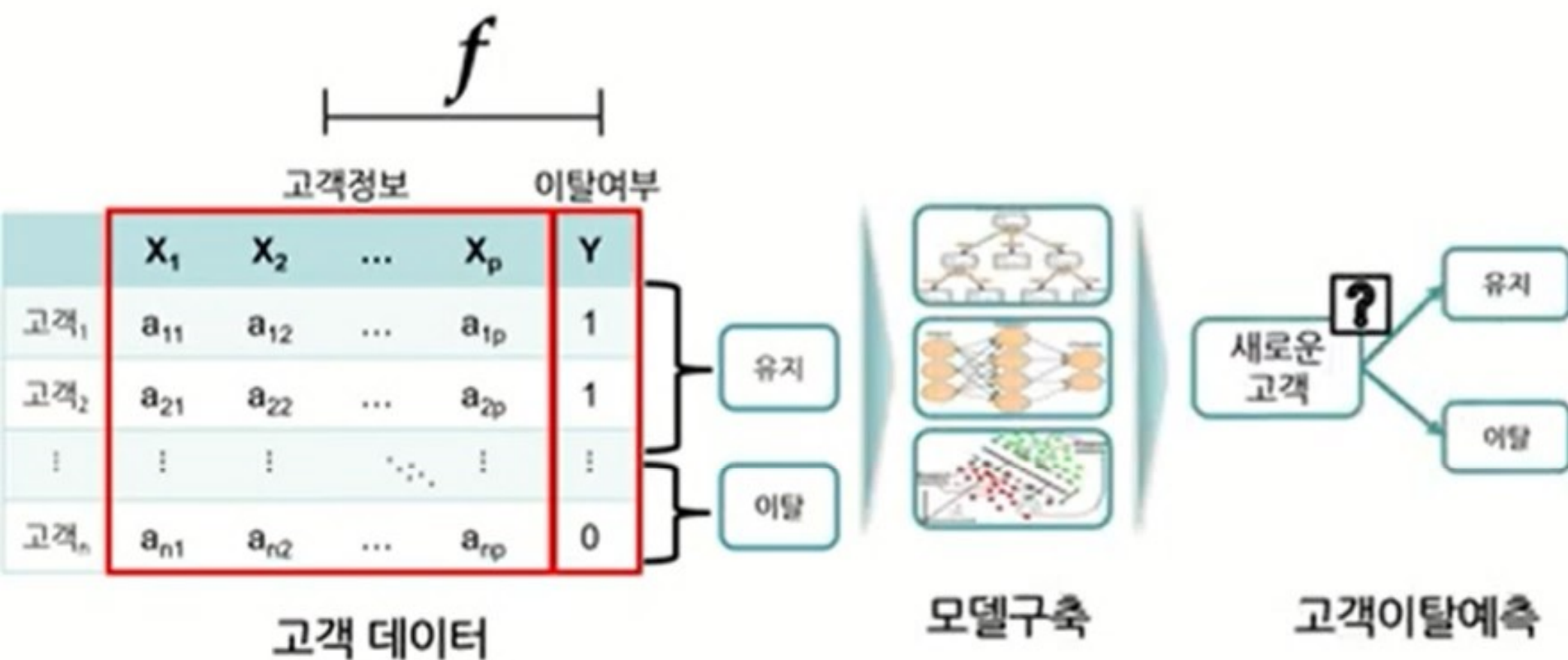


점주예측 예제 - 고객 이탈 예측

- 고객의 정보(성별, 연령, 직업, 연봉 등)를 이용하여,
고객 이탈 여부를 예측



다변량) 데이터

인자 (변수) 예측치	X_1	...	X_i	...	X_p	Y
N_1	x_{11}	...	x_{1i}	...	x_{1p}	20.5
N_2	x_{21}	...	x_{2i}	...	x_{2p}	22.2
...
N_{n-1}	x_{n-11}	...	x_{n-1i}	...	x_{n-1p}	72.3
N_n	x_{n1}	...	x_{ni}	...	x_{np}	82.8

