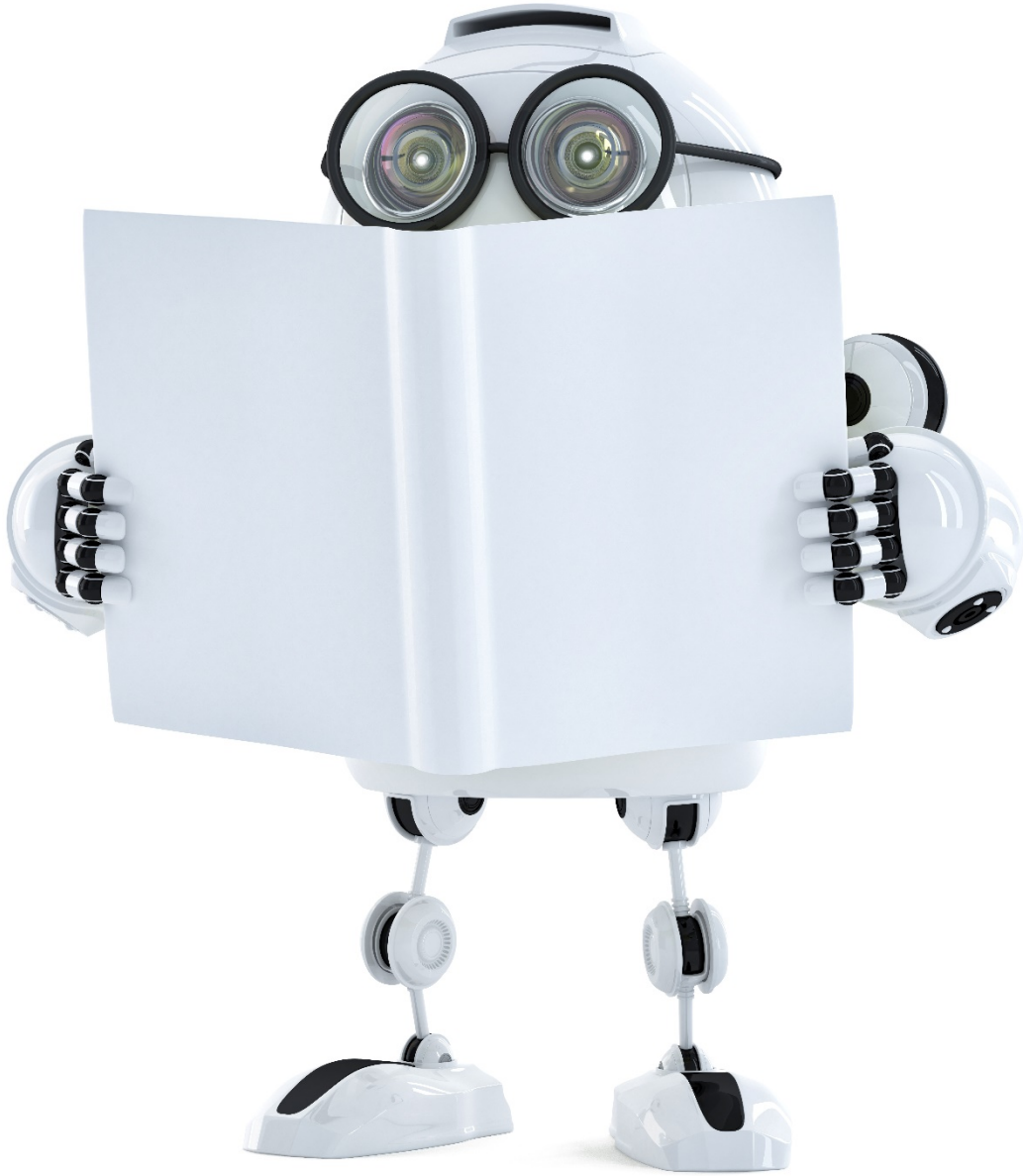
Overview

Logistic Regression Classifier



Director of TEAMLAB
Sungchul Choi

머신러닝의 학습 방법들

- Gradient descent based learning
- Probability theory based learning
- Information theory based learning
- Distance similarity based learning

