

# 데이터 사이언스와 인공지능

제주대학교 변영철

나, 연서









양, 들끼
























나만의 메리골드 

날짜	길이 (전체)	너비	길이 (앞다리)
2019.06.16	0	0	0
2019.06.20	1.6cm	0.28cm	0.9cm
06.22	2.6cm	3.8cm	1.4cm
6.25	<del>3.0</del> 3cm	4.6cm	1.6cm
2019.6.29	4.4	<del>5.8</del> 5.8cm	2cm
	<del>5.0</del>	0.58cm	
	5.0	0.6	2.2

② 돌고래

① 큰따개비



2019.06.16	0	0	0
2019.06.20	1.6cm	0.28cm	0.9cm
2019.06.22	3.3cm	4.5cm	1.5cm
6.25	3.5cm	5cm	1.8cm
6.29	5cm	0.65cm	2.5cm

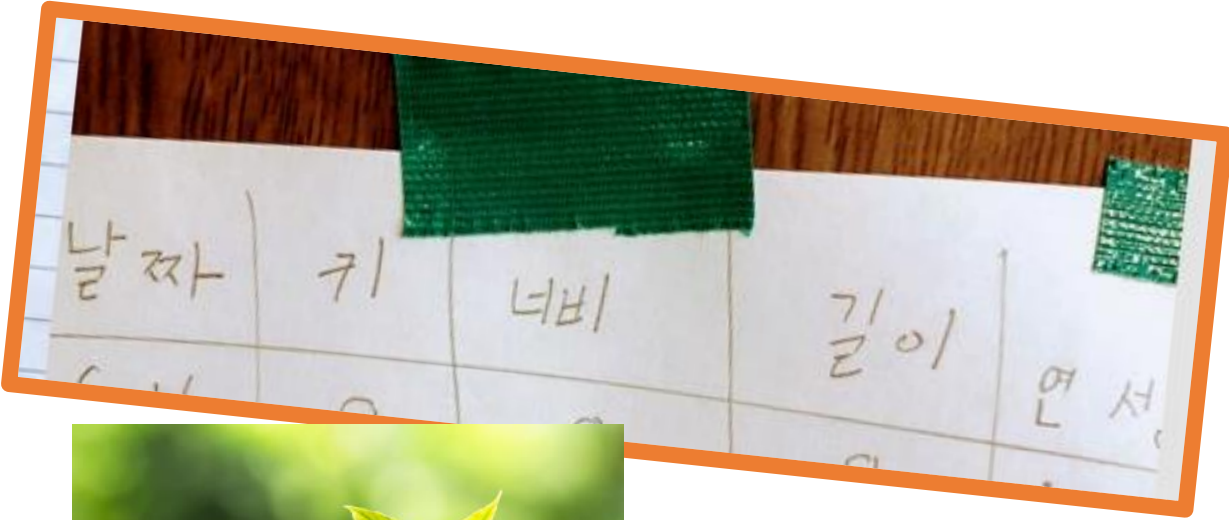
날짜	길이	너비	길이
	0	0	0
	12mm	2.9mm	8mm
6.22	25mm	4.5mm	13mm
25	31mm	5.5mm	15mm
6.29	45.5mm	60mm	19mm

③ 세이재



지난날짜, 키, 잎 너비, 잎 길이, 주인

1, 0, 0, 0, 1  
1, 0, 0, 0, 2  
1, 0, 0, 0, 3  
5, 16, 28, 9, 1  
5, 16, 2.8, 9, 2  
5, 12, 2.9, 8, 3  
7, 33, 4.5, 15, 1  
7, 26, 3.8, 14, 2  
7, 25, 4.5, 13, 3  
10, 35, 5, 18, 1  
10, 30, 4.6, 16, 2  
10, 31, 5.5, 15, 3  
14, 50, 6.5, 25, 1  
14, 44, 5.8, 20, 2  
14, 45.5, 6, 19, 3  
20, 56, 6.8, 27, 1  
20, 50, 6, 22, 2  
20, 51, 6.5, 21, 3





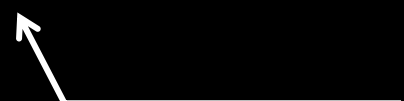
'날짜'가 지남에 따라  
'잎 너비'는 얼마나 자랐을까?  
점으로 찍어보라(plot!)  
(주인에 따라 다른 색으로 표시)

