

---

PART 1

# 단층 퍼셉트론(SLP)

1장 회귀분석

2장 이진판단

3장 선택분류

---

# 이 장에서 다룰 내용

1. 펄서 판정 문제
2. 이진 판단 문제와 시그모이드 함수
3. 확률분포, 정보 엔트로피, 교차 엔트로피
4. 시그모이드 교차 엔트로피 편미분
5. 계산값 폭주와 시그모이드 관련 함수의 안전한 계산법
6. 펄서 여부 판정 신경망 구현과 실행

실습에 사용할 데이터셋 : Pulsar Dataset

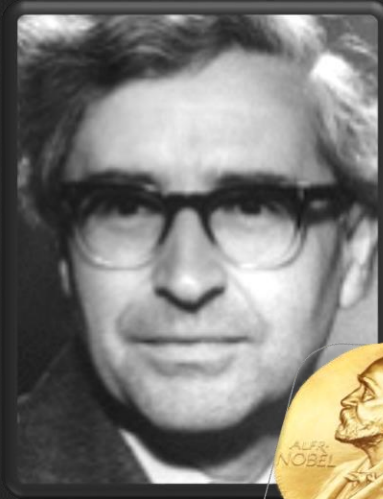
접근 방법 : Google ▶ 'pulsar dataset' 검색 ▶ kaggle에서 다운로드

## 2.1 펄서 판정 문제

1968년



Dame Susan  
Jocelyn Bell Burnell  
(1943년 7월 15일 ~ )



Anthony Hewish  
(1924년 5월 11일 ~ )



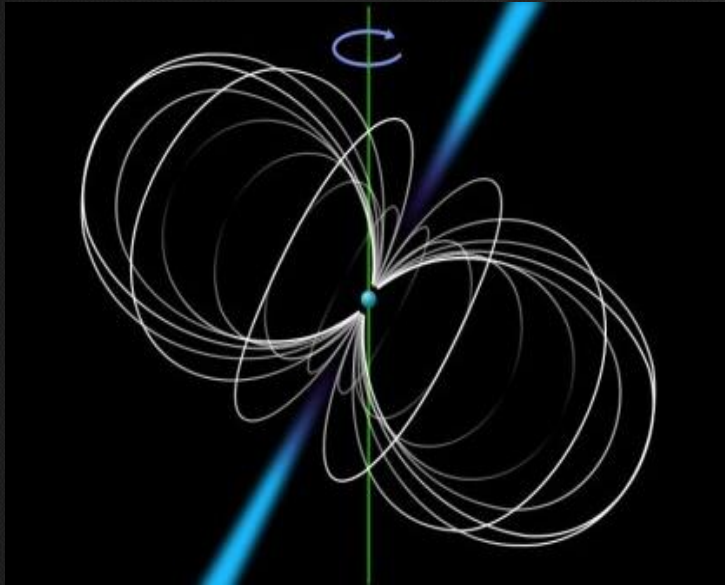
[ 전자기파 등을 방출하는 펄서 ]



[ 보통의 별 ]



## 2.1 펄서 판정 문제



[펄서의 개념도]