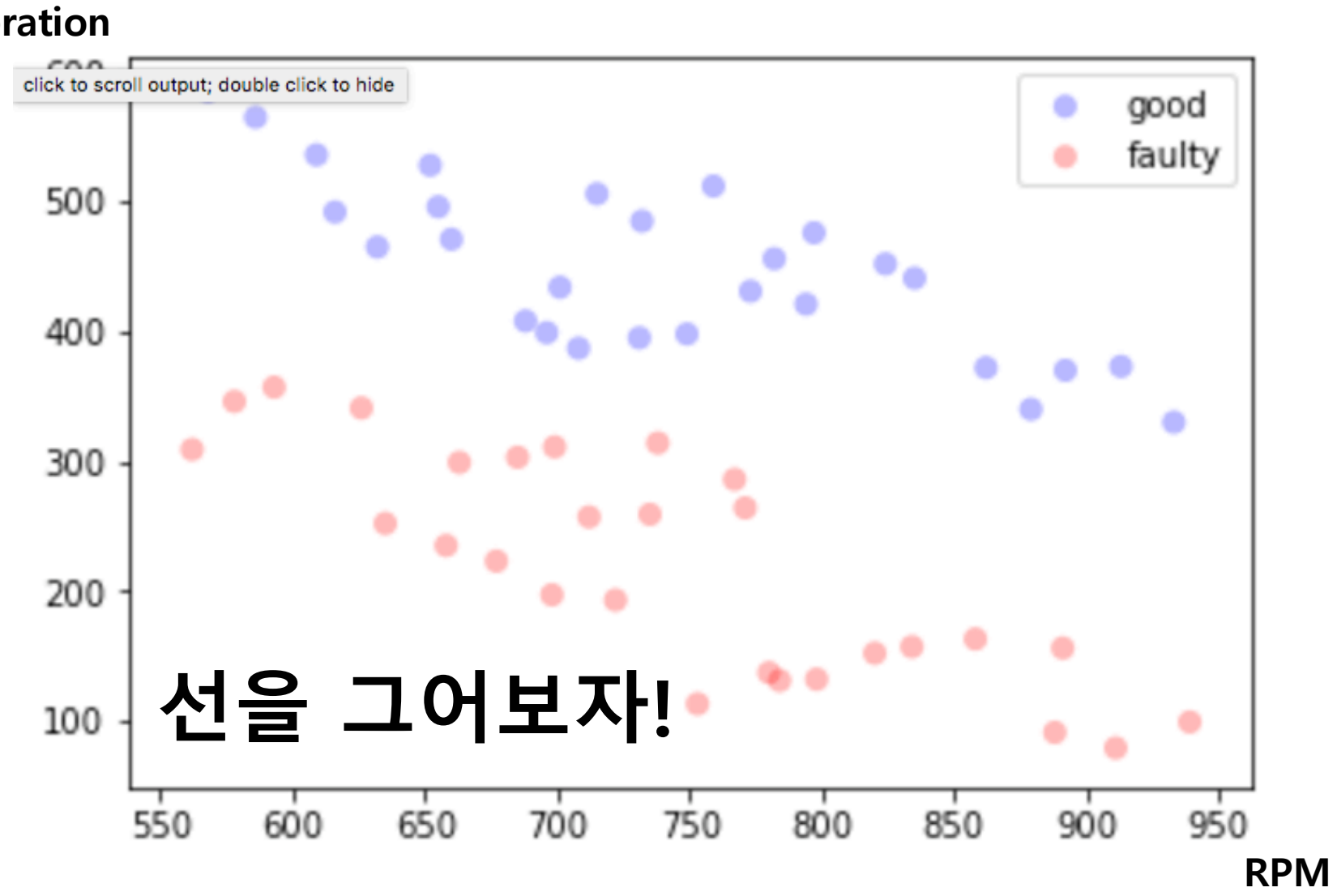
Classification Problem

- 발전소 발전기가 제대로 돌아가고 있는가?

