24 24: [[31, 110, 105, 84, 84, 105]] [[0.00673056 0.9932694 ]]
25 # 25 25: [[31, 110, 105, 100, 63, 105]] [[0.14531459 0.85468537]]
26 # 26 26: [[30, 105, 100, 105, 95, 100]] [[1.02046346e-04 9.99897957e-01]]
27 # 27 27: [[30, 60, 80, 105, 105, 105]] [[0.21118811 0.78881186]]
28 # 28 28: [[14, 102, 105, 101, 53, 105]] [[1.0695771e-04 9.9989307e-01]]
29 # 29 29: [[0, 45, 100, 27, 105, 105]] [[0.00427067 0.99572927]]
30 # 30 30: [[0, 93, 105, 105, 105, 46]] [[1.0195414e-04 9.9989808e-01]]
31 # 31 31: [[0, 82, 105, 17, 105, 105]] [[0.00351803 0.9964819 ]]
32 # 32 32: [[0, 107, 105, 105, 115, 46]] [[1.0195414e-04 9.9989808e-01]]
33 # 33 33: [[0, 116, 105, 58, 8, 105]] [[0.01774505 0.982255 ]]
34 # 34 34: [[39, 105, 73, 100, 105, 105]] [[1.02927654e-04 9.99897003e-01]]
35 # 35 35: [[32, 105, 105, 81, 90, 105]] [[0.00255538 0.9974446 ]]
36 # 36 36: [[30, 108, 116, 91, 116, 105]] [[1.0195988e-04 9.9989808e-01]]

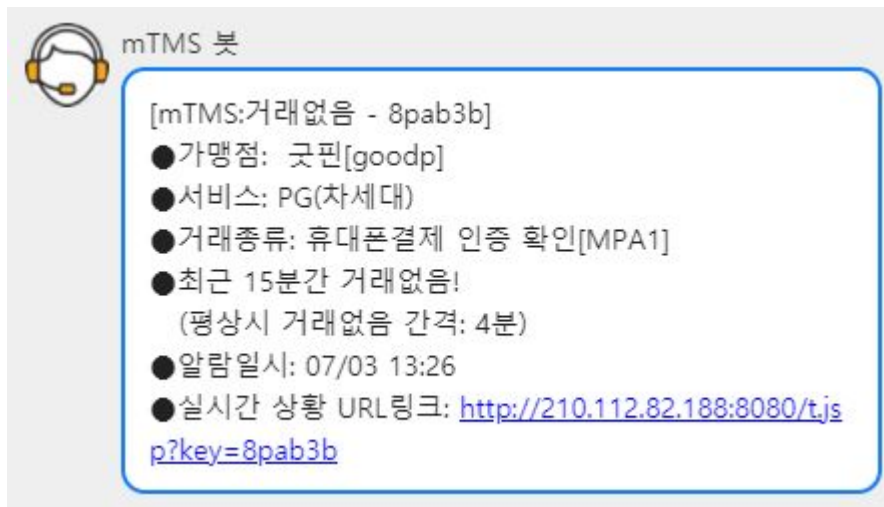
```

# 03

## 개선할 점



## 기존 시스템

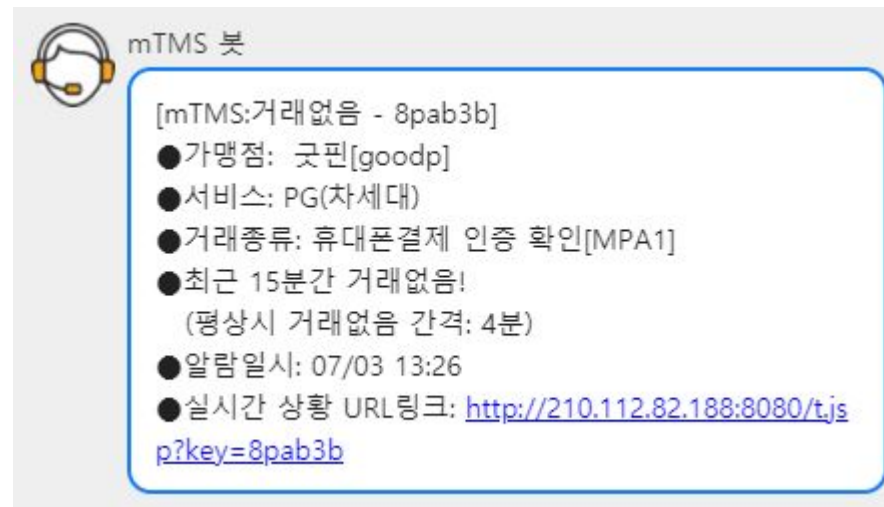


가맹점 이름

서비스 및 거래종류

실시간 상황 **URL**링크

## 추가할 점



## 기사 및 시사적 문제 **URL** 링크

- 거래량이 비정상적으로 확인 되어  
알람이 뜨는 회사의 이름을 네이버나  
특정 사이트에서 크롤링하게 한다.  
모델은 기사의 내용을 읽고 판단하여  
거래량에 영향을 준다고 판단되는  
기사를 우리에게 알려준다.

# 04

## 나의

# 05

## 마무리