# 그럼 어떤 천체가 펄서지?

▶ 전자기파 변화 탐색

너무 많은 정보가 포착



### 2.1 펄서 판정 문제

Mean of the integrated profile

Standard deviation of the integrated profile

Excess kurtosis of the integrated profile

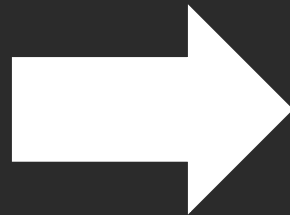
Standard deviation of the DM-SNR curve

Mean of the DM-SNR curve

Excess kurtosis of the DM-SNR curve

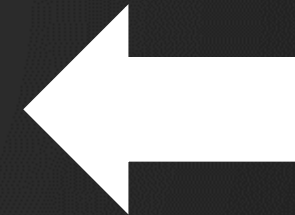
Skewness of the integrated profile

Skewness of the DM-SNR curve



천체 17,898개의 기록  
8가지 특징값

데이터 기반



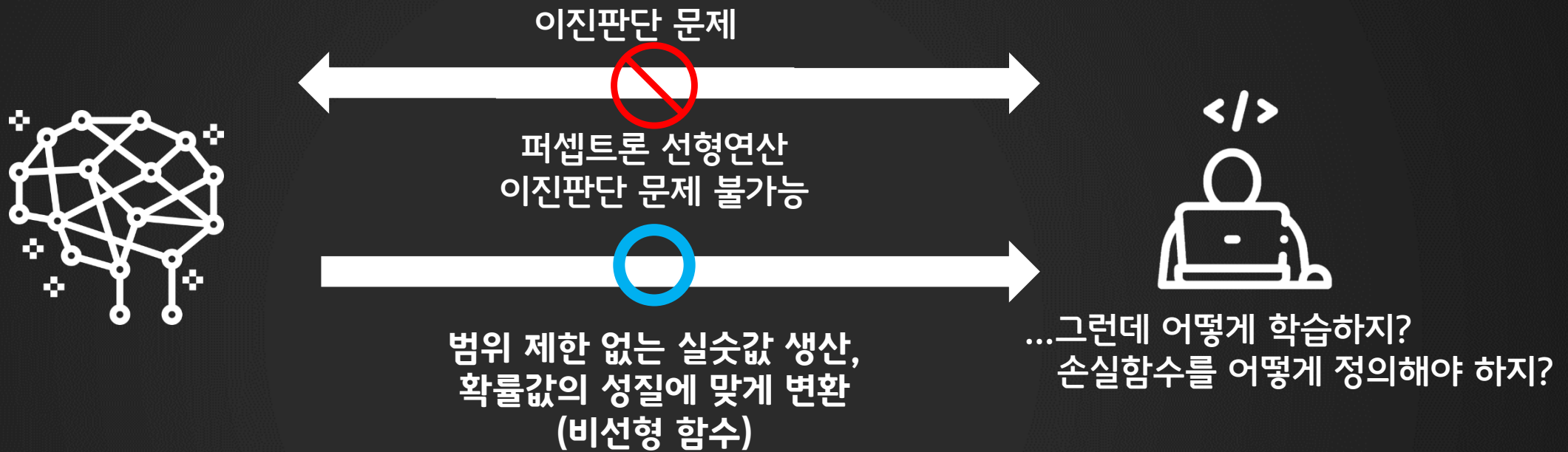
딥러닝  
문제해결

target\_class



0 / 1

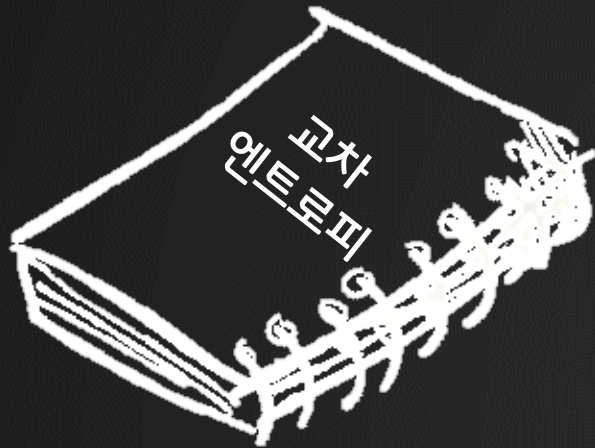
### 2.2 이진 판단 문제의 신경망 처리



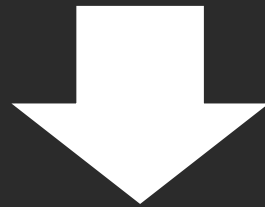
한동안의 딥러닝 발전을 가로막은 '난제'



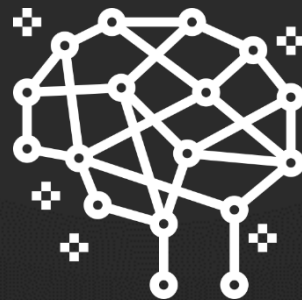
### 2.2 이진 판단 문제의 신경망 처리



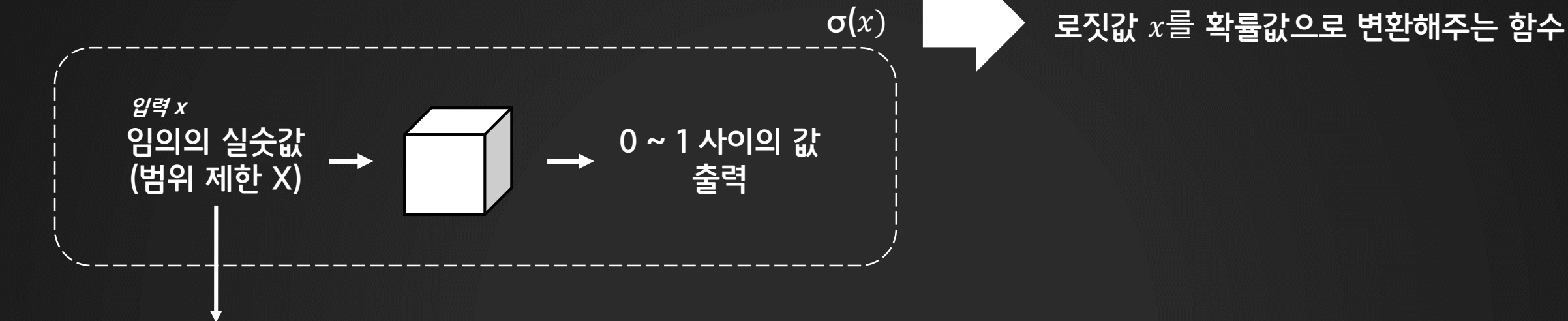
- 두 가지 확률분포가 얼마나 다른지를 숫자 하나로 표현
- 항상 양수
- 두 확률분포가 비슷해질수록 값이 작아짐



이진 판단문제 적합  
신경망 학습가능



### 2.3 시그모이드 함수



#### 어떤 확률값의 로짓(logit) 표현

- 실제 표현하려는 값을 로그값으로
- 로짓값은 상대적

EX) [롤드컵에서 T1이 우승할 확률의 로짓값은 2 다.] 는 무엇을 뜻할까요?