X (원인): 독립변수, 예측변수, 입력변수

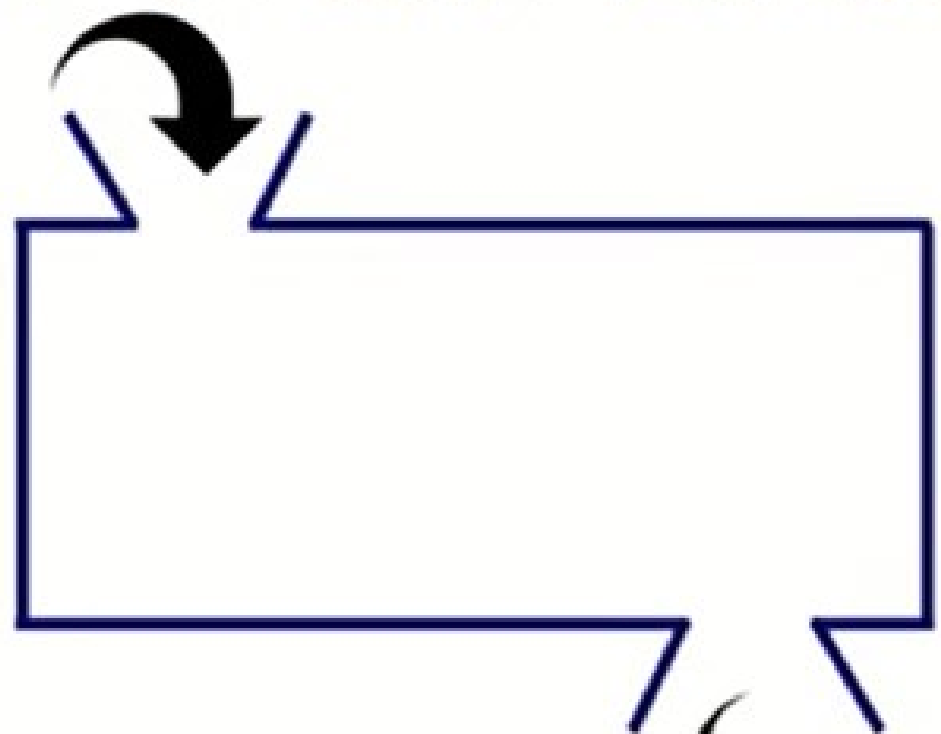
Y (결과): 종속변수, 반응변수, 출력변수

X_1	X_2	Y
0	2	2
1	3	4
2	4	6
3	5	8



X1	X2	Y
0	2	2
1	3	4
2	4	6
3	5	8

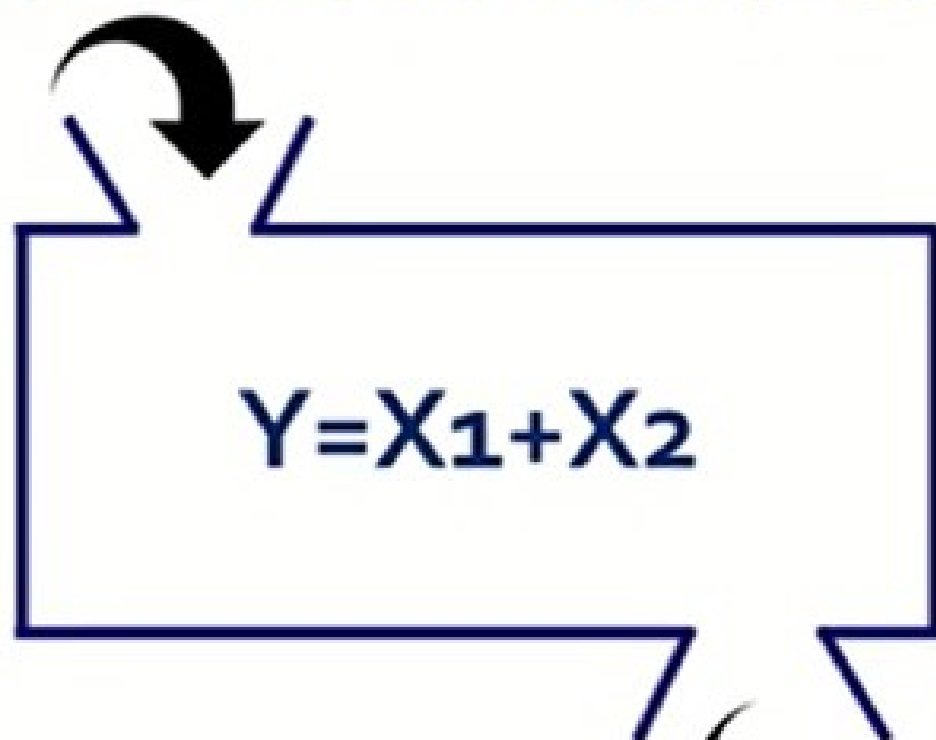
$(0, 2), (1, 3), (2, 4), (3, 5)$



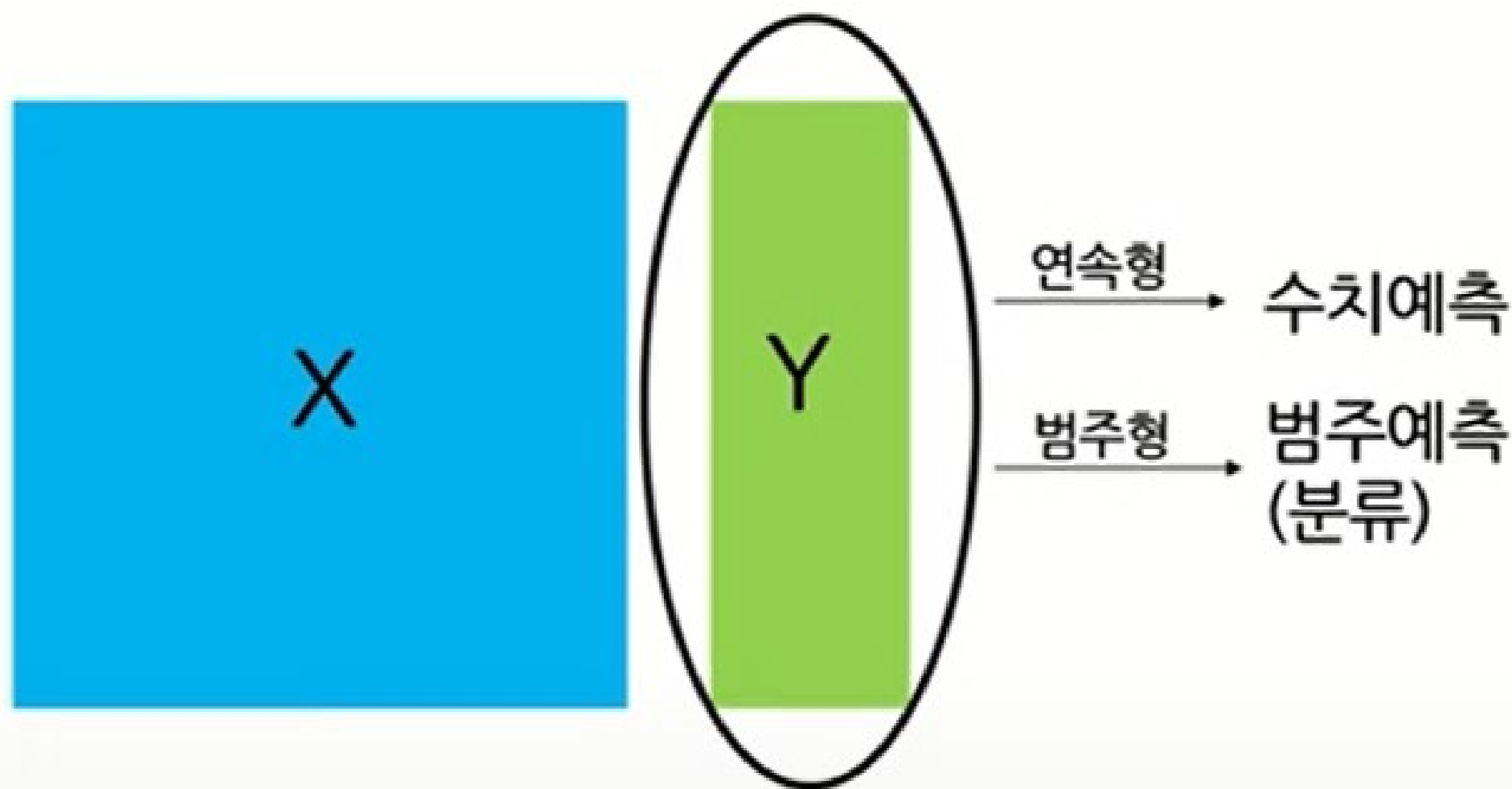
