

13장. 머신러닝 및 딥러닝



머신러닝(Machine Learning)

+ 머신러닝에 대해 설명해 줘



머신러닝이란? 명시적으로 규칙을 코딩하지 않아도, 데이터로부터 패턴을 학습해 예측·판단하는 인공지능 기술

머신러닝의 핵심 개념

1 데이터

- 학습의 재료
- 많고, 정확할수록 성능이 좋아짐

2 모델(Model)

- 데이터를 학습한 결과물
- 예: “이메일이 스팸일 확률은 92%”

3 학습(Training)

- 데이터를 보고 모델이 규칙을 조정하는 과정

4 예측(Prediction)

- 새로운 데이터에 대해 결과를 맞혀보는 것

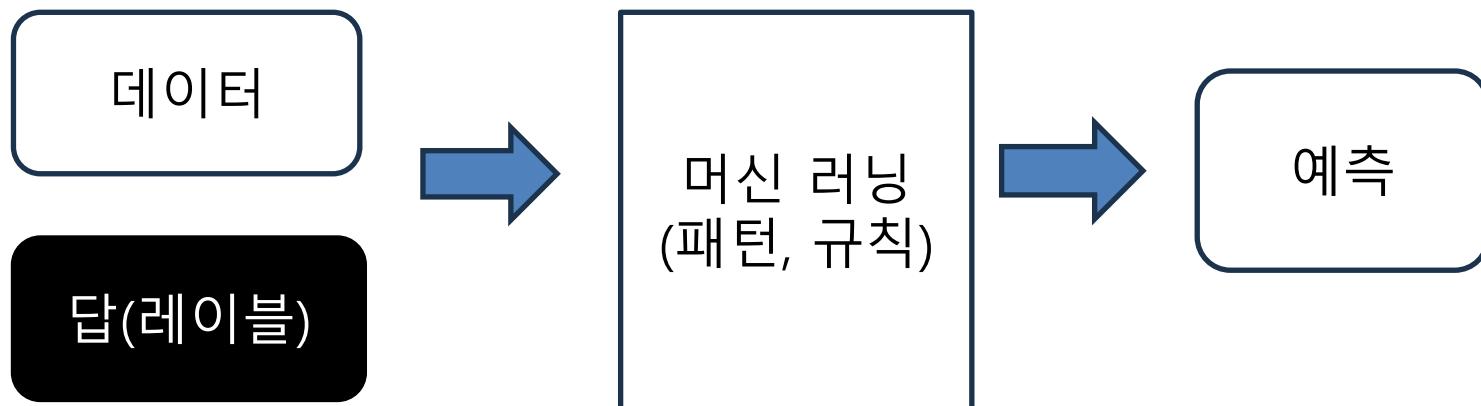
머신러닝(Machine Learning)

❖ 지도 학습 vs 비지도 학습

	지도 학습	비지도학습
분석 모형 (알고리즘)	<ul style="list-style-type: none">회귀분류	<ul style="list-style-type: none">군집차원 축소
특징	<ul style="list-style-type: none">정답(레이블)을 알고 있는 상태에서 학습회귀 – 가격, 매출, 주가 예측분류 – 고객분류, 질병 진단, 스팸 메일 필터링	<ul style="list-style-type: none">정답(레이블)이 없는 상태에서 서로 비슷한 데이터를 찾아서 그룹화(클러스터링)구매 패턴 분석, 유튜브 추천 등