[롤드컵에서 GEN.G가 우승할 확률의 로짓값은 5 다.]

[롤드컵에서 G2가 우승할 확률의 로짓값은 1 이다.]

} 비교 로짓값

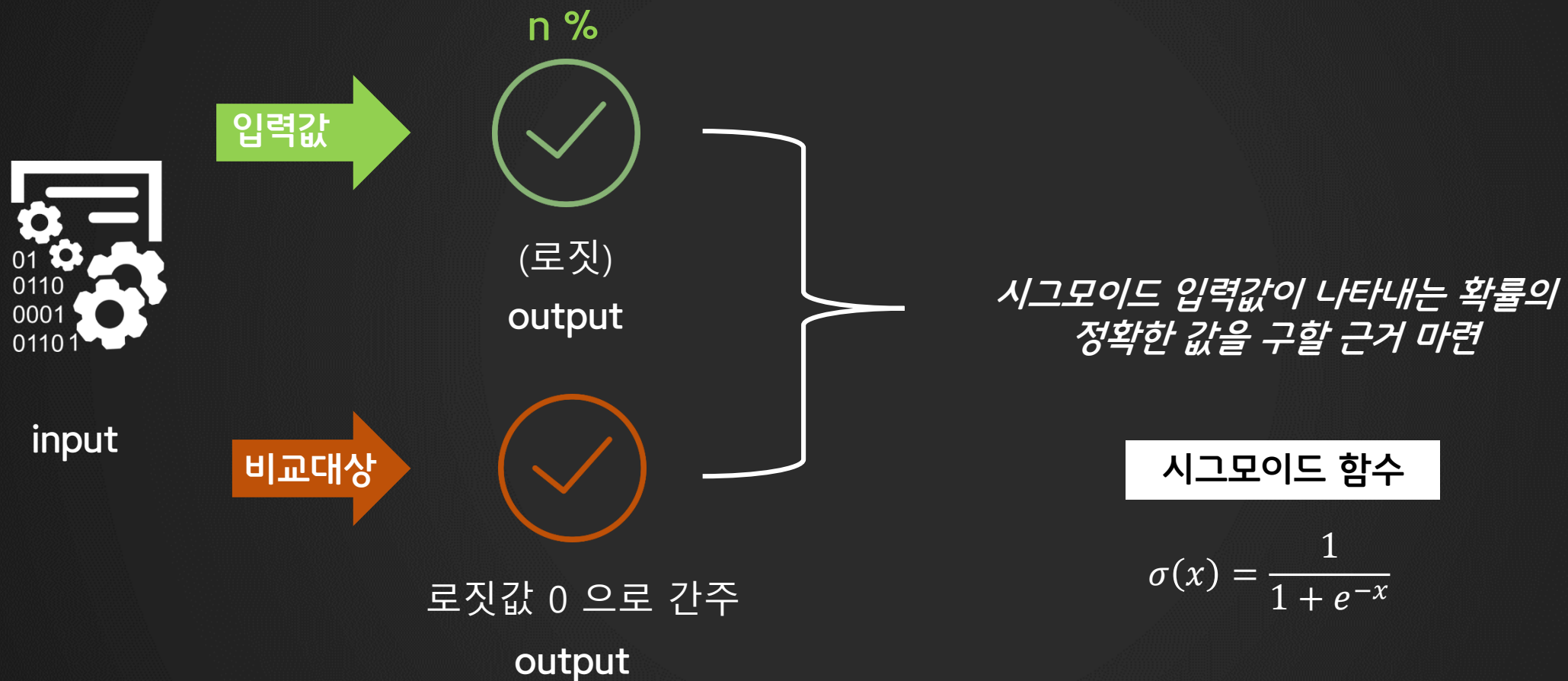
로그함수는 지수함수의 역함수 ->

GEN.G가 우승할 확률은 T1이 우승할 확률보다  $e^{5-3}=e^2 = 20.08$  배 가량 크다.

T1이 우승할 확률은 G2이 우승할 확률보다  $e^{2-1}=e^1 = 2.7$  배 가량 크다.