$2, 4, 6, 8$

X_1	X_2	Y
0	2	2
1	3	4
2	4	6
3	5	8

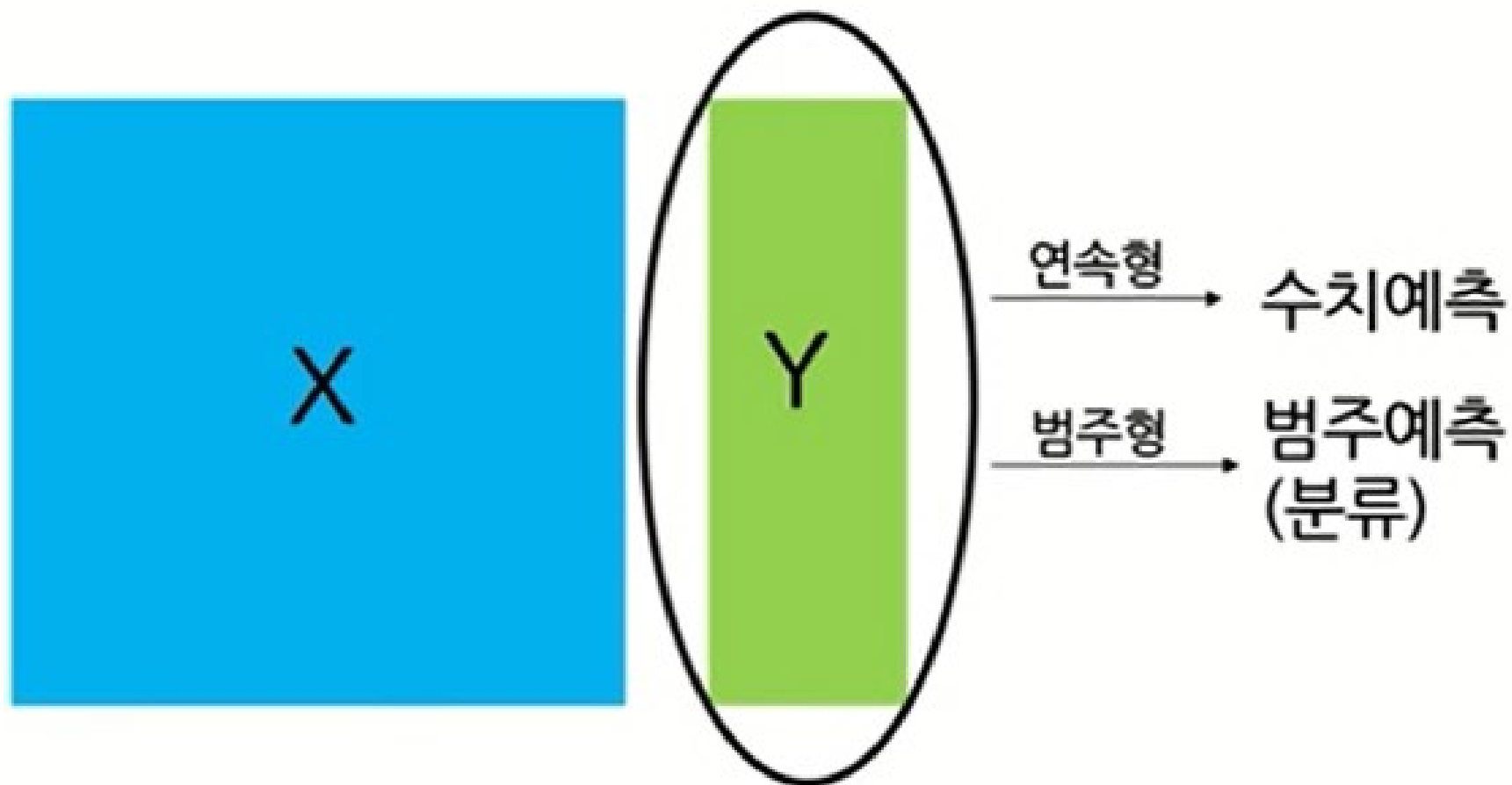
$(0, 2), (1, 3), (2, 4), (3, 5)$



$2, 4, 6, 8$

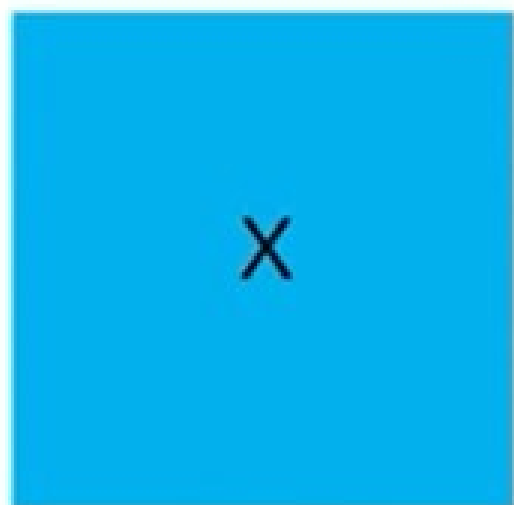


수치예측 / 범주예측 (분류)



- 연속형 데이터: 데이터 자체를 숫자로 표현
예) 가격, 길이, 압력, 두께, ...
- 범주형 데이터: 원칙적으로 숫자로 표시할 수 없는 데이터