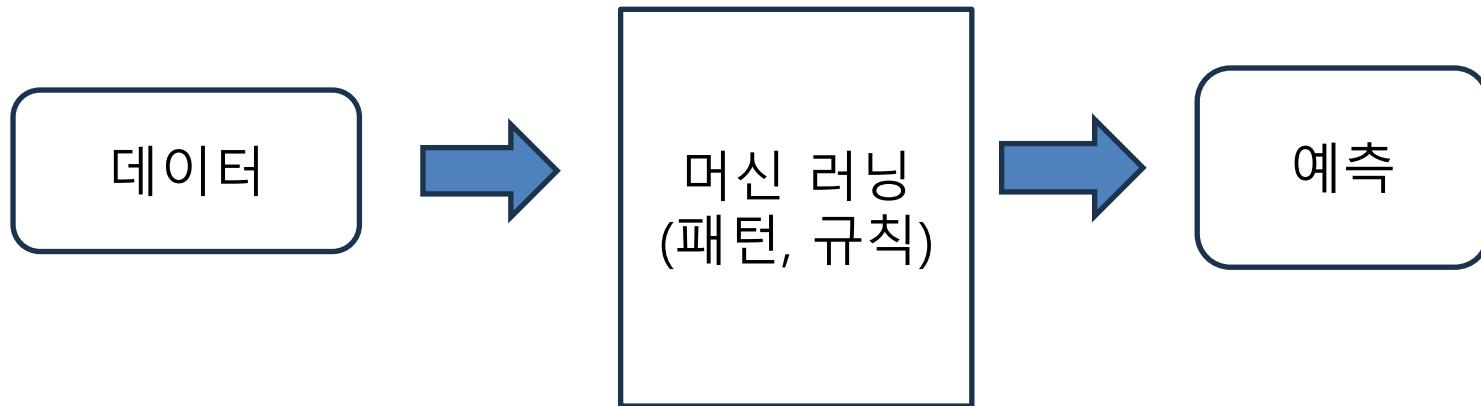
머신러닝(Machine Learning)

❖ 지도 학습



머신러닝(Machine Learning)

❖ 비지도 학습

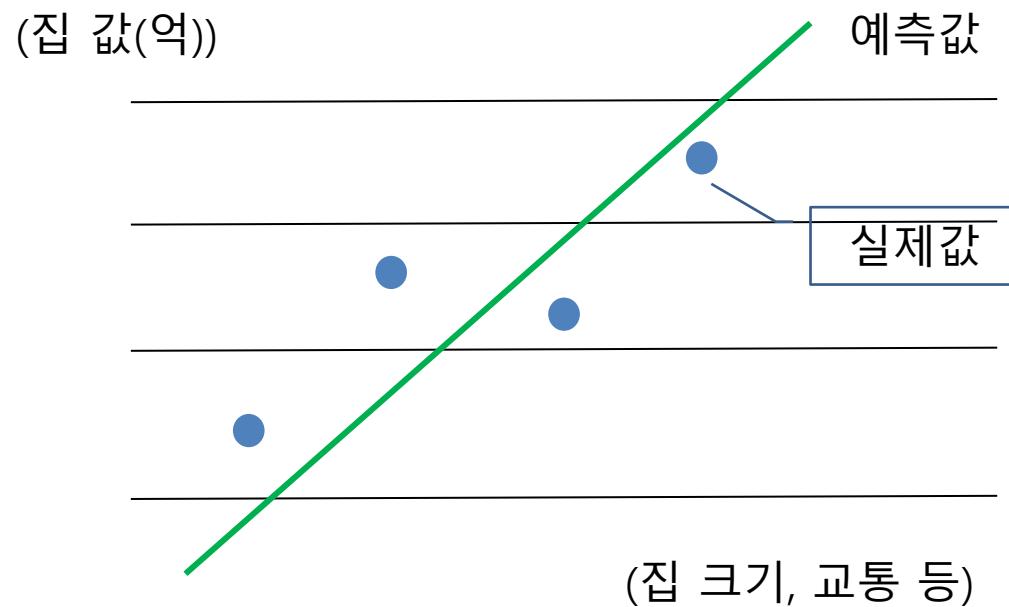


머신러닝(Machine Learning)