### 2.3 시그모이드 함수



2.3 시그모이드 함수

시그모이드 함수 정의식 도출 과정

일반화 과정



이 상자가 내 상자일 가능성은 로짓값 0.5

내 상자일 가능성 : 내 상자가 아닐 가능성

$$e^{0.5} : e^0$$

$$1.649 : 1$$



두 경우의 확률합 1

내 상자일 가능성

$$= \frac{1.469}{1.649+1} = 0.622$$

내 상자가 아닐 가능성

$$= \frac{1}{1.649+1} = 0.378$$

	참	거짓
가능성의 로짓값	$x$	0
실제 확률	$P_T = ae^x$	$P_F = a$



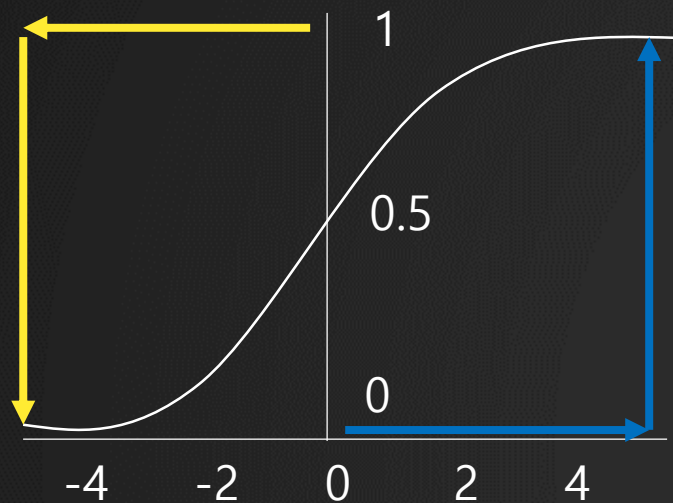
$$P_T + P_F = 1$$

$$ae^x + a = 1$$
$$a(e^x + 1) = 1$$

$$a = \frac{1}{e^x + 1}$$

$$P_T = \frac{e^x}{e^x + 1} = \frac{1}{1 + e^{-x}}$$

### 2.3 시그모이드 함수



시그모이드 함수의 그래프

*차후 파라미터 수정 불필요*

0.51의 확률로 '참' 선택 = 0.99의 확률로 '참' 선택

*차후 파라미터 대폭 수정*



학습의 진행



0 or 1



2.3 시그모이드 함수

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$f(x) = \frac{1}{x} = x^{-1}$$

$$g(x) = 1 + e^1 \rightarrow f(g(h(x)))$$

$$h(x) = -x$$

시그모이드 함수의 도함수

시그모이드 함수의 미분

$$\sigma'(x) = \frac{df}{dg} \cdot \frac{dg}{dh} \cdot \frac{dh}{dx}$$

$$= -\frac{1}{g(h(x))^2} \cdot e^{h(x)} \cdot -1 = \frac{e^{-x}}{(1+e^{-x})^2}$$

$$= \frac{1}{1+e^{-x}} \cdot \frac{1+e^{-x}-1}{1+e^{-x}} = \frac{1}{1+e^{-x}} \left(1 - \frac{1}{1+e^{-x}}\right)$$

$$= \sigma(x)(1 - \sigma(x))$$

※ 미분의 성질

$$f(x) = \frac{1}{x} \rightarrow \frac{df}{dx} = -\frac{1}{x^2}$$

$$f_c(x) = c + x \rightarrow \frac{df}{dx} = 1$$

$$f(x) = e^x \rightarrow \frac{df}{dx} = e^x$$

$$f_a(x) = ax \rightarrow \frac{df}{dx} = a$$

## 2.4 확률 분포와 정보 엔트로피

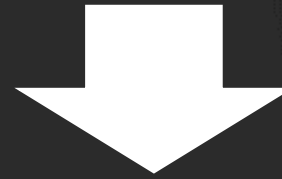


Claude Elwood Shannon

(1916년 4월 30일 ~  
2001년 2월 24일)

『Entropy』

분자들의 무질서도 or 에너지의 분산 정도

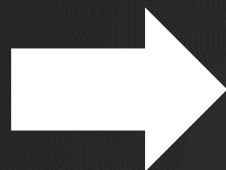


확률 분포의 무질서도나 불확실성 혹은  
정보 표현의 부담 정도를 나타내는 정보 엔트로피 개념 고안

## 2.4 확률 분포와 정보 엔트로피

### 정보 엔트로피

정보량  
(엔트로피)



불확실한 정보를 숫자로 정량화하려는 노력

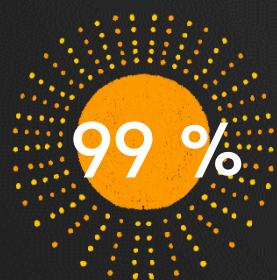
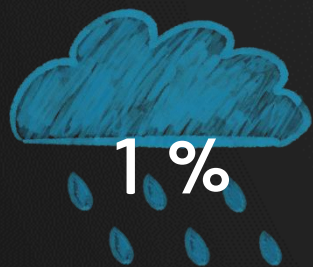
$$h(x) = \log \frac{1}{p(x)} = -\log p(x)$$

※ 어떠한 사건  $p(x)$   
엔트로피  $h(x)$

※ 확률에 역수를 취하는 이유

‘확률이 높다.’ == ‘높은 확률로 그 일이 일어 난다.’

↳ 당연히 여겨진다. == 새로운 정보(량)은 적다.



$$h(\text{비}) = -\log 0.01 = 4.605 \quad \text{정보량} \blacktriangle$$

$$h(\text{맑음}) = -\log 0.99 = 0.010 \quad \text{정보량} \blacktriangledown$$



### 2.4 확률 분포와 정보 엔트로피

엔트로피의 기댓값

$$E(X) = -\underline{p(x)} \log p(x)$$

```
Rain_E = -0.01*np.log(0.01)
```

```
Sunny_E = -0.99*np.log(0.99)
```

```
print(Rain_E)
```

```
print(Sunny_E)
```

```
0.04605170185988091
```

```
0.009949832494966436
```