예) 제품불량여부 (양품/불량), 보험사기여부(정상/비정상), ...

수치예측 데이터

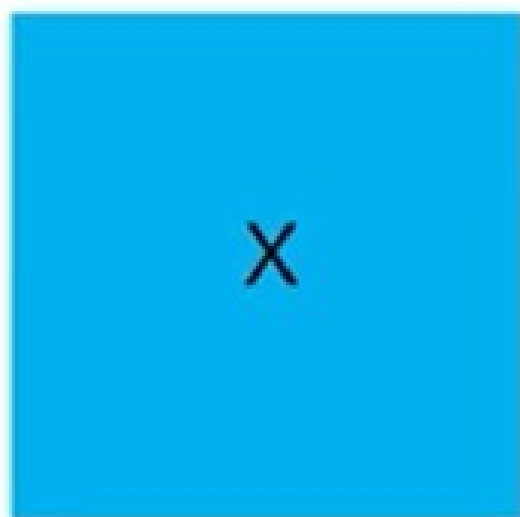


연속형 → 수치예측 (Regression)

인자 (변수) 관측치	X_1	...	X_i	...	X_p
N_1	x_{11}	...	x_{1i}	...	x_{1p}
N_2	x_{21}	...	x_{2i}	...	x_{2p}
...
N_{n-1}	x_{n-11}	...	x_{n-1i}	...	x_{n-1p}
N_n	x_{n1}	...	x_{ni}	...	x_{np}