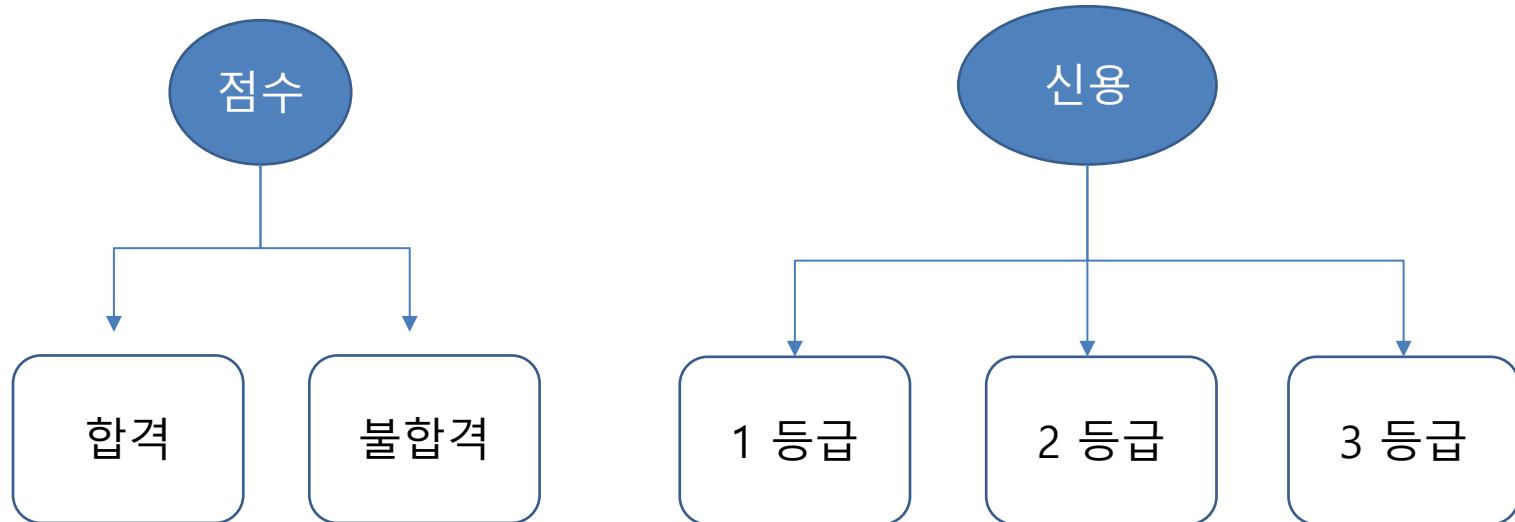
❖ 강화 학습

- 강화란 '어떤 것의 수준이나 정도를 높인다'는 의미로 인공지능의 수준을 높이는 것을 말한다.
 - 에이전트가 환경과 상호작용하면서 보상 신호에 따라 최적의 정책을 찾아내는 방법
-
- 벽돌깨기 게임 – 딥마인드(구글) 회사
 - 알파고(딥마인드) – 바둑을 학습한 인공지능
 - 자율주행 자동차, 인공지능 로봇으로 확장되고 있음

회귀(Regression)

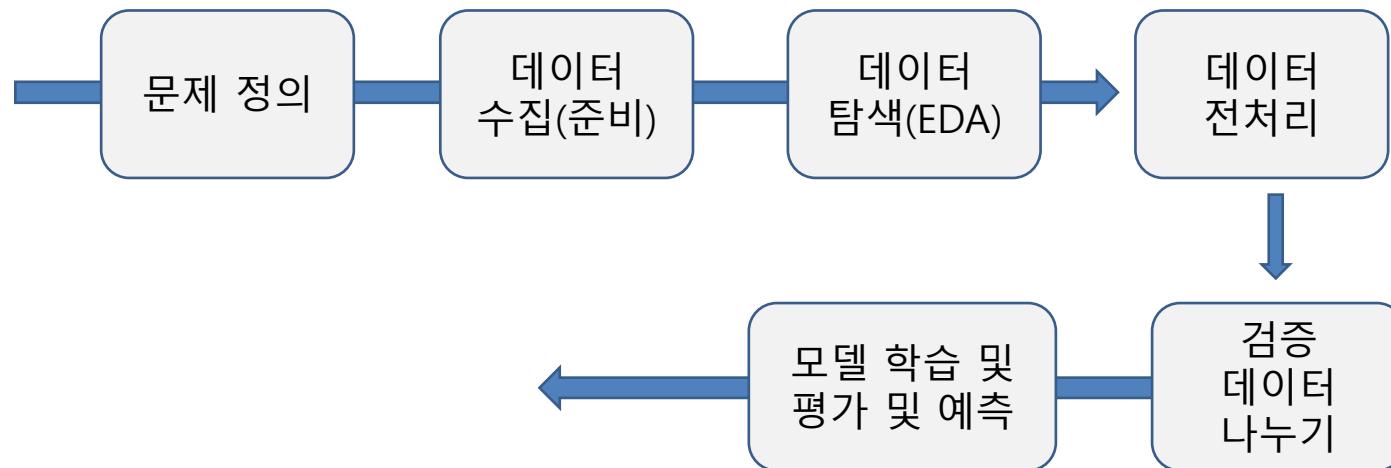


분류(Classification)



머신러닝 프로세스

- 머신러닝 모델 예측 프로세싱



데이터 수집(준비)

- 데이터는 학습용과 검증으로 준비한다.