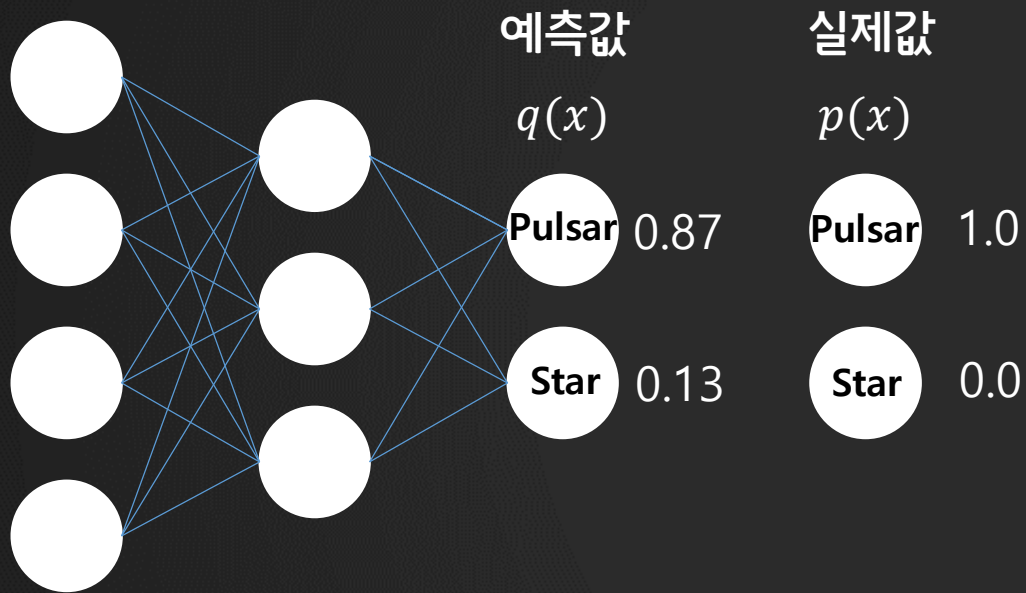
*#불확실한 정보를 수치화*

통계에서의 기댓값

### 2.5 교차 엔트로피

엔트로피의 기댓값  $\neq$  교차 엔트로피

### 2.5 교차 엔트로피

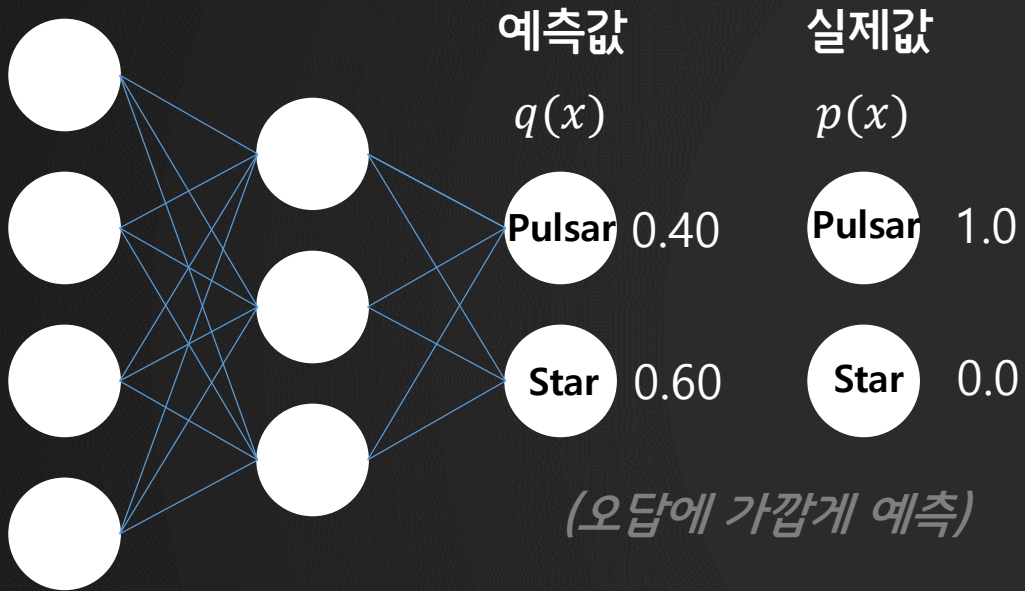


$$CCE = -\frac{1}{n} \sum_{j=1}^n p(x) \log q(x)$$

$$CCE_1 = -\frac{1}{2} (1.0 * \log 0.87 + 0.0 * \log 0.13) = 0.0696 \dots$$

```
CCE_1 = -0.5*(1*np.log(0.87)+0.0*np.log(0.13))  
print(CCE_1)  
0.06963103366675383
```