Y
20.5
22.2
...
72.3
82.8

범주 예측 데이터

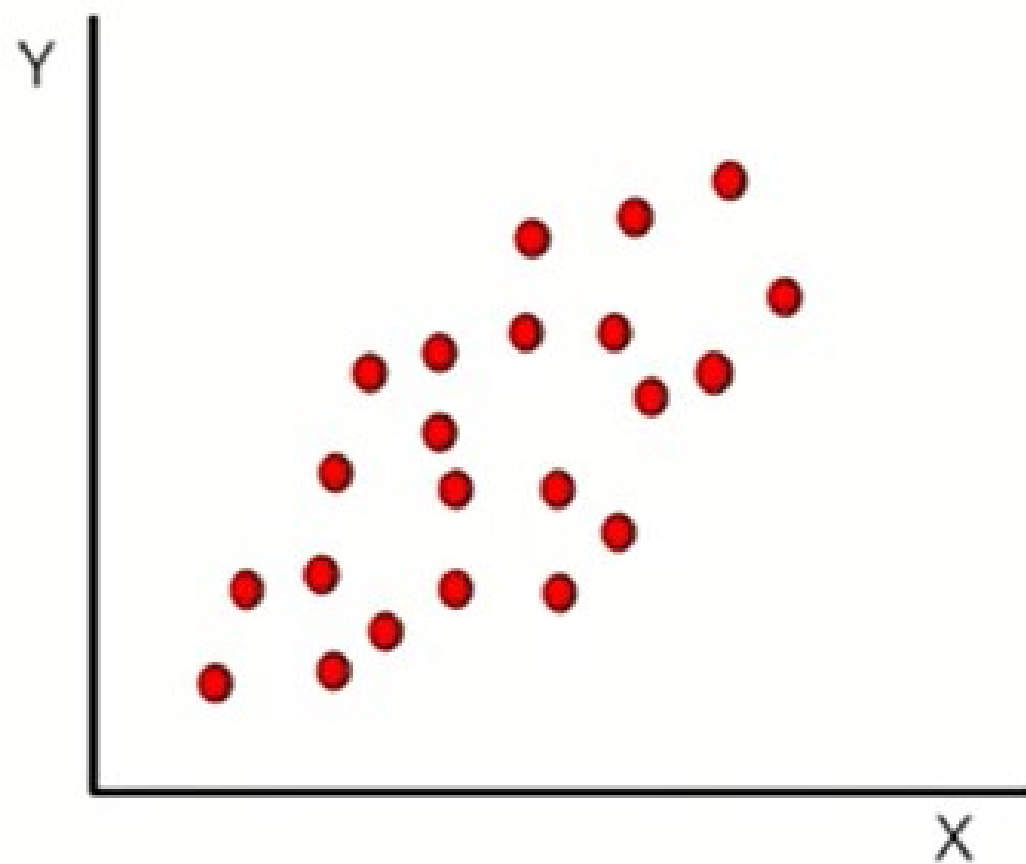


범주형

