

# +·\*~ 2021 SMARCLE 겨울방학 인공지능 스터디 ~·\*

넷째마당 - 10장~11장-3

3주차 1팀 2021.1.19.

김건우 박지하 송혜원





# 목차

## 1. 딥러닝 활용하기

- (1) 폐암 수술 환자의 생존율 예측
- (2) 교차 엔트로피



## 2. 딥러닝과 데이터

- (1) 피마 인디언 데이터 분석하기
- (2) Pandas를 활용한 데이터 조사



# 폐암 수술 환자의 생존율 예측하기

표 2-1 폐암 수술 환자의 의료 기록 데이터

줄 항목	속성																	클래스
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	293	1	3.8	2.8	0	0	0	0	0	0	12	0	0	0	1	0	62	0
2	1	2	2.88	2.16	1	0	0	0	1	1	14	0	0	0	1	0	60	0
3	8	2	3.19	2.5	1	0	0	0	1	0	11	0	0	1	1	0	66	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
470	447	8	5.2	4.1	0	0	0	0	0	0	12	0	0	0	0	0	49	0

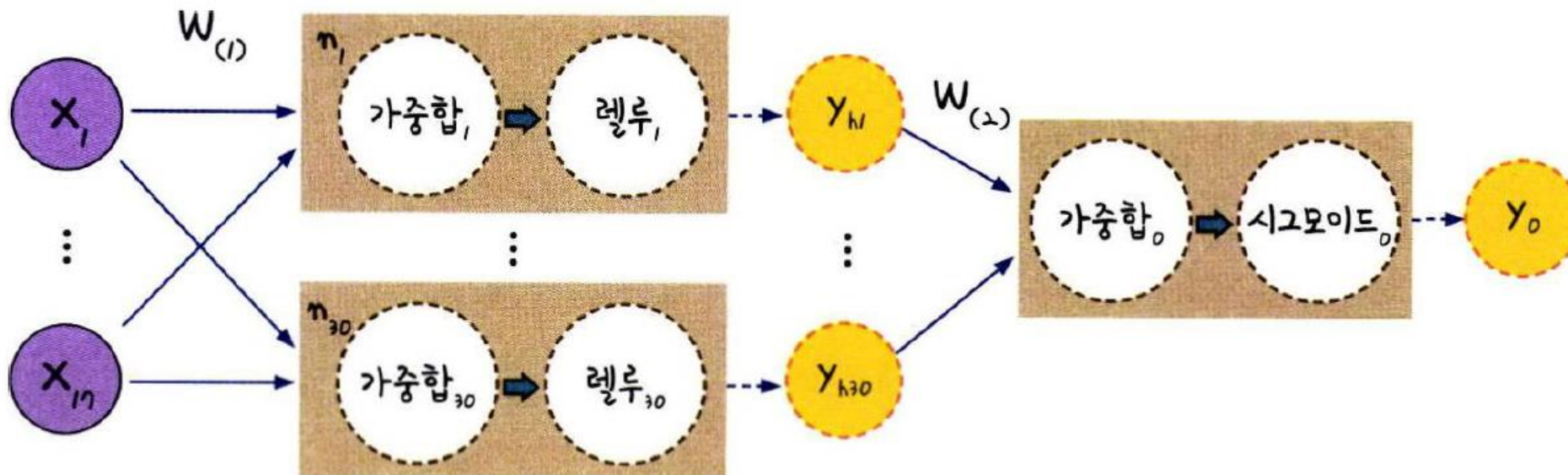
# ~실습 타임~

\* Google colab을 실행해주세요.



```
model.compile(loss='mean_squared_error', optimizer='adam', metrics=['accuracy'])
```

---



\* 실제 값을  $y_t$ , 예측 값을  $y_o$ 라고 가정할 때

평균 제곱 계열	mean_squared_error	평균 제곱 오차 계산: $\text{mean}(\text{square}(y_t - y_o))$
	mean_absolute_error	평균 절대 오차(실제 값과 예측 값 차이의 절댓값 평균) 계산: $\text{mean}(\text{abs}(y_t - y_o))$
	mean_absolute_percentage_error	평균 절대 백분율 오차(절댓값 오차를 절댓값으로 나눈 후 평균) 계산: $\text{mean}(\text{abs}(y_t - y_o) / \text{abs}(y_t))$ (단, 분모 $\neq 0$ )
	mean_squared_logarithmic_error	평균 제곱 로그 오차(실제 값과 예측 값에 로그를 적용한 값의 차이를 제곱한 값의 평균) 계산: $\text{mean}(\text{square}((\log(y_o) + 1) - (\log(y_t) + 1)))$
교차 엔트로피 계열	categorical_crossentropy	범주형 교차 엔트로피(일반적인 분류)
	binary_crossentropy	이항 교차 엔트로피(두 개의 클래스 중에서 예측할 때)



## 교차 엔트로피

`model.fit(X, Y, epochs=100, batch_size=10)`

속성

클래스

샘플

	정보 1	정보 2	정보 3	...	정보 17	생존 여부
1번째 환자	293	1	3.8	...	62	0
2번째 환자	1	2	2.88	...	60	0
3번째 환자	8	3	3.19	...	66	1
...	...	...	...	...	...	...
470번째 환자	447	8	5.2	...	49	0



## 데이터 나누기 - 좋은데이터와 나쁜데이터?

### 〈나쁜데이터〉

1. 충분하지 않은 양의 데이터
2. 대표성 없는 훈련 데이터
3. 낮은 품질의 데이터
4. 관련 없는 특성



### 〈좋은데이터〉

1. 충분한 양의 데이터
2. 대표성 있는 훈련 데이터
3. 좋은 품질의 데이터
4. 관련 있는 특성



# Pandas를 활용한 피마 인디언 데이터 분석하기



```
import pandas as pd
```

```
.
6,148,72,35,0,33.6,0.627,50,1
1,85,66,29,0,26.6,0.351,31,0
8,183,64,0,0,23.3,0.672,32,1
1,89,66,23,94,28.1,0.167,21,0
0,137,40,35,168,43.1,2.288,33,1
5,116,74,0,0,25.6,0.201,30,0
3,78,50,32,88,31.0,0.248,26,1
10,115,0,0,0,35.3,0.134,29,0
2,197,70,45,543,30.5,0.158,53,1
8,125,96,0,0,0,0.232,54,1
4,110,92,0,0,37.6,0.191,30,0
10,168,74,0,0,38.0,0.537,34,1
10,139,80,0,0,27.1,1.441,57,0
1,189,60,23,846,30.1,0.398,59,1
5,166,72,19,175,25.8,0.587,51,1
7,100,0,0,0,30.0,0.484,32,1
0,118,84,47,230,45.8,0.551,31,1
7,107,74,0,0,29.6,0.254,31,1
1,103,30,38,83,43.3,0.183,33,0
1,115,70,30,96,34.6,0.529,32,1
3,126,88,41,235,39.3,0.704,27,0
8,99,84,0,0,35.4,0.388,50,0
7,196,90,0,0,39.8,0.451,41,1
2,118,88,27,0,32.0,0.484,32,1
```

# ~실습 타임~

\* Google colab을 실행해주세요.



# ~질문 타임~

\*질문 환영



**1팀이었습시다.**  
**잘 들어주셔서 감사합니당!!**