2.5 교차 엔트로피



$$CCE = -\frac{1}{n} \sum_{j=1}^n p(x) \log q(x)$$

$$CCE_1 = -\frac{1}{2} (1.0 * \log 0.4 + 0.0 * \log 0.6) = 0.0696 ...$$

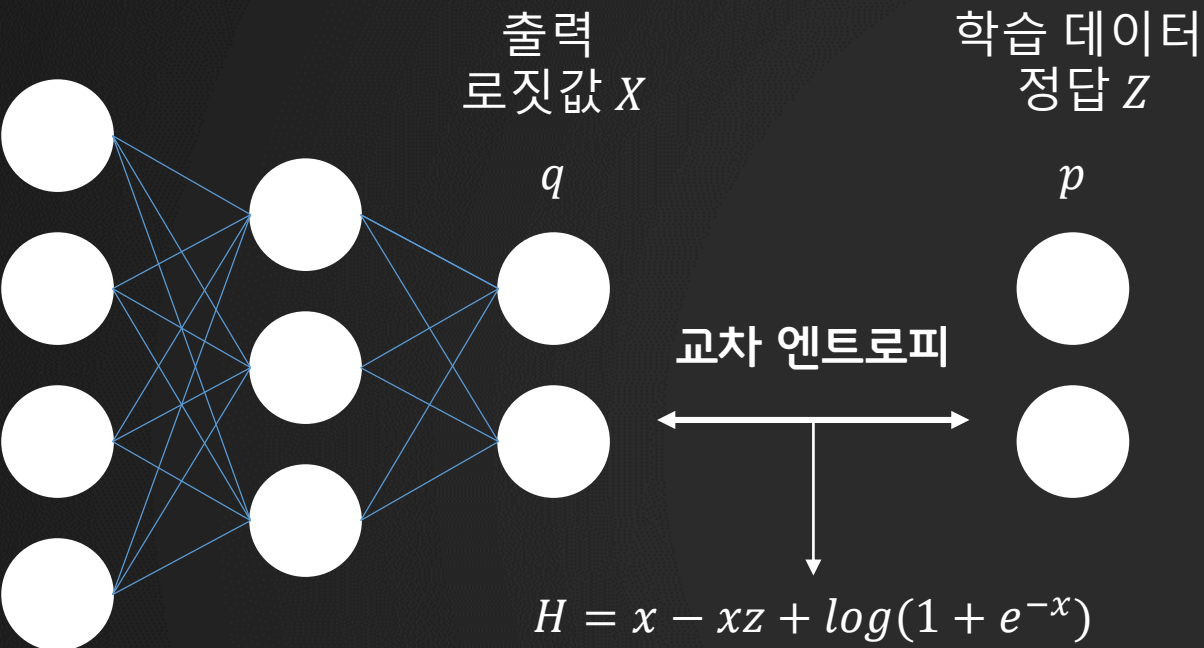
$$CCE_2 = -\frac{1}{2} (0.0 * \log 0.4 + 1.0 * \log 0.6) = 0.4581 ...$$

손실의 일종 ←

★ 신경망 학습의 원리



2.6 시그모이드 교차 엔트로피와 편미분



시그모이드 교차 엔트로피 정의식

데이터의 결과가 참일 확률  $p_T$ , 거짓일 확률  $p_F$

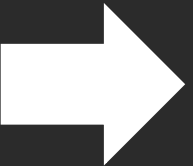
$$p_T = Z, \quad p_F = 1 - Z \quad (* Z = 0 \text{ or } 1)$$
$$q_T = \sigma(x), \quad q_F = 1 - \sigma(x)$$

대입

$$H = - \sum p(x) \log q(x)$$

$$= -p_T \log q_T - p_F \log q_F$$
$$= -z \log \sigma(x) - (1 - z) \log (1 - \sigma(x))$$

시그모이드 함수의 정의  
 $\sigma(x) = 1 / 1 + e^{-x}$



$$H = -z \log \frac{1}{1 + e^{-x}} - (1 - z) \log (1 - \frac{1}{1 + e^{-x}})$$
$$\vdots$$

$z = 0$ 일 때  $H = x + \log(1 + e^{-x})$

$z = 1$ 일 때  $H = \log(1 + e^{-x})$

-----  $= x - xz + \log(1 + e^{-x})$

2.6 시그모이드 교차 엔트로피와 편미분

$$H = x - xz + \log(1 + e^{-x})$$
$$\sigma(x) = 1 / 1 + e^{-x}$$



```
import numpy as np

def sigmoid(x):
    return 1.0 / (1.0+np.exp(-x))

print(sigmoid(1.0))
print(sigmoid(-1000))

0.7310585786300049
0.0
```

RuntimeWarning: overflow encountered in exp after removing the cwd from sys.path.

"x 가 음수일 때 한하여 계산을 달리해 보자!



시그모이드

$$\sigma(x) = \frac{1}{1 + e^{-x}} * \frac{e^x}{e^x} = \frac{e^x}{1 + e^x}$$

시그모이드 교차 엔트로피 함수

$$H = x - xz + \log(1 + e^{-x})$$
$$= -xz + \log e^x (1 + e^{-x}) = -xz + \log(e^x + 1)$$

음수 / 양수



부호에 맞춰 처리  
일괄 처리

미니배치 처리





2.6 시그모이드 교차 엔트로피와 편미분

변형된 식

기존 정의식

$$\sigma(x) = \frac{e^{-\max(-x,0)}}{1 + e^{-|x|}}$$

$x = 0$ 
$$\frac{e^{-\max(-0,0)}}{1 + e^{-|0|}} = \frac{1}{1 + 1}$$

$x = 3$ 
$$\frac{e^{-\max(-3,0)}}{1 + e^{-|3|}} = \frac{e^0}{1 + e^{-3}} = \frac{1}{e^{-3} + 1}$$