**plot**(df, '날짜', '잎 너비', '주인')

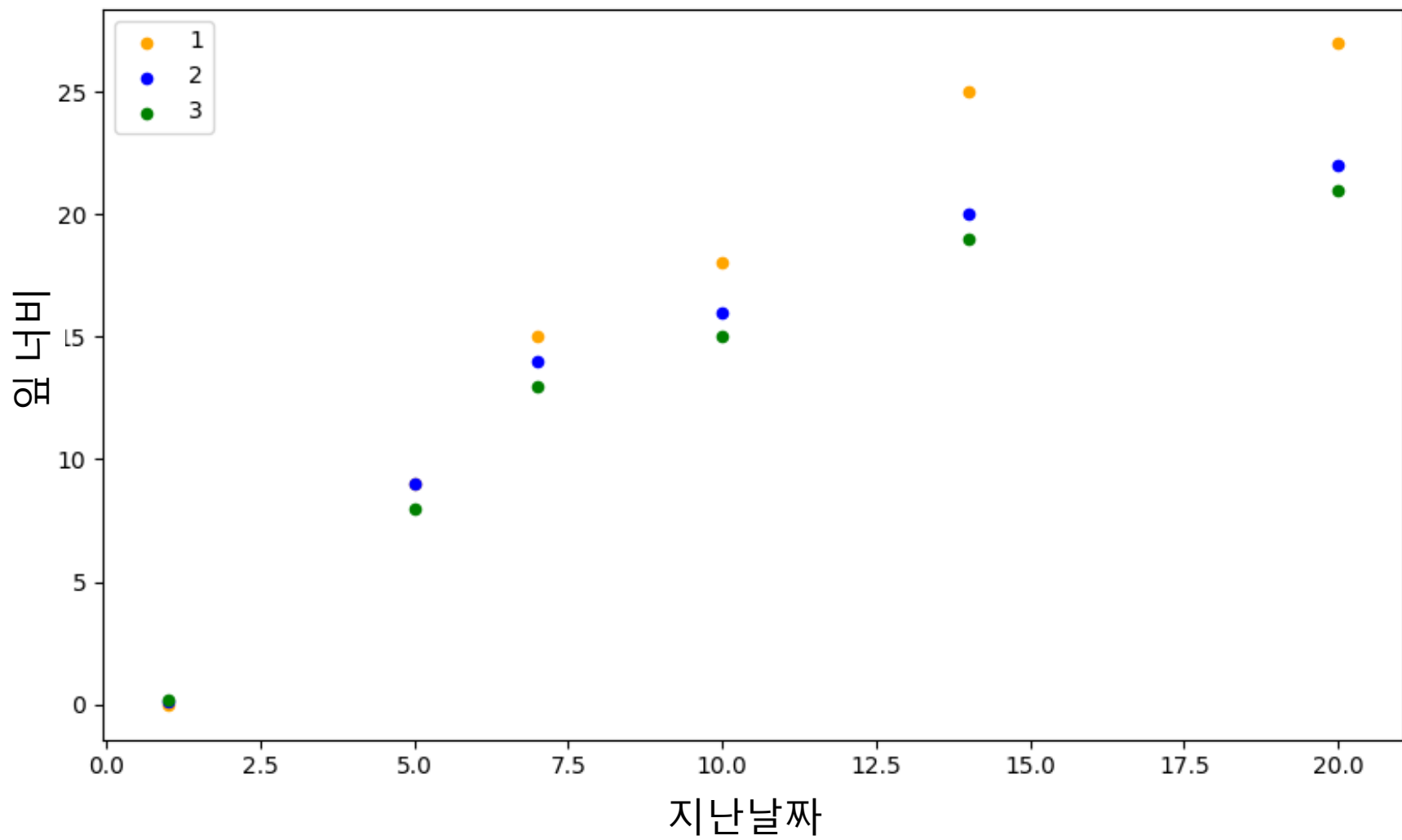
아까 그 데이터



첫째 아이, 둘째  
셋째

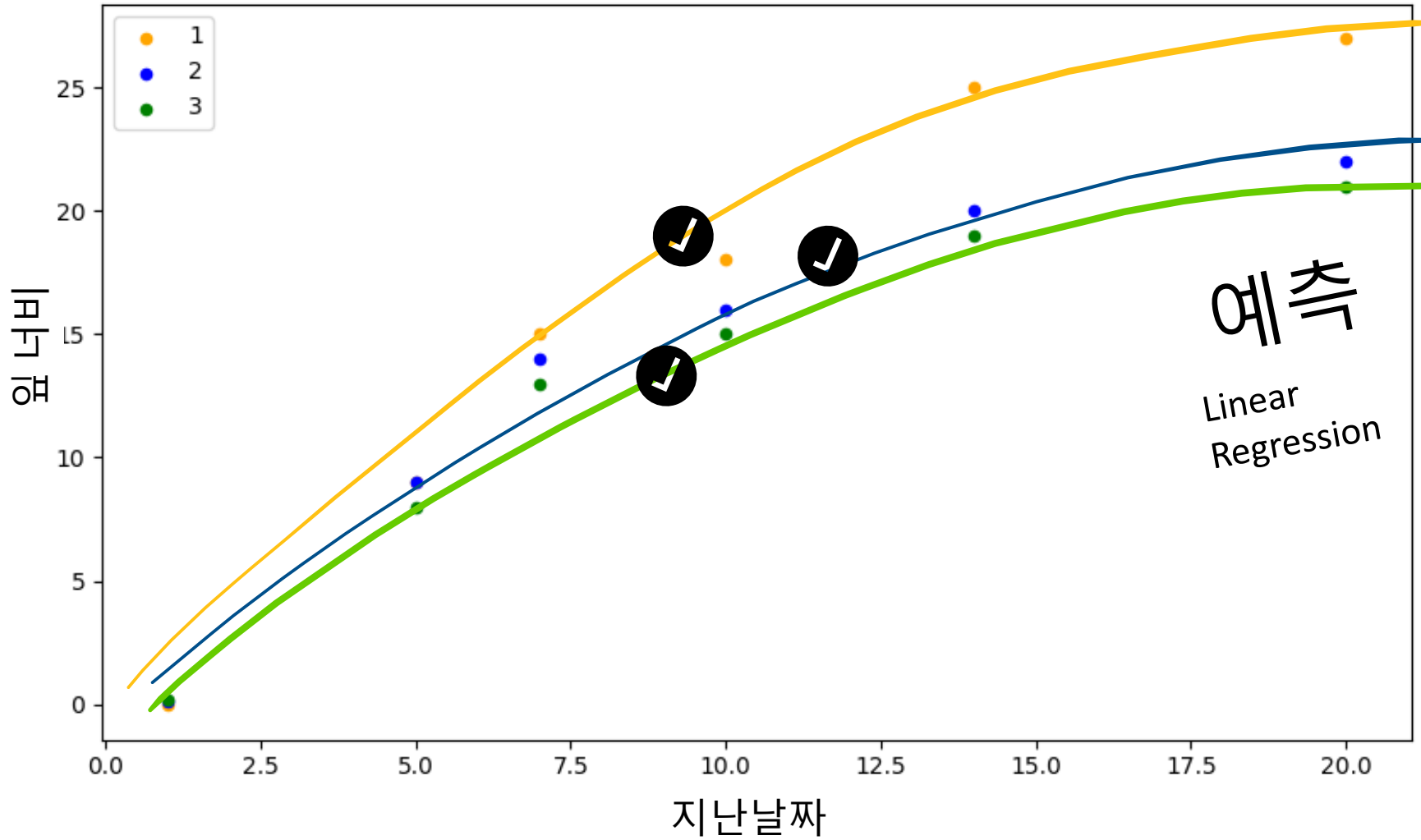




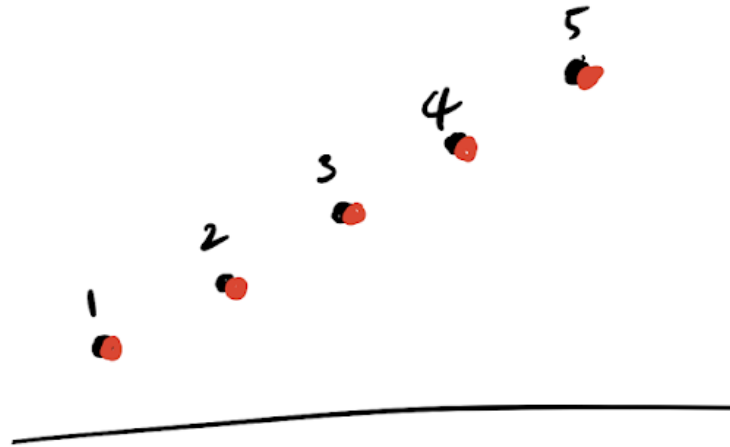




지난날짜 vs. 앞 너비



# 스코어(score)



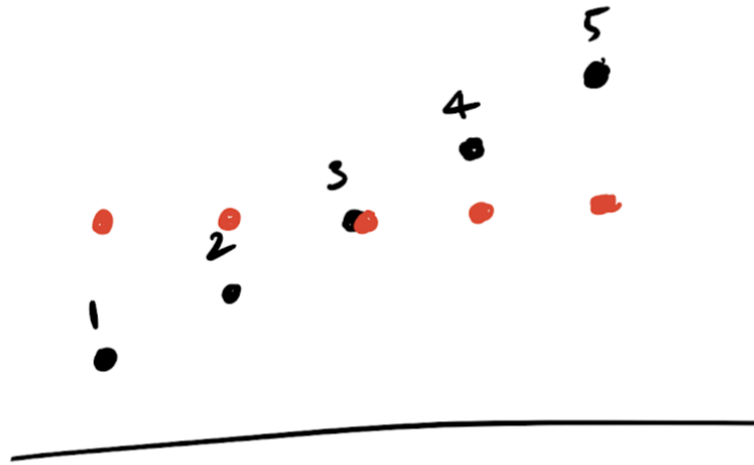
$$\text{score} = 1 - \frac{(\bullet - \bullet)^2}{(\bullet - \underbrace{\overline{(\bullet)}}_{\text{mean}})^2} = 1 - \frac{0}{10} = 1$$

↓

$$\begin{aligned} & (1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2 \\ &= 2^2 + 1^2 + 0^2 + 1^2 + 2^2 = 10 \end{aligned}$$



# 스코어(score)



$$\text{score} = 1 - \frac{(\bullet - \bullet)^2}{(\bullet - \overline{\text{color}})^2} = 1 - \frac{10}{10} = 1 - 1 = 0$$

$$\begin{aligned} &= (1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2 \\ &= 2^2 + 1^2 + 0^2 + 1^2 + 2^2 = 10 \end{aligned}$$

# 예측 알고리즘

Machine Learning

- KNeighborsRegressor (K-근접)
- DecisionTreeRegressor (결정 트리)
- RandomForestRegressor (랜덤 포레스트)
- Linear Regressor (선형 회귀)
- GradientBoostingRegressor (부스팅)
- XGBRegressor (부스팅)
- CatBoostRegressor (부스팅)