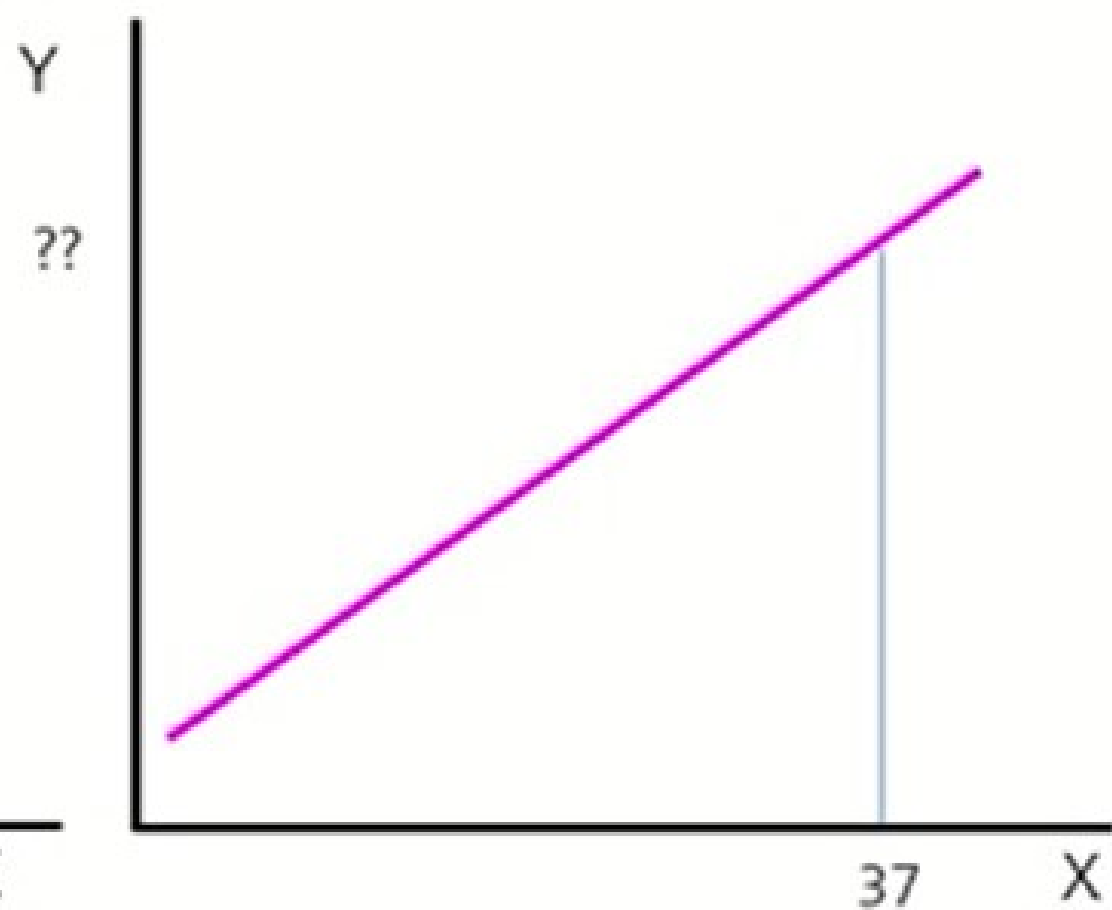
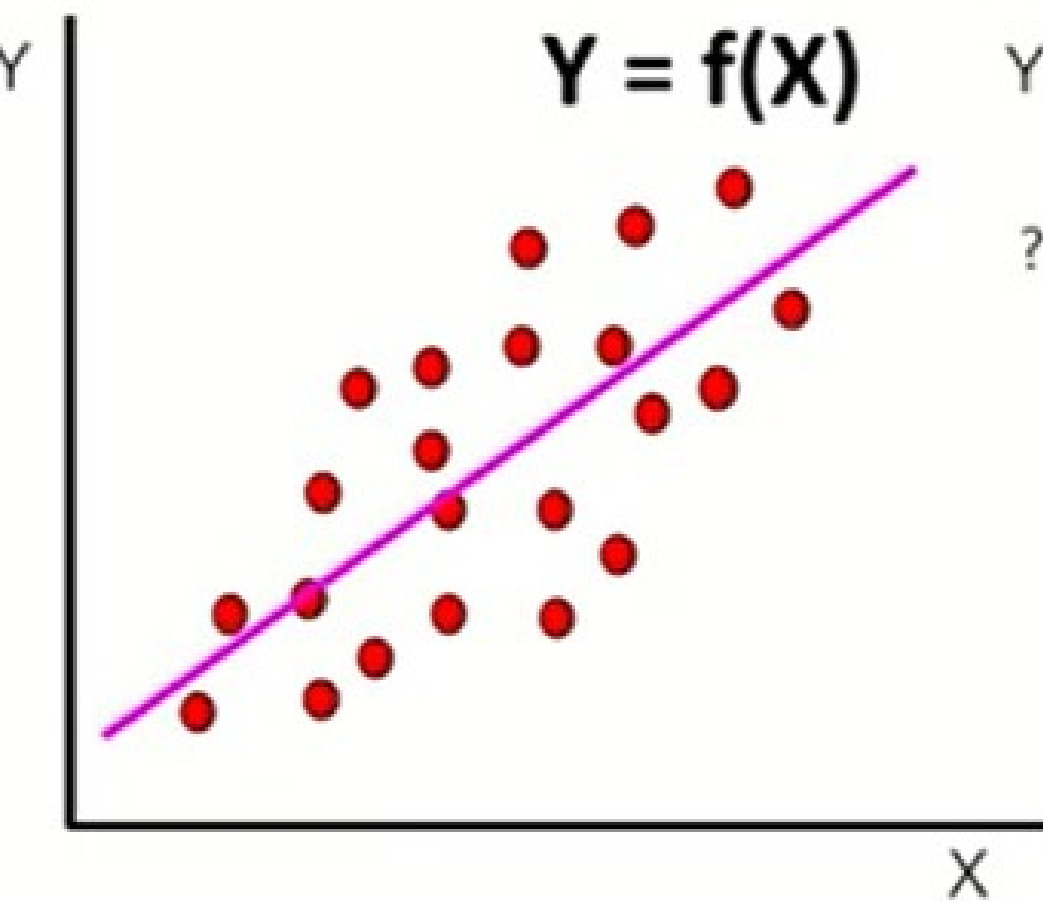
→ 범주 예측, 분류 (Classification)