$x = -3$ 
$$\frac{e^{-\max(-(-3),0)}}{1 + e^{-|-3|}} = \frac{e^{-3}}{1 + e^{-3}} = \frac{1}{e^3 + 1}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$\frac{1}{1 + e^{-0}} = \frac{1}{1 + 1}$

$\frac{1}{1 + e^{-(3)}} = \frac{1}{1 + e^{-3}}$

$\frac{1}{1 + e^{-(-3)}} = \frac{1}{1 + e^3}$

※  $\sigma(x) = \frac{1}{1+e^{-x}}$  에 대한 교차 엔트로피의 편미분

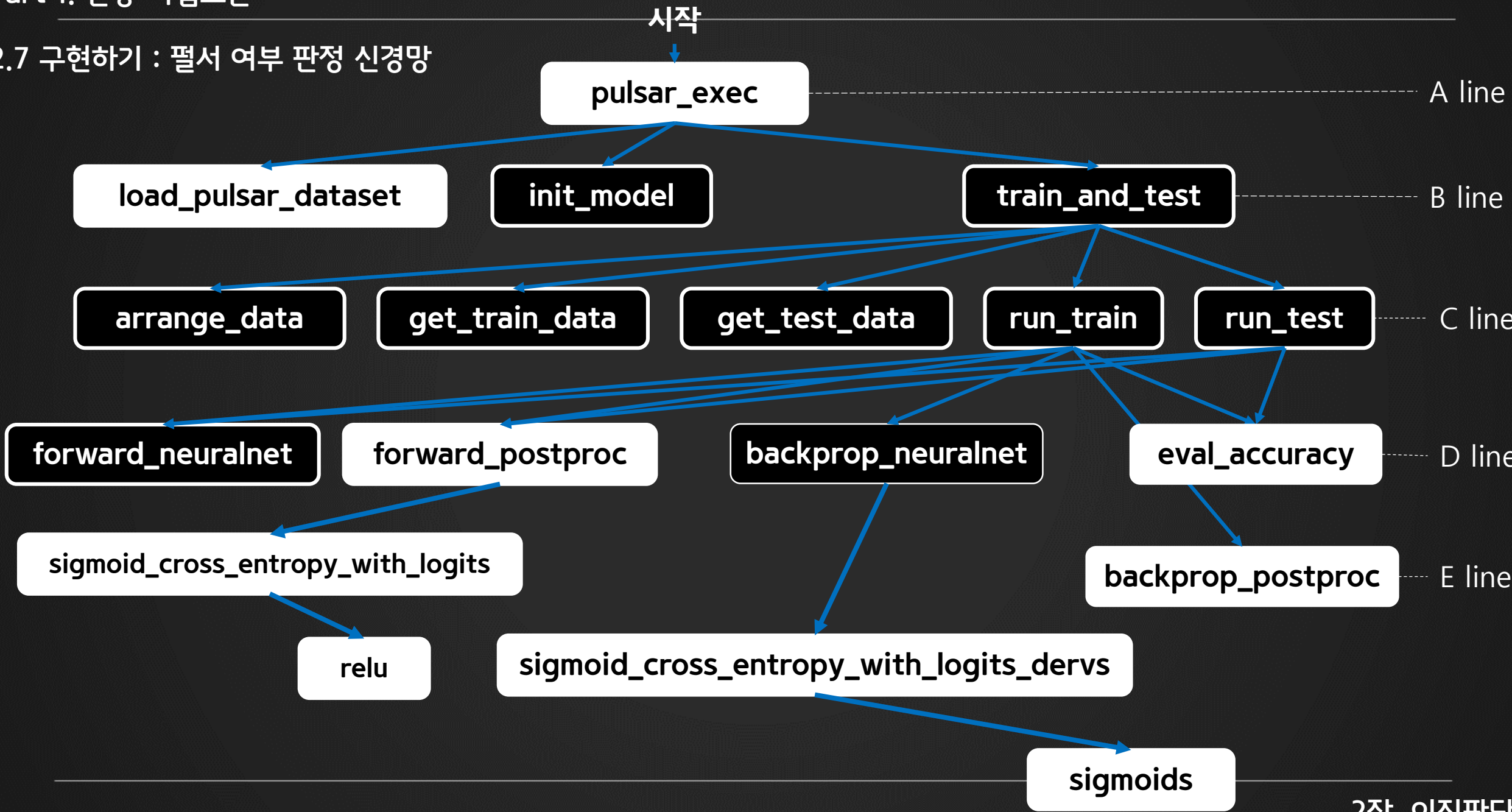
$$\frac{\partial H}{\partial x} = -z + \sigma(x)$$

$$H = \max(x, 0) - xz + \log(1 + e^{-|x|})$$


$$H = x - xz + \log(1 + e^{-x})$$



2.7 구현하기 : 펄서 여부 판정 신경망



2.8 확장하기 : 균형잡힌 데이터셋과 착시 없는 평가방법



```
pulsar_exec()

Epoch 1: loss=0.138, accuracy=0.960/0.970
Epoch 2: loss=0.134, accuracy=0.966/0.976
Epoch 3: loss=0.134, accuracy=0.967/0.966
Epoch 4: loss=0.140, accuracy=0.967/0.975
Epoch 5: loss=0.129, accuracy=0.966/0.977
Epoch 6: loss=0.118, accuracy=0.968/0.967
Epoch 7: loss=0.127, accuracy=0.968/0.973
Epoch 8: loss=0.126, accuracy=0.969/0.976
Epoch 9: loss=0.124, accuracy=0.969/0.977
Epoch 10: loss=0.131, accuracy=0.968/0.975

Final Test: final accuracy = 0.975
```