- NN-based LinearRegression
- MLPRegressor
- RNN/LSTM/GRU

Deep Learning





# 키 몸무게 발 크기 학년 성별



번호, 키, 몸무게, 발 크기, 학년, 성별

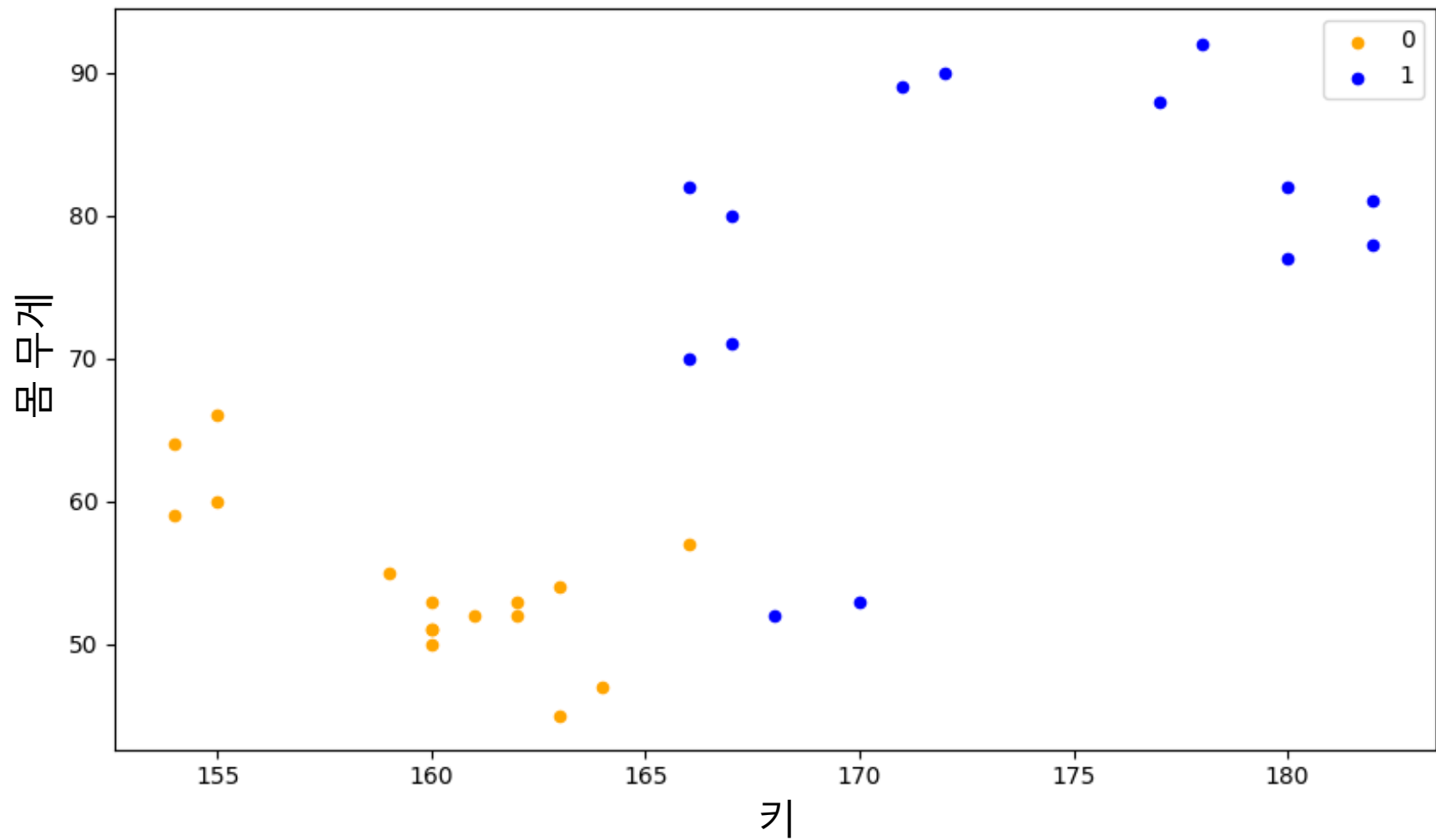
1,166,57,240,1,0  
2,178,92,265,1,1  
3,167,80,270,1,1  
4,168,52,245,2,1  
5,155,60,235,2,0  
6,163,45,230,2,0  
7,160,53,235,3,0  
8,180,77,260,4,1  
9,167,71,260,2,1  
10,160,51,245,2,0  
11,162,53,240,2,0  
12,180,82,280,6,1  
13,172,90,255,6,1  
14,160,51,245,5,0  
15,155,66,245,5,0  
16,163,54,242,5,0  
17,177,88,263,5,1  
18,166,82,268,6,1  
19,170,53,247,6,1  
20,154,59,234,1,0  
21,164,47,232,1,0