✓ 훈련 : 검증 비율

80 : 20 or 70 : 30

train
(훈련용)

X_train

y_train

test
(검증용)

test

y_test

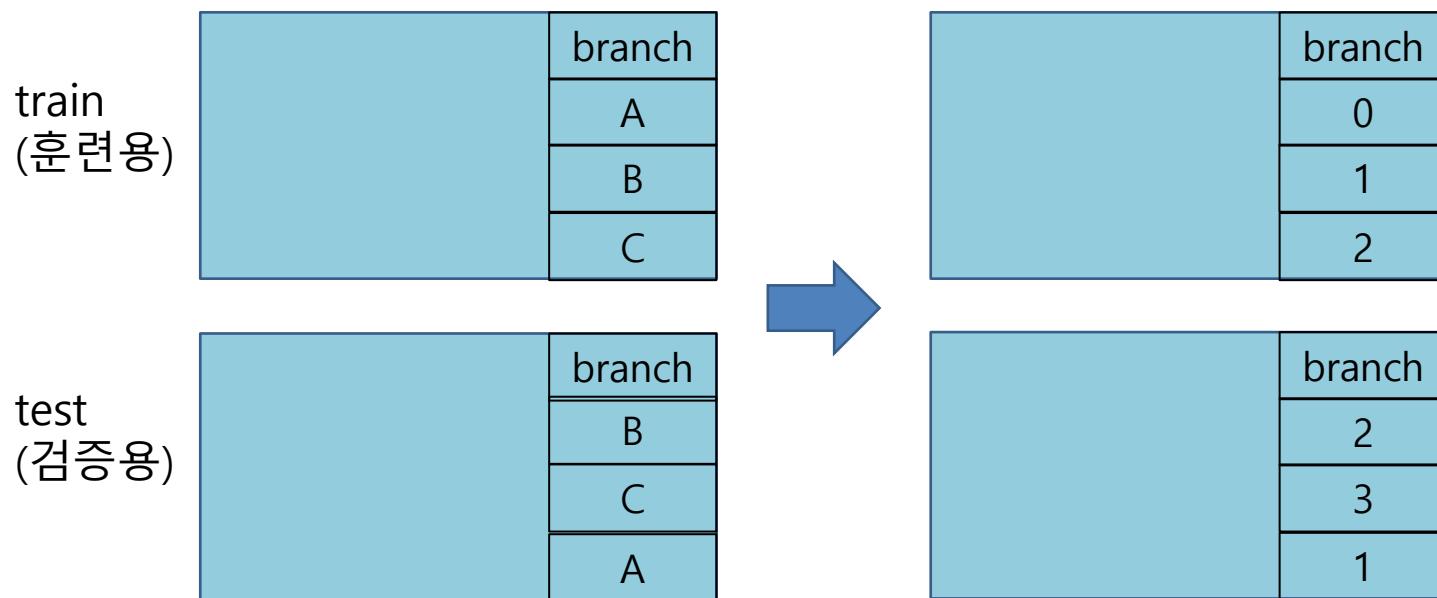
데이터 전처리

- 데이터 처리를 위한 모듈

	알고리즘	필수 여부
레이블 인코딩	sklearn.preprocessing(모듈) LabelEncoder	필수 (원-핫 인코딩중 선택)
스케일링	sklearn. preprocessing(모듈) StandardScaler	선택
데이터 분할	sklearn. model_selection(모듈) train_test_split	필수