ID	RPM	VIBRATION	STATUS
1	568	585	good
2	586	565	good
3	609	536	good
4	616	492	good
5	632	465	good
6	652	528	good
7	655	496	good
8	660	471	good
9	688	408	good
10	696	399	good
11	708	387	good
12	701	434	good
13	715	506	good
14	732	485	good
15	731	395	good
16	749	398	good
17	759	512	good
18	773	431	good
19	782	456	good
20	797	476	good
21	794	421	good
22	824	452	good
23	835	441	good
24	862	372	good
25	879	340	good
26	892	370	good
27	913	373	good
28	933	330	good

ID	RPM	VIBRATION	STATUS
29	562	309	faulty
30	578	346	faulty
31	593	357	faulty
32	626	341	faulty
33	635	252	faulty
34	658	235	faulty
35	663	299	faulty
36	677	223	faulty
37	685	303	faulty
38	698	197	faulty
39	699	311	faulty
40	712	257	faulty
41	722	193	faulty
42	735	259	faulty
43	738	314	faulty
44	753	113	faulty
45	767	286	faulty
46	771	264	faulty
47	780	137	faulty
48	784	131	faulty
49	798	132	faulty
50	820	152	faulty
51	834	157	faulty
52	858	163	faulty
53	888	91	faulty
54	891	156	faulty
55	911	79	faulty
56	939	99	faulty

