

# 선수 지식 - 통계

## 조건부 확률

조건부 확률 | 딥러닝의 기초가 되는 확률 개념 알아보기

강사 나동빈

# 선수 지식 - 통계

조건부 확률

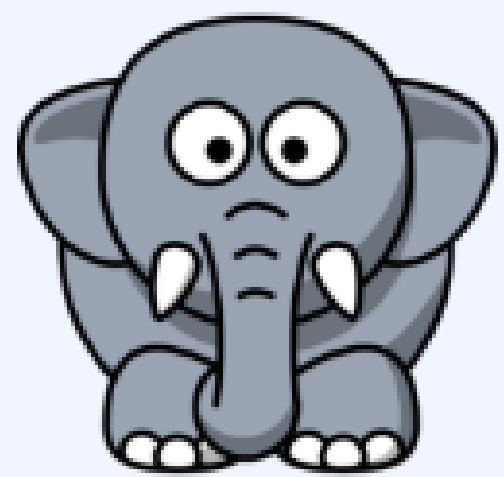
## 조건부 확률 (Conditional Probability)

- 어떠한 사건이 일어나는 경우에 다른 사건이 일어날 확률을 의미한다.
- 딥러닝 분야에서는 “ $X$  사건이 단서로 주어졌을 때,  $Y$  사건이 발생할 확률”로 이해할 수 있다.

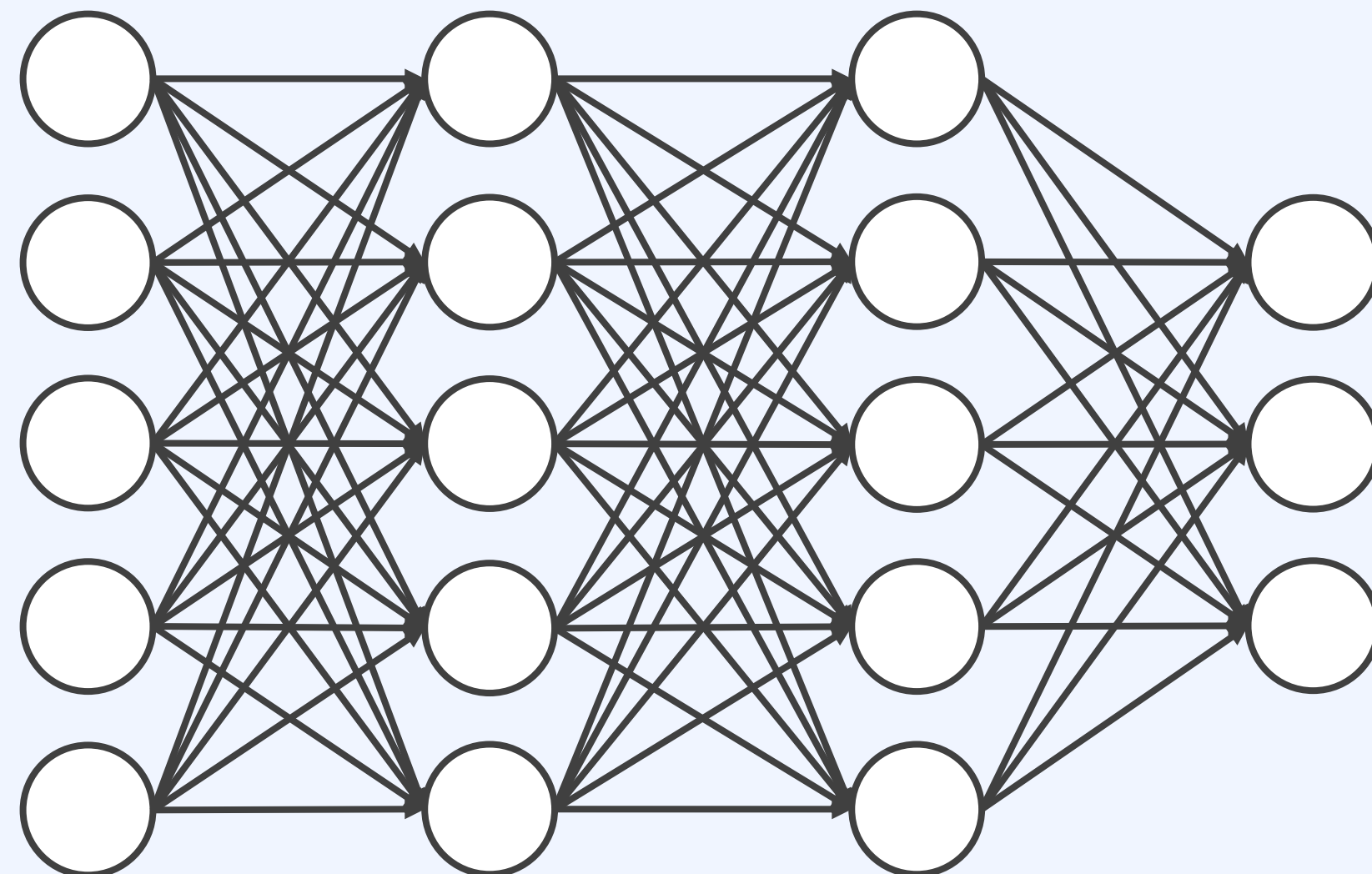
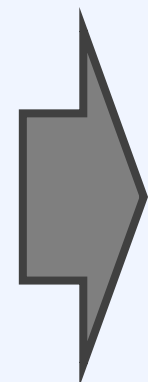
## 조건부 확률 (Conditional Probability)