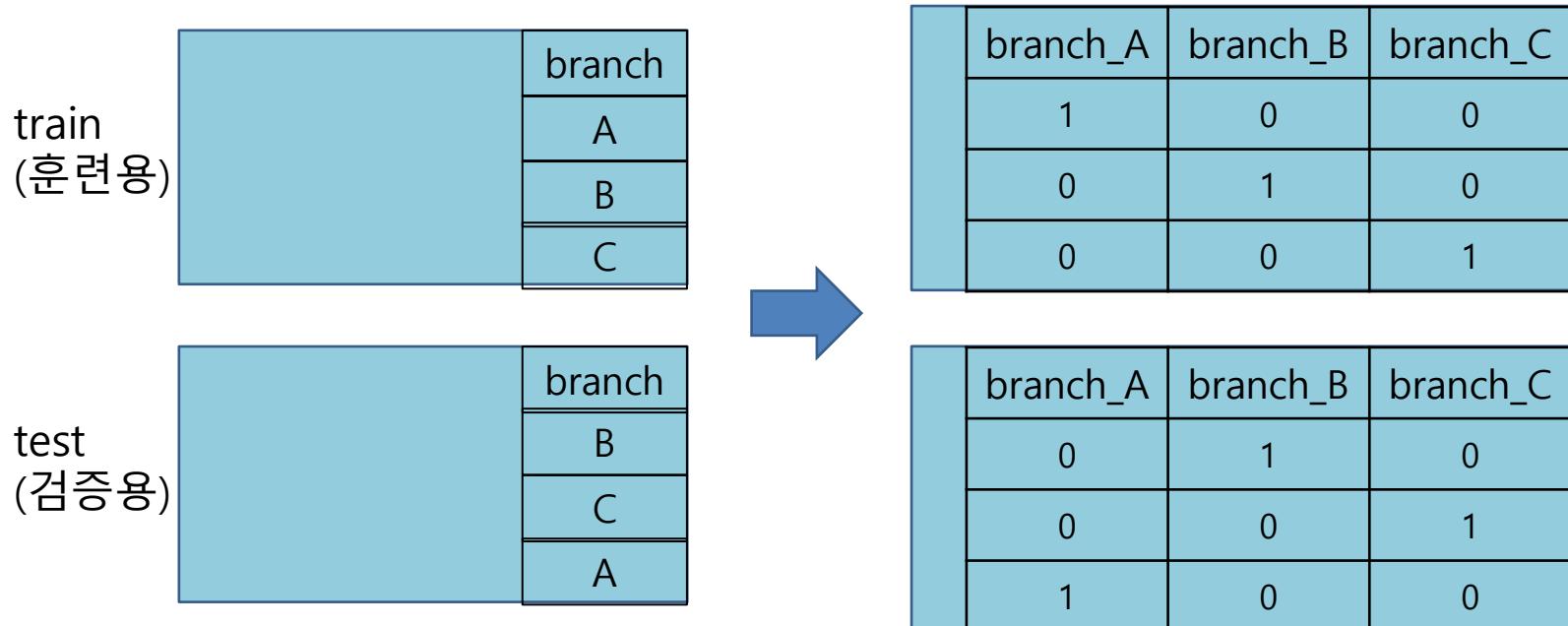
데이터 전처리

- 인코딩(Encoding) – 문자형을 숫자형으로 변환
 - 레이블 인코딩(Label Encoding) : 각 칼럼을 숫자로 맵핑하는 인코딩 방법이다.



데이터 전처리

- 인코딩(Encoding) – 문자형을 숫자형으로 변환
 - 원-핫 인코딩(One-Hot Encoding) : 각 칼럼에 대해 새로운 칼럼을 만들고 새로운 칼럼에 0 또는 1이 저장된다.



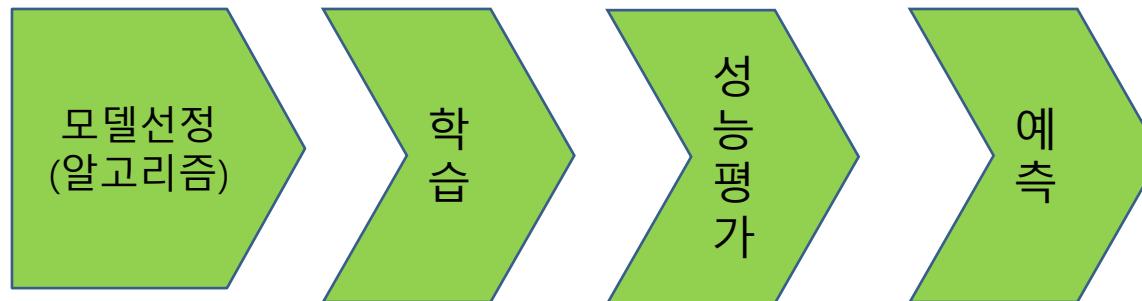
모델 학습 및 평가

● 모델 학습 및 평가

지도 학습	알고리즘	평가 지표
분류	sklearn.ensemble(모듈) RandomForestClassifier	sklearn.metrics(모듈) roc_auc_score, f1_score
회귀	sklearn.linear_model(모듈) LinearRegression sklearn.ensemble(모듈) RandomForestRegressor	sklearn.metrics(모듈) mean_squared_error mean_absolute_error