인자 (변수)	X_1	...	X_i	...	X_p	Y
관측치						
N_1	x_{11}	...	x_{1i}	...	x_{1p}	0 (정상)
N_2	x_{21}	...	x_{2i}	...	x_{2p}	0 (정상)
...
N_{n-1}	x_{n-11}	...	x_{n-1i}	...	x_{n-1p}	1 (불량)
N_n	x_{n1}	...	x_{ni}	...	x_{np}	1 (불량)

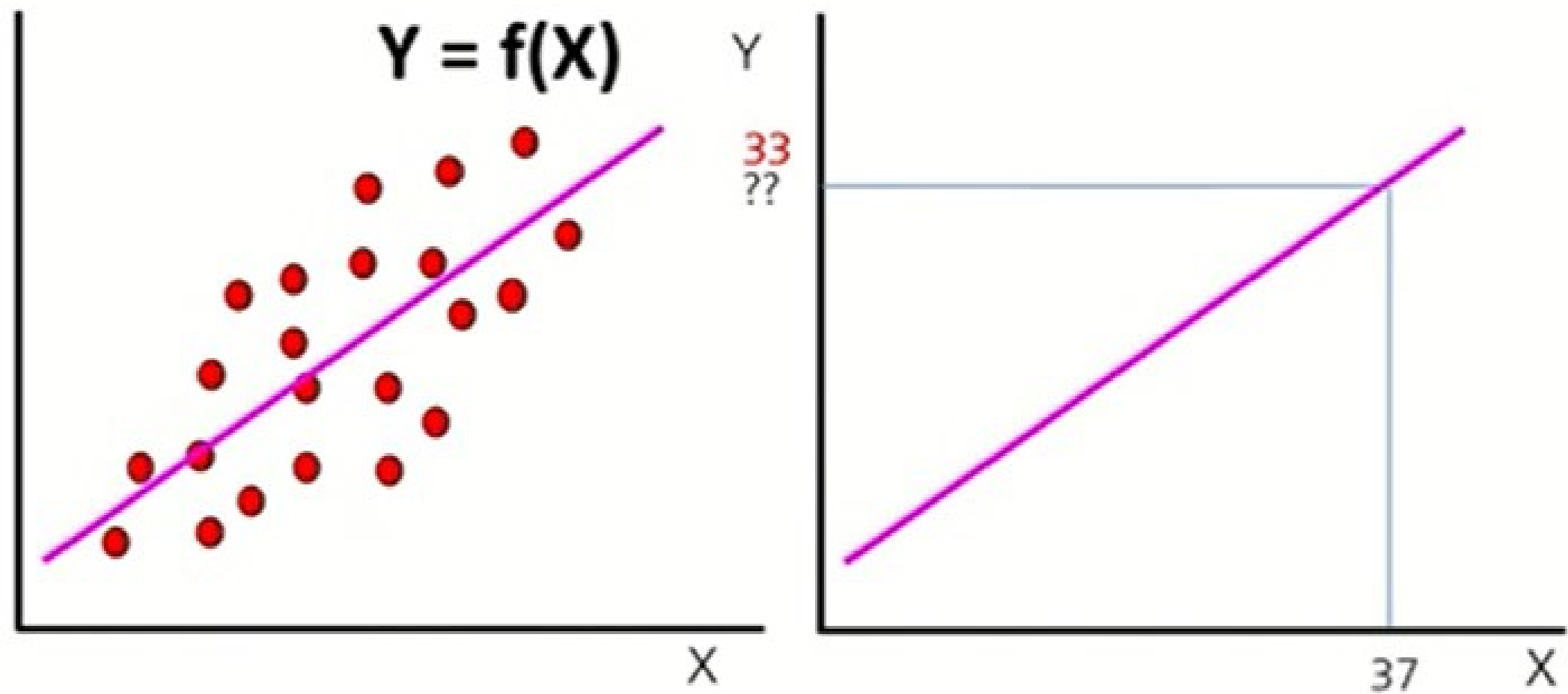
수치예측 모델링 개요



회귀분석 모형 개요



치예측 모델링 개요

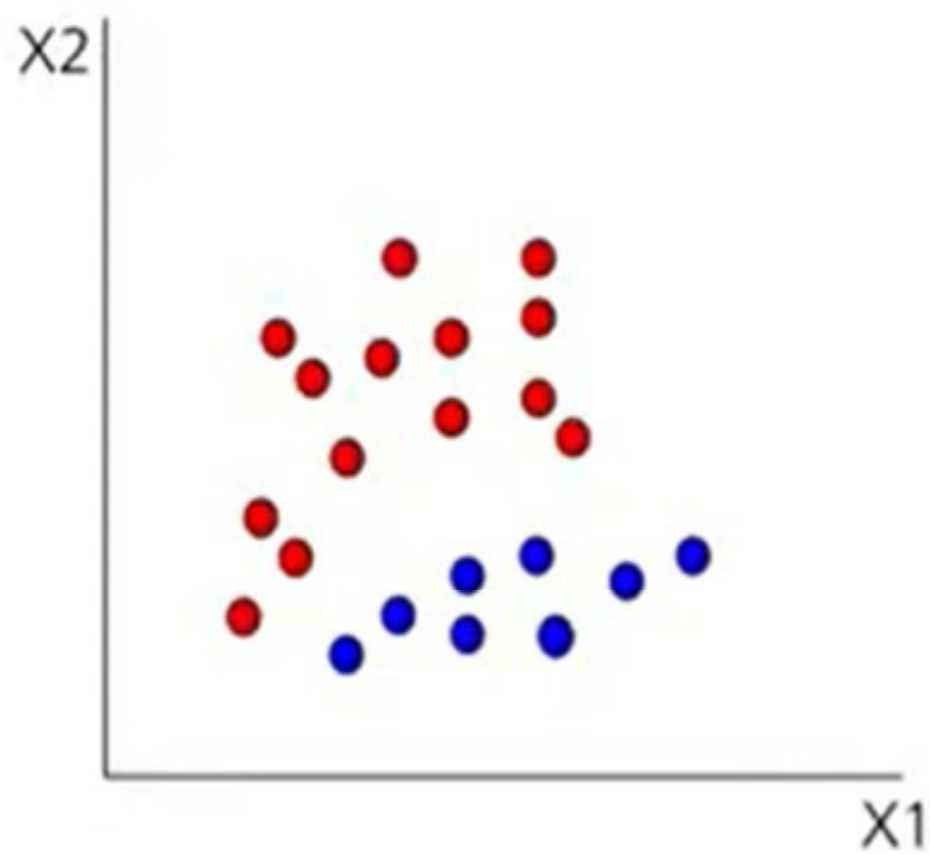


수치예측 예제 - 중고차 가격 예측

모델	X			Y
	주행거리	마력	용량 (CC)	가격
TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors	46986	90	2000	13500
TOYOTA Corolla 1800 T SPORT VVT I 2/3-Doors	19700	192	1800	21500
TOYOTA Corolla 1.9 D HATCHB TERRA 2/3-Doors	71138	69	1900	12950
TOYOTA Corolla 1.8 VVT-i T-Sport 3-Dr 2/3-Doors	31461	192	1800	20950
TOYOTA Corolla 1.8 16V VVT-i 3DR T SPORT BNS 2/3-Doors	43610	192	1800	19950
TOYOTA Corolla 1.6 VVTi Linea Terra Comfort 2/3-Doors	21716	110	1600	17950
TOYOTA Corolla 1.6 16v LSOL 2/3-Doors	25563	110	1600	16750
TOYOTA Corolla 1.6 16V VVT I 3DR TERRA 2/3-Doors	64359	110	1600	16950
TOYOTA Corolla 1.6 16V VVT I 3DR SOL AUT4 2/3-Doors	43905	110	1600	16950
TOYOTA Corolla 1.6 16V VVT I 3DR SOL 2/3-Doors	56349	110	1600	15950
TOYOTA Corolla 1.4 VVTi Linea Terra 2/3-Doors	9750	97	1400	12950
TOYOTA Corolla 1.4 16V VVT I 3DR 2/3-Doors	27500	97	1400	14750
TOYOTA Corolla 1.4 16V VVT I 3DR 2/3-Doors	49059	97	1400	13950
TOYOTA Corolla 1.4 16V VVT I 3DR 2/3-Doors	44068	97	1400	16750
TOYOTA Corolla 1.4 16V VVT I 3DR 2/3-Doors	46961	97	1400	13950
TOYOTA Corolla 2.0 D4D 90 SDR TERRA COMFORT 4/5-Doors	110404	90	2000	16950
TOYOTA Corolla 2.0 D4D 90 SDR TERRA COMFORT 4/5-Doors	100250	90	2000	16950
TOYOTA Corolla 2.0 D4D 90 SDR SOL 4/5-Doors	84000	90	2000	19000
TOYOTA Corolla 2.0 D4D 90 SDR TERRA 4/5-Doors	79375	90	2000	17950
TOYOTA Corolla 1.4 16V VVT I SDR TERRA COMFORT 4/5-Doors	75048	97	1400	15800