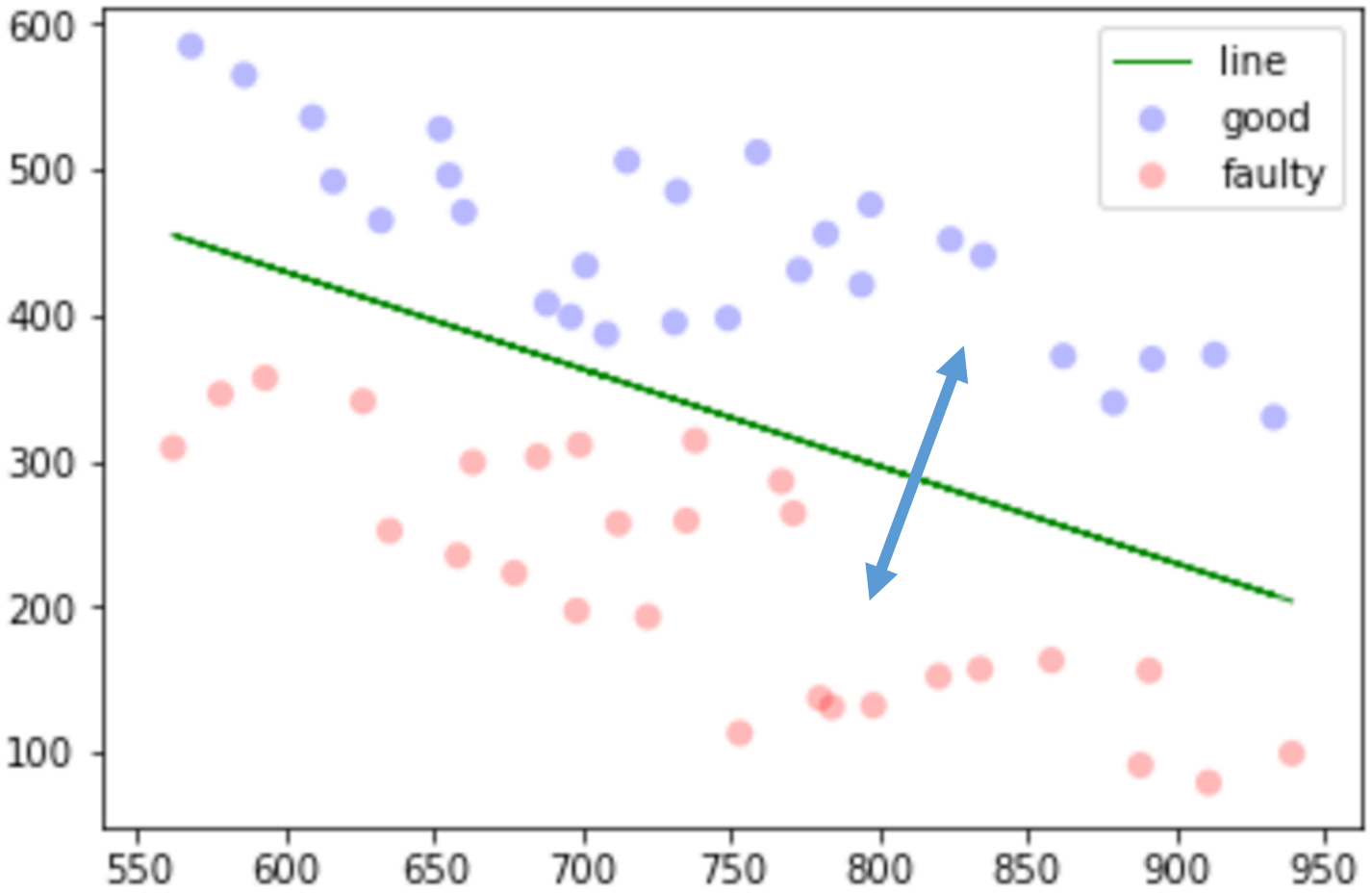
Classification Problem



Classification Problem

Vibration

$$\text{Vibration} = 830 - 0.667 * \text{RPM}$$



RPM

Classification Problem

$$f(x) = 803 - 0.667 * RPM - 1 * Vibration = 0$$

$$Status = \begin{cases} 1 & \text{if } f(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Classification Problem

어떻게 학습을 시킬까?

$$f(x) = \underline{803} - \underline{0.667} * RPM - \underline{1} * Vibration = 0$$

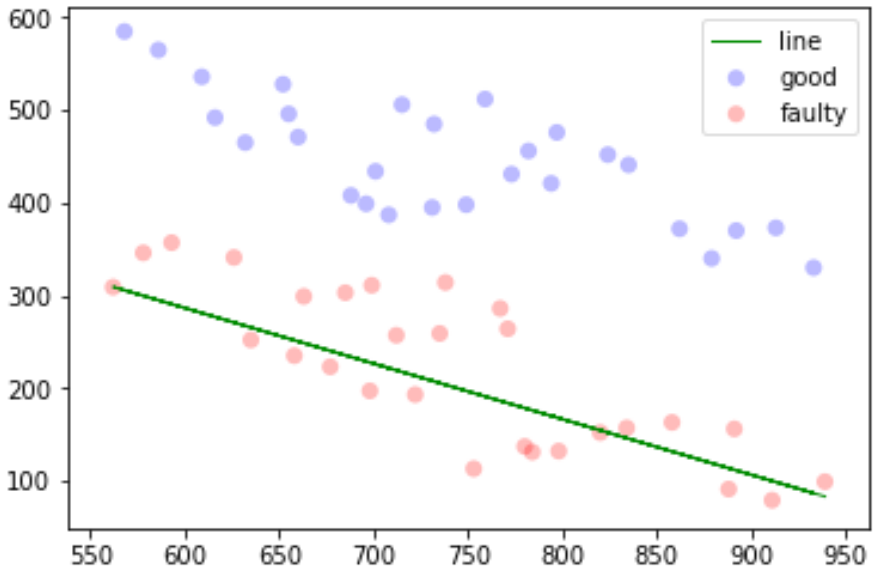
$$Status = \begin{cases} 1 & \text{if } f(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Classification Problem

Linear Regression으로 학습해보자

$$f(x) = -2.450750883259760 6 + 0.00227488 * RPM + 0.00379006 * Vibration$$

```
array([  1.05856314e+00,   1.02370973e+00,   9.66120163e-01,
click to scroll output; double click to hide
         9.19162092e-01,   8.35784996e-01,   6.60707802e-01,
         6.44796277e-01,   6.26614080e-01,   7.88822723e-01,
         1.09355523e+00,   1.05263689e+00,   7.09256697e-01,
         7.61574640e-01,   1.21639013e+00,   9.41243609e-01,
         1.05646897e+00,   1.16639329e+00,   9.51115412e-01,
         1.13685352e+00,   1.12018650e+00,   9.20094079e-01,
         8.37485080e-01,   9.80760238e-01,   1.03990281e+00,
         9.22427788e-01,  -1.14240215e-03,   1.75487796e-01,
         2.51301583e-01,   2.65731543e-01,  -5.11098276e-02,
        -6.32186858e-02,   1.90719471e-01,  -6.54767525e-02,
         2.55926977e-01,  -1.16245894e-01,   3.18095711e-01,
         1.43005910e-01,  -7.68091088e-02,   2.02908173e-01,
         4.18186047e-01,  -3.09492679e-01,   3.78035795e-01,
         3.03754000e-01,  -1.57109613e-01,  -1.70750464e-01,
        -1.35112143e-01,  -9.26369286e-03,   4.15348645e-02,
         1.18872240e-01,  -8.57657338e-02,   1.67412730e-01,
        -7.89242969e-02,   6.05734081e-02])
```