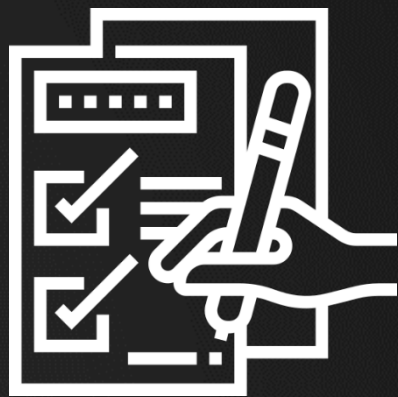
정확도 90%의 신경망 모델



새로운 신경망 평가지표의 필요성

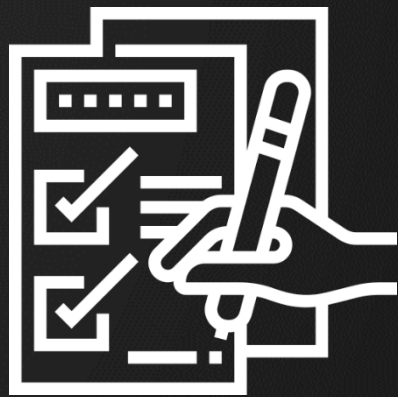
*Pulsar = 10%*

*Star = 90%*



*Pulsar = 50%*

*Star = 50%*

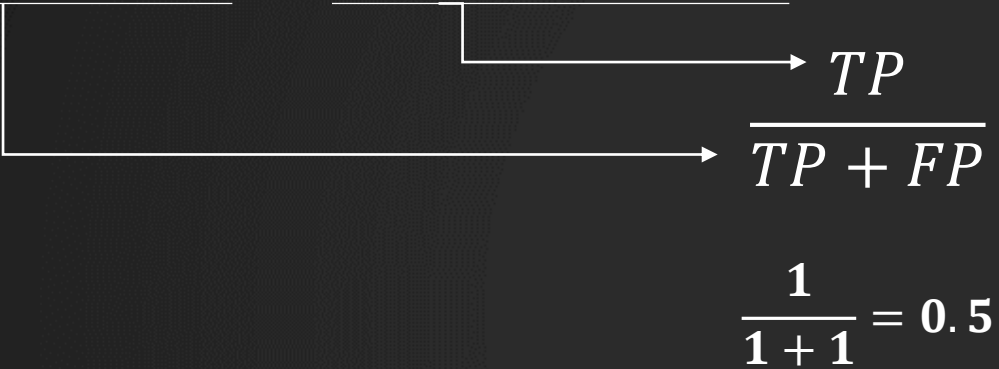




2.8 확장하기 : 균형잡힌 데이터셋과 착시 없는 평가방법

정밀도

“분류 결과 양성으로 식별된 사례 중 실제로 양성이었던 사례의 비율은 어느 정도인가요?”

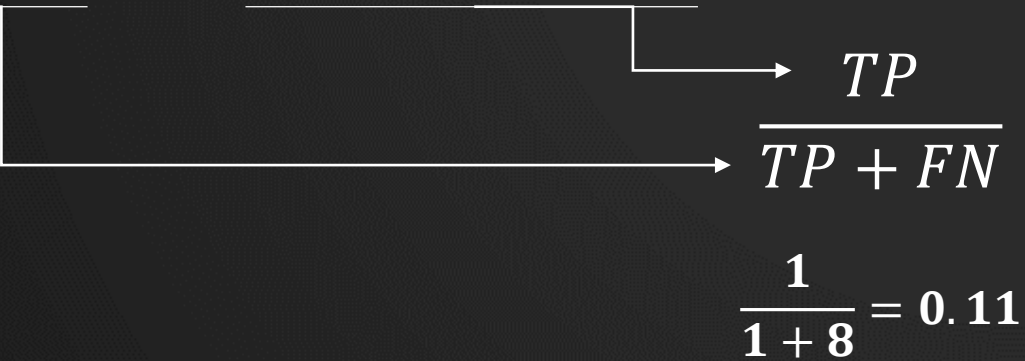


정답 레이블  
( T / F )

	TP : 1	FP : 1
정답과 예측 (N/P)	FN : 8	TN : 90

재현율

“실제 양성 중 정확히 양성이라고 식별된 사례의 비율은 어느 정도 인가요?”



$F1 \rightarrow \frac{2}{\frac{1}{\text{정밀도}} + \frac{1}{\text{재현율}}}$

$= 2 * \frac{\text{정밀도} * \text{재현율}}{\text{정밀도} + \text{재현율}}$