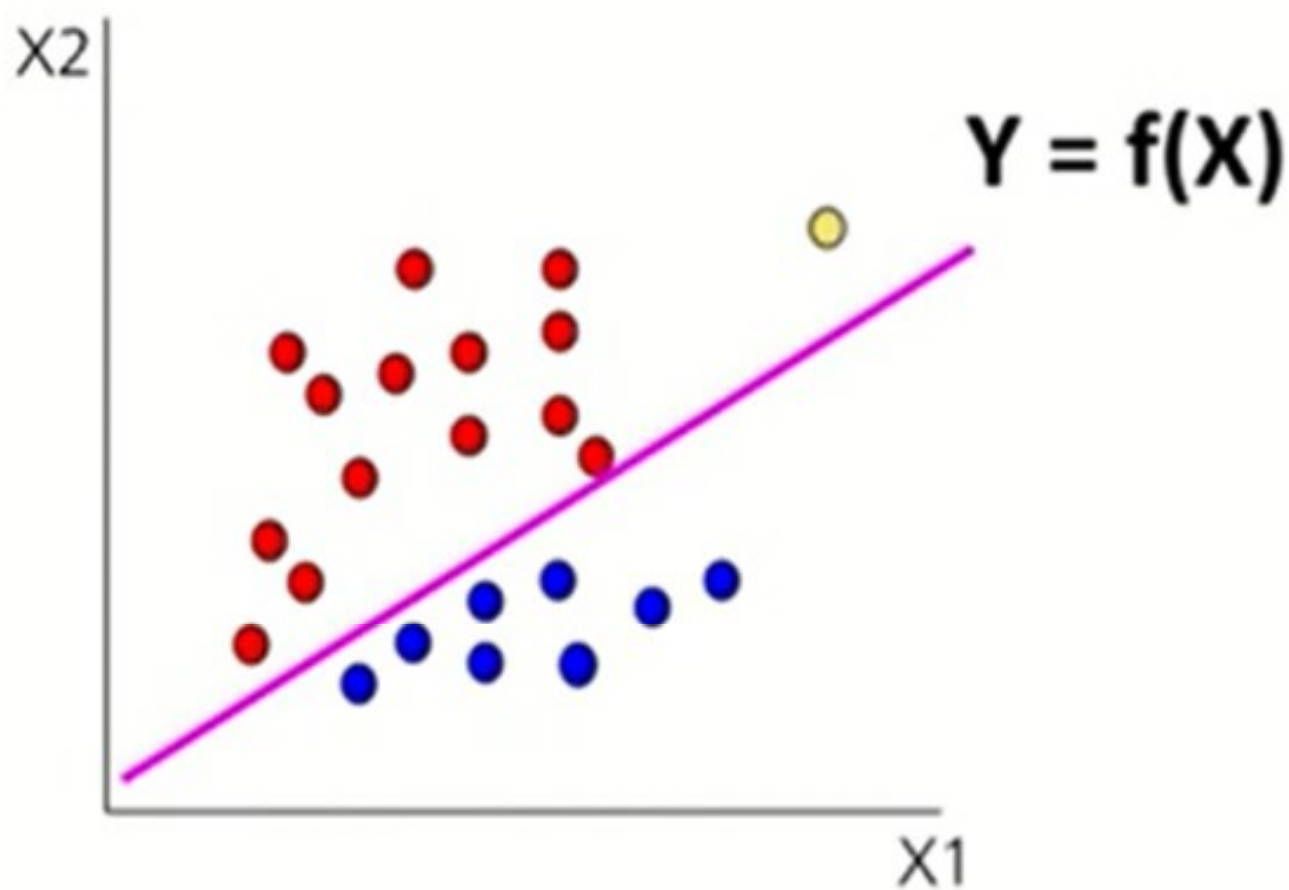
범주예측 모델링 개요

- 불량범주
- 양품범주



주예측 모델링 개요

- 불량범주
- 양품범주



범주예측 예제 - 불량 예측

- 디스플레이 공정에서 공정 파라미터의 측정값들을 이용하여, 해당 글라스가 양품인지 불량품인지를 여부를 예측

공정파라미터

	X_1	X_2	...	X_p	Y
제품 ₁	a_{11}	a_{12}	...	a_{1p}	1
제품 ₂	a_{21}	a_{22}	...	a_{2p}	1
⋮	⋮	⋮		⋮	⋮
제품 _n	a_{n1}	a_{n2}	...	a_{np}	0

양품

불량품

디스플레이공정 데이터

주예측 예제 - 불량 예측

디스플레이 공정에서 공정 파라미터의 측정값들을 이용하여,
해당 글라스 가 양품인지 불량품인지를 여부를 예측

	공정파라미터				글라스 상태	
	X_1	X_2	...	X_p	Y	
제품 ₁	a_{11}	a_{12}	...	a_{1p}	1	양품
제품 ₂	a_{21}	a_{22}	...	a_{2p}	1	
...	
제품 _n	a_{n1}	a_{n2}	...	a_{np}	0	불량품

디스플레이공정 데이터

주예측 예제 - 불량 예측

디스플레이 공정에서 공정 파라미터의 측정값들을 이용하여,
해당 글라스 가 양품인지 불량품인지의 여부를 예측

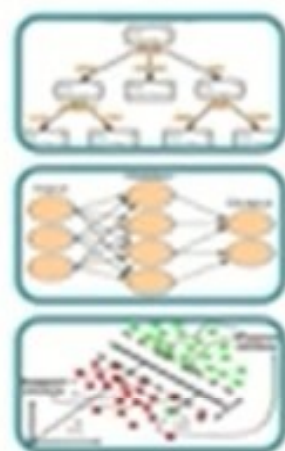
$$f$$

공정파라미터 글라스 상태

	X_1	X_2	...	X_p	Y
제품 ₁	a_{11}	a_{12}	...	a_{1p}	1
제품 ₂	a_{21}	a_{22}	...	a_{2p}	1
⋮	⋮	⋮	⋱	⋮	⋮
제품 _n	a_{n1}	a_{n2}	...	a_{np}	0

양품

불량품



새로운
제품



양품

불량품

디스플레이공정 데이터

모델구축

불량 글라스 예