- 분류 모델은 일반적으로 다음과 같이 동작한다.
- $x$ 라는 이미지가 입력으로 주어졌을 때, 클래스(class)  $y$ 가 나올 확률은 다음과 같이 표현한다.

$$P(Y = y | X = x)$$



이미지  $x$



$$P(Y = \text{새} | X = x) = 0.06$$

$$P(Y = \text{코끼리} | X = x) = 0.87$$

$$P(Y = \text{벌} | X = x) = 0.07$$

## 조건부 확률 (Conditional Probability)

- 특정한 사건이 발생하는 경우에 다른 사건이 발생할 확률을 의미한다.

	스팸 메일	일반 메일	합계
학교 계정	40	30	70
회사 계정	50	60	110
합계	90	90	180

- 하나의 메일을 뽑았을 때, 학교 계정으로 온 메일일 확률  
→  $P(\text{학교}) = 70/180$
- 하나의 메일을 뽑았을 때, 학교 계정으로 온 메일이면서 스팸 메일일 확률 (**결합 확률**)  
→  $P(\text{학교} \cap \text{스팸}) = 40/180$
- 스팸 메일 중 하나를 뽑았을 때, 학교 계정으로 온 메일을 확률 (**조건부 확률**)  
→  $P(\text{학교}|\text{스팸}) = 40/90$

## 조건부 확률 질량 함수

- 조건부 확률 질량 함수 공식:  $P_{Y|X}(y|x) = \frac{P_{XY}(x,y)}{P_X(x)}$
- 영어 성적(Y)이 1등급일 때, 수학 성적(X)이 1등급일 확률은?

Y \ X	1	2	3	4	5	$P_Y(y)$
1	2 / 30	1 / 30	0 / 30	0 / 30	0 / 30	3 / 30
2	1 / 30	3 / 30	4 / 30	0 / 30	0 / 30	8 / 30
3	1 / 30	3 / 30	5 / 30	2 / 30	0 / 30	11 / 30
4	0 / 30	0 / 30	0 / 30	3 / 30	2 / 30	5 / 30
5	0 / 30	0 / 30	0 / 30	1 / 30	2 / 30	3 / 30
$P_X(x)$	4 / 30	7 / 30	9 / 30	6 / 30	4 / 30	

➡ 2 / 3

## 조건부 확률 질량 함수

- Python을 이용해 특정한 조건부 확률 질량 함수를 표현할 수 있다.

```
index = 0
x = [0, 1, 2, 3, 4]
plt.bar(x, pmf.iloc[index] / marginal_pmf_y[index + 1])
plt.xticks(x, ["1", "2", "3", "4", "5"])
plt.title(f"P(X|Y={index + 1})")
plt.show()
```

- Python을 이용해 특정한 조건부 확률 질량 함수를 표현할 수 있다.