키에 따라  
몸무게는 어떻게 변할까?  
(성별에 따라 다른 색으로 표시)

```
plot(df, '키', '몸무게', '성별')
```

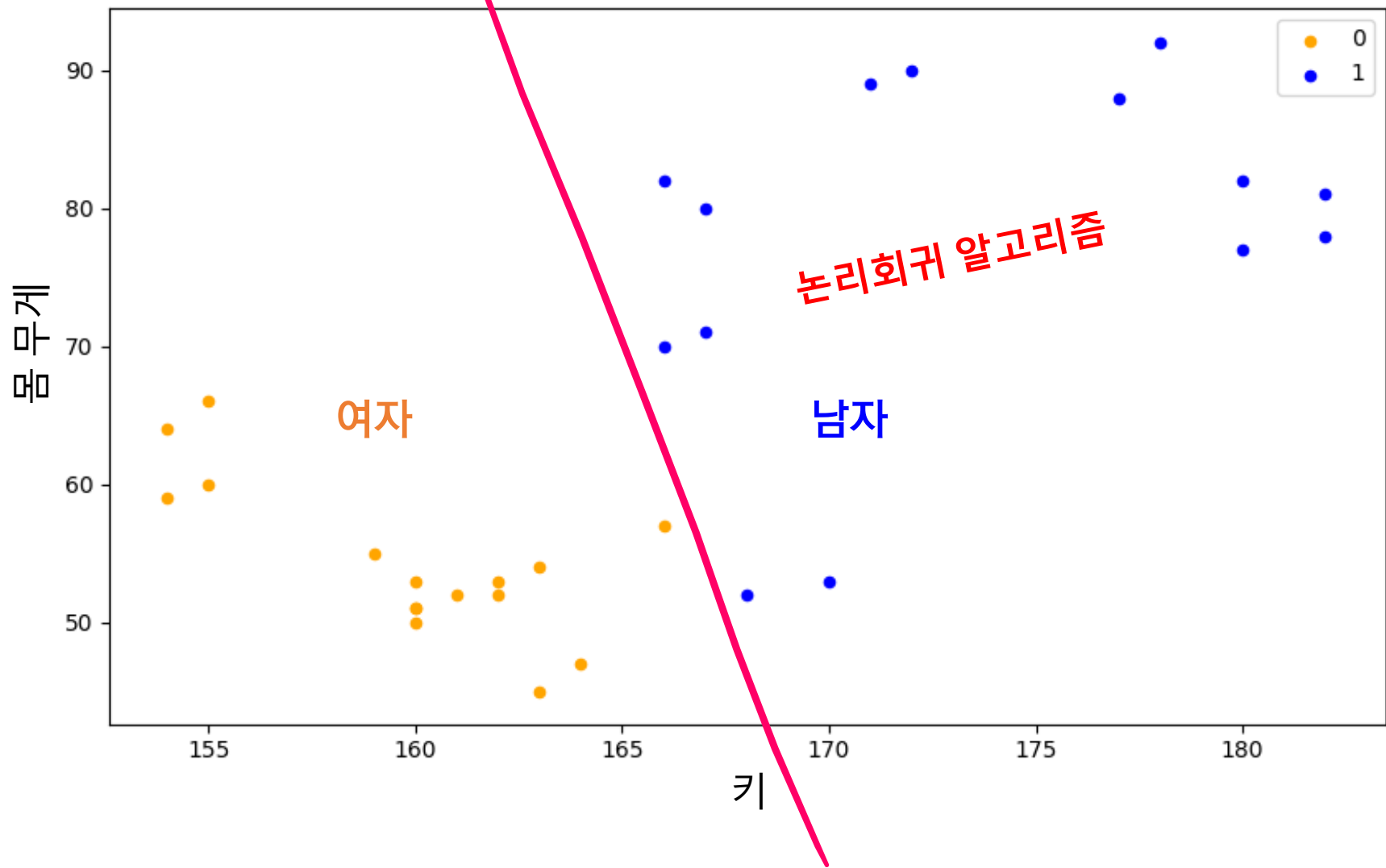
키 vs. 몸무게





분류

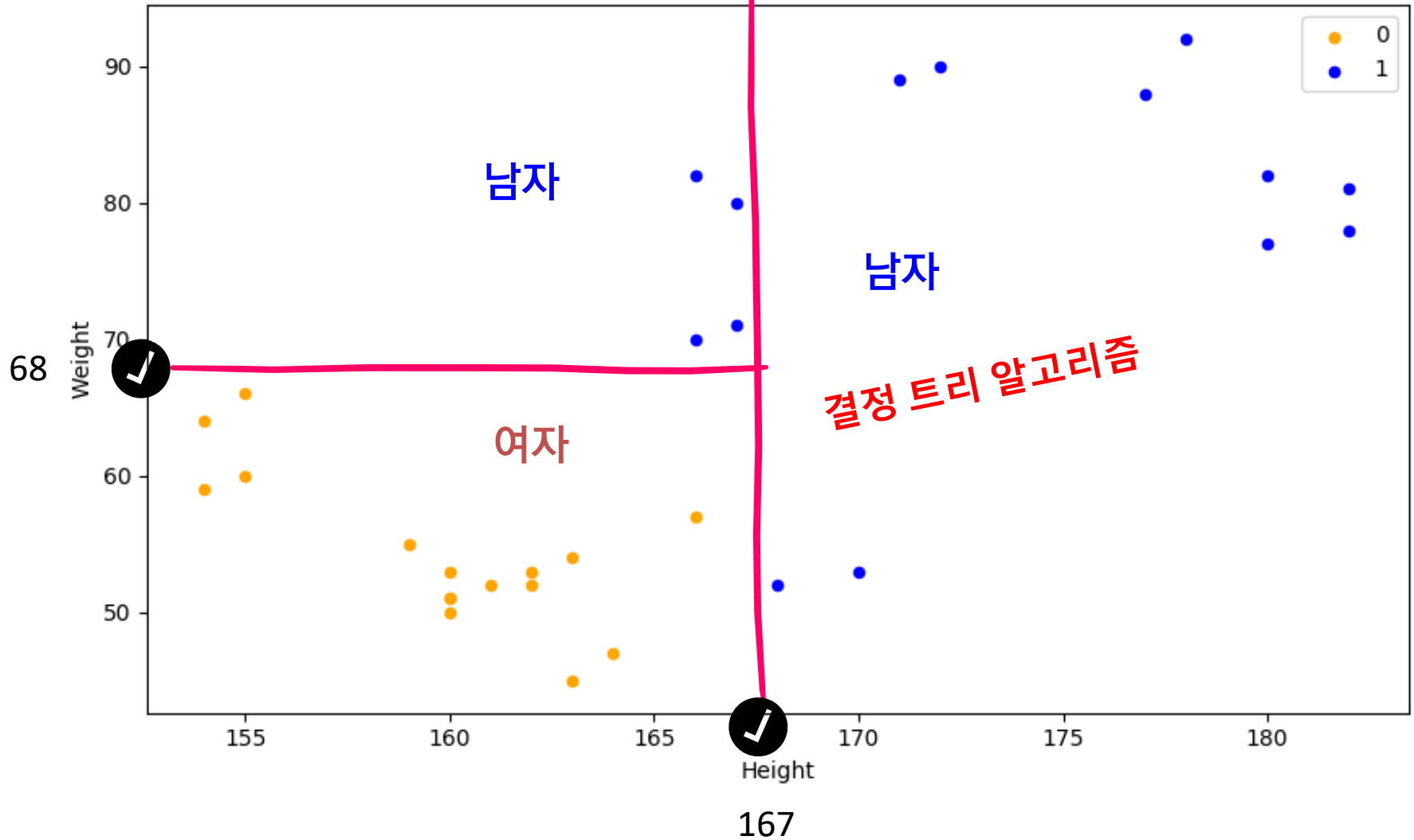
Classification



분류

Classification

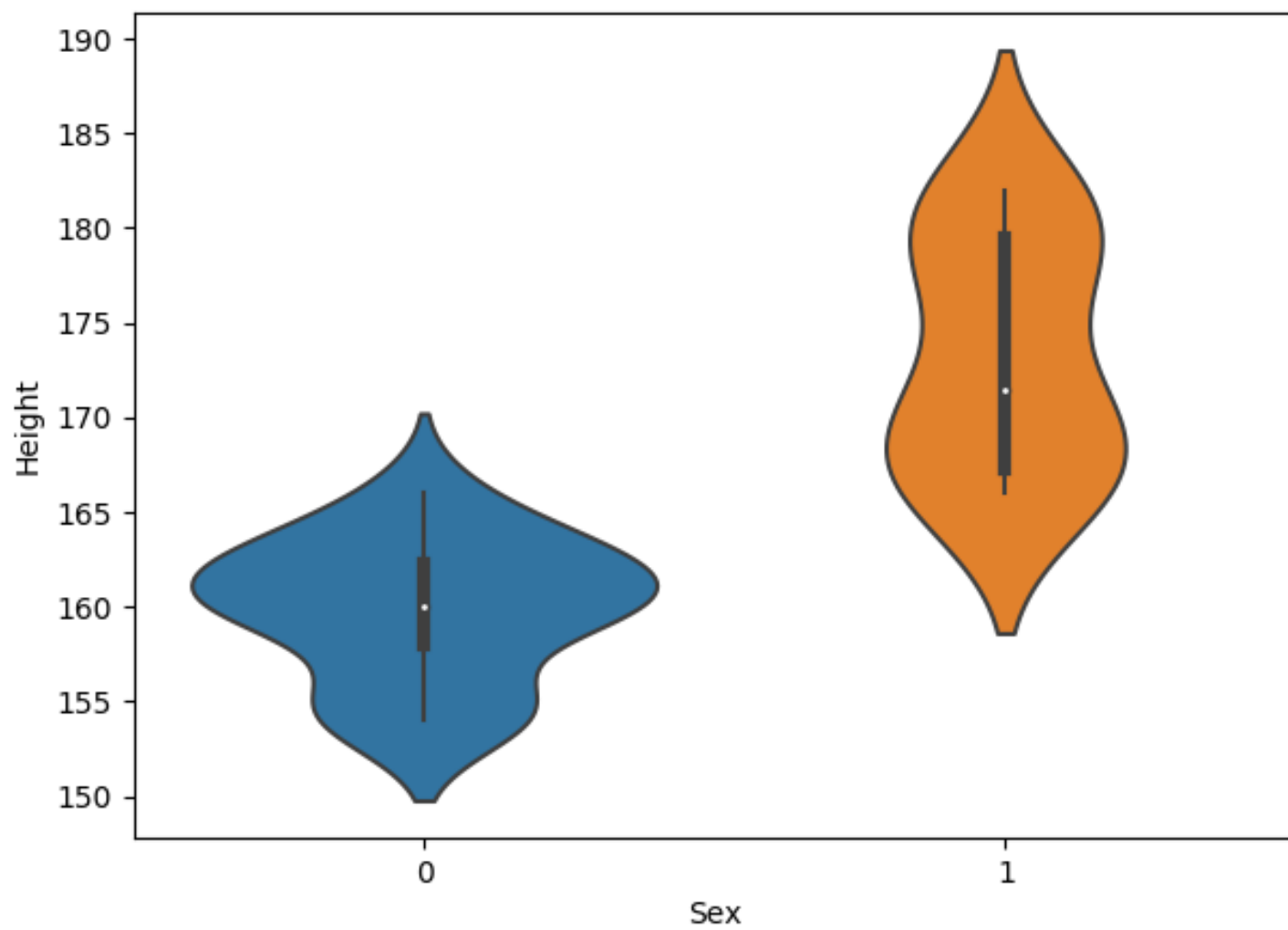
Height vs. Weight

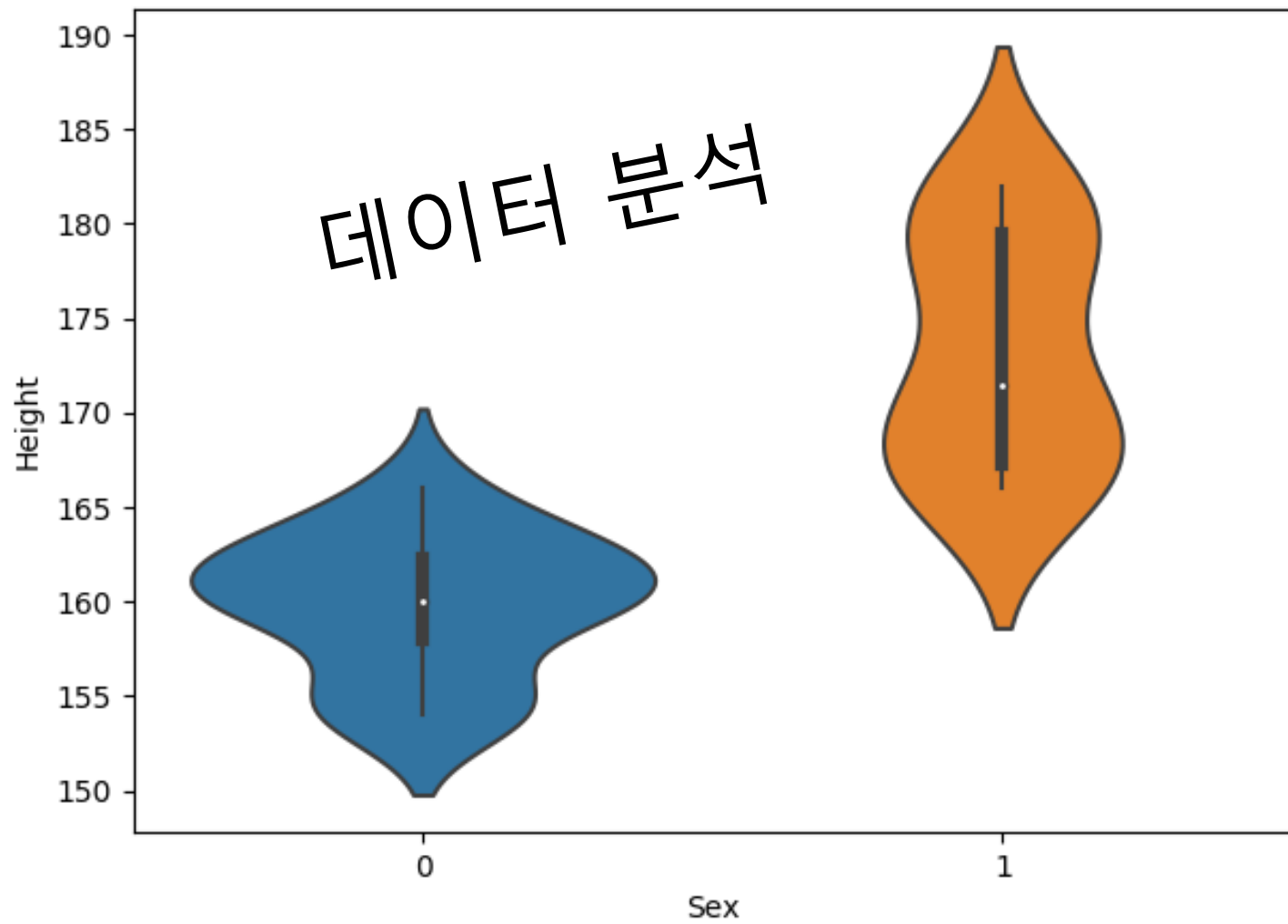


```
violinplot(df, '성별', '키')
```

성별에 따라 키가 어떻게 변하는지  
바이올린 모양으로 표시해보라!







키, 몸무게, 발 크기, 학년, 성별

학습용  
문제

166,57,240,1,0  
178,92,265,1,1  
167,80,270,1,1  
168,52,245,2,1  
155,60,235,2,0  
163,45,230,2,0  
160,53,235,3,0  
180,77,260,4,1  
167,71,260,2,1  
160,51,245,2,0  
162,53,240,2,0  
180,82,280,6,1  
172,90,255,6,1  
160,51,245,5,0

정답

테스트용  
문제

155,66,245,5,0  
163,54,242,5,0  
177,88,263,5,1  
166,82,268,6,1  
170,53,247,6,1  
154,59,234,1,0  
164,47,232,1,0

정답



키, 몸무게, 발 크기, 학년, 성별

학습용  
문제

166,57,240,1,0  
178,92,265,1,1  
167,80,270,1,1  
168,52,245,2,1  
155,60,235,2,0  
163,45,230,2,0  
160,53,235,3,0  
180,77,260,4,1  
167,71,260,2,1  
160,51,245,2,0  
162,53,240,2,0  
180,82,280,6,1  
172,90,255,6,1  
160,51,245,5,0