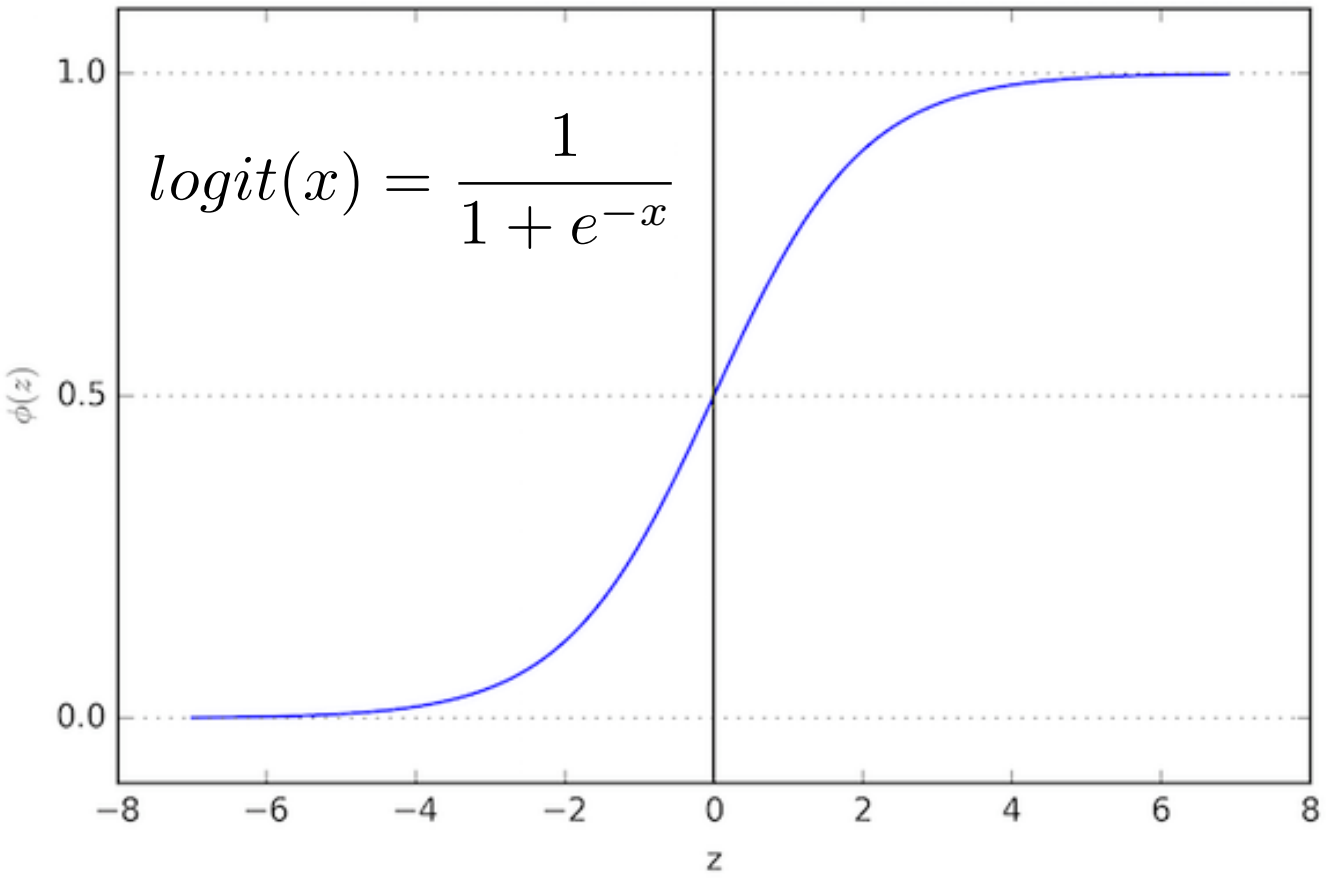
기존 접근의 문제점들

- 1이상 또는 0이하의 수들이 나오는 걸 어떻게 해석?
- 1 또는 0으로 정확히 표현 가능한가? $Status = \begin{cases} 1 & \text{if } f(x) \geq 0 \\ 0 & \text{otherwise} \end{cases}$
- 변수가 Y에 영향을 주는정도가 비례하는 가?
- **확률로** 발생할 사건의 가능성을 표현해야 함

$$f(x) = -2.450750883259760 6 + 0.00227488 * RPM + 0.00379006 * Vibration$$

Solution

확률로 나타내자!



Human knowledge belongs to the world.