정답

테스트용  
문제

155,66,245,5,0  
163,54,242,5,0  
177,88,263,5,1  
166,82,268,6,1  
170,53,247,6,1  
154,59,234,1,0  
164,47,232,1,0

정답

```
youngJa = svm.SVC()
```

```
youngJa.fit('학습용문제', '정답')
```

```
prediction=youngJa.predict('테스트용 문제')
```

# 분류 알고리즘

Machine Learning

- SVC (서포트벡터머신)
- DecisionTreeClassifier (결정트리)
- RandomForestClassifier (랜덤포레스트)
- XGBClassifier (XGBoost, eXtreme Gradient Boosting, Boosting or Additive Training) (부스팅)

- LogisticRegression (논리회귀)
- Multilayer Neural Networks
- CNN/RCNN/GCNN

Deep Learning



# 머신러닝 인공지능



# 지능이란?

(지능, intelligence , 知能)

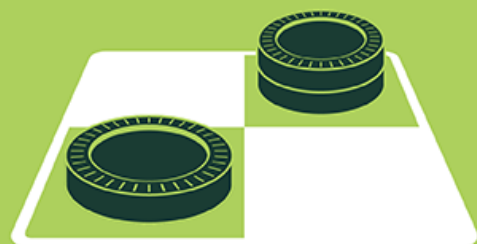
새로운 사물 현상에 부딪쳐 그 의미를  
이해하고 처리 방법을 알아내는 지적  
활동 능력

# 인공지능

AI (Artificial Intelligence),  
사람의 지능을 컴퓨터에  
구현한 지능

# ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



## MACHINE LEARNING

Machine learning begins to flourish.



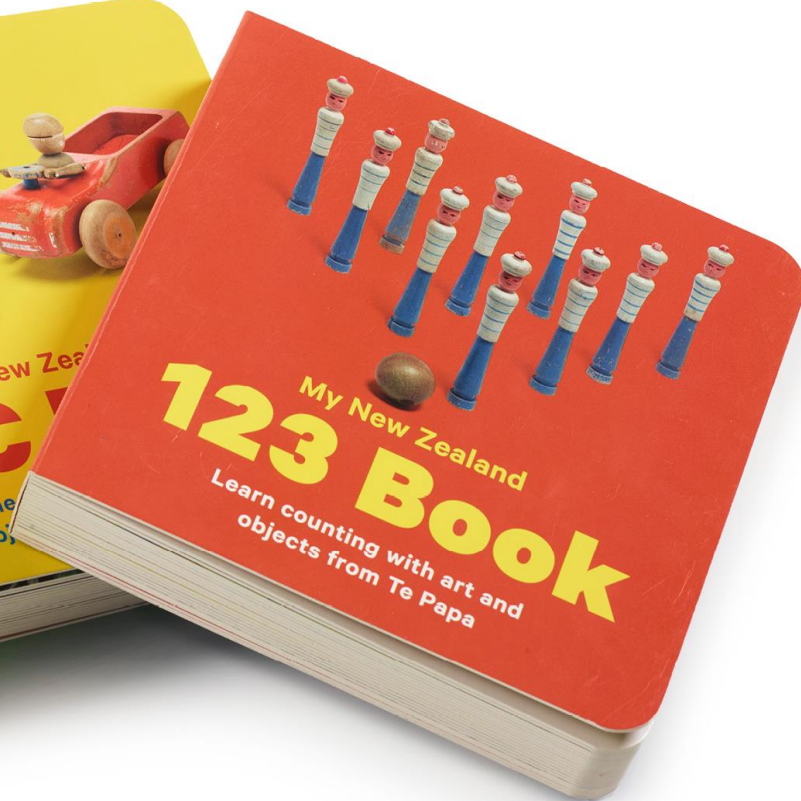
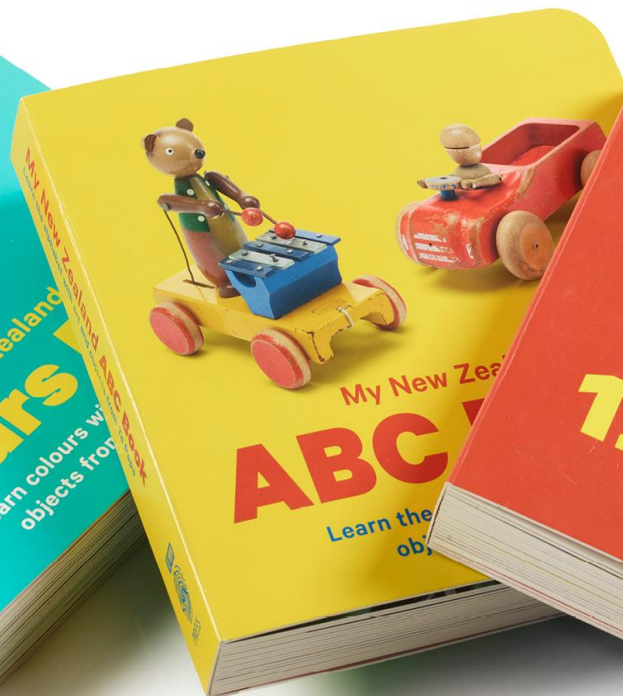
## DEEP LEARNING

Deep learning breakthroughs drive AI boom.



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.







개





토끼



돼지





오리



# 지도학습

## Supervised Learning

## *Machine Learning Techniques*