모델 학습 및 평가

- 모델 학습, 평가 및 예측



회귀(Regression)

❖ 선형 회귀



선형 회귀에 대해서 설명해줘

- ◆ 선형 회귀(Linear Regression)는 가장 기본적이고 널리 사용되는 통계 및 머신러닝 기법 중 하나입니다. 한 개 이상의 독립 변수(설명 변수)와 종속 변수(예측 변수) 사이의 선형 관계를 모델링하는 데 사용됩니다. 즉, 입력 값(공부 시간)과 출력 값(시험 점수) 사이에 가장 잘 맞는 직선을 찾아 그 관계를 설명하고 미래 값을 예측하는 것이 목

[공부 시간과 시험 점수 데이터를 사용하여 선형 회귀 모델을 만들어줘](#)

[선형 회귀의 수학적 원리에 대해 설명](#)



빌드하는 데 어떤 도움이 필요하신가요?



Gemini 2.5 Flash ▾ ▶

Gemini는 실수할 수 있으므로, 대답을 다시 한번 확인하고 신중하게 코드를 사용하시기 바랍니다. [자세히 알아보기](#)

선형 회귀 분석

● 단순 선형 회귀

- $y = Wx + b$
- W 를 가중치(Weight), b 를 편향(bias)라고 부른다
- 그래프의 형태는 **직선**

● 다중 선형 회귀

- $y = W_1x_1 + W_2x_2 + \dots + W_nx_n + b$
- 만약 2개의 독립 변수면 그래프는 **곡선**으로 나타남

선형 회귀 분석

● 공부한 시간의 차이에 따른 시험 성적 예측하기

- 성적을 변하게 하는 요소를 x 라 하고 이 x 값에 따라 변하는 '성적'을 y 라 하자.
- 독립 변수(x) : 공부한 시간, 사교육비 지출, 학생의 지능지수 등
- 종속 변수(y) : 성적
- 어떤 변수가 다른 변수에 영향을 준다면 두 변수 사이에 선형관계가 있다.

선형 회귀 분석

- 공부한 시간의 차이에 따른 시험 성적 예측하기

공부 시간	2시간	4시간	6시간	8시간
성적	81점	93점	91점	97점

$$x = \{2, 4, 6, 8\}$$

$$y = \{81, 93, 91, 97\}$$

직선의 방정식을 구한다.

$$y = ax + b \text{(일차함수)}$$

기울기 a 와 절편 b 를 알아야 함

최소 제곱법

- 키의 차이에 따른 몸무게 예측하기

- 기울기 a 구하기

$$a = \frac{(x - x\text{평균})(y - y\text{평균})\text{의 합}}{(x - x\text{평균})^2\text{의 합}}$$

공부 시간(x) 평균 : $(2 + 4 + 6 + 8) \div 4 = 5$

성적(y) 평균: $(81 + 93 + 91 + 97) \div 4 = 90.5$

기울기 $a = 2.3$

x의 편차(각 값과 평균과의 차이)를 제곱해서 합한 값을 분모로 놓고,
x와 y의 편차를 곱해서 합한 값을 분자로 놓으면 기울기가 나옴

최소 제곱법

- 공부한 시간의 차이에 따른 시험 성적 예측하기

- 편차(deviation)

- 관측값에서 평균을 뺀 값

- 분산 (variance)

- 편차(관측값-평균)를 제곱하고 그것을 모두 더한 후 전체 개수로 나눈것. 즉 편차의 제곱의 평균을 말한다.

- 표준편차(standard deviation)

- 분산의 제곱근, 즉 분산 값에 루트를 씌운것을 말하며 stdev라고 함.

최소 제곱법

● 공부한 시간의 차이에 따른 시험 성적 예측하기

- y 절편 b 값 구하기

$$b = y\text{의 평균} - (\text{기울기 } a \times (x\text{의 평균}))$$

$$b = 90.5 - (2.3 \times 5) = 79$$

최적의 직선(방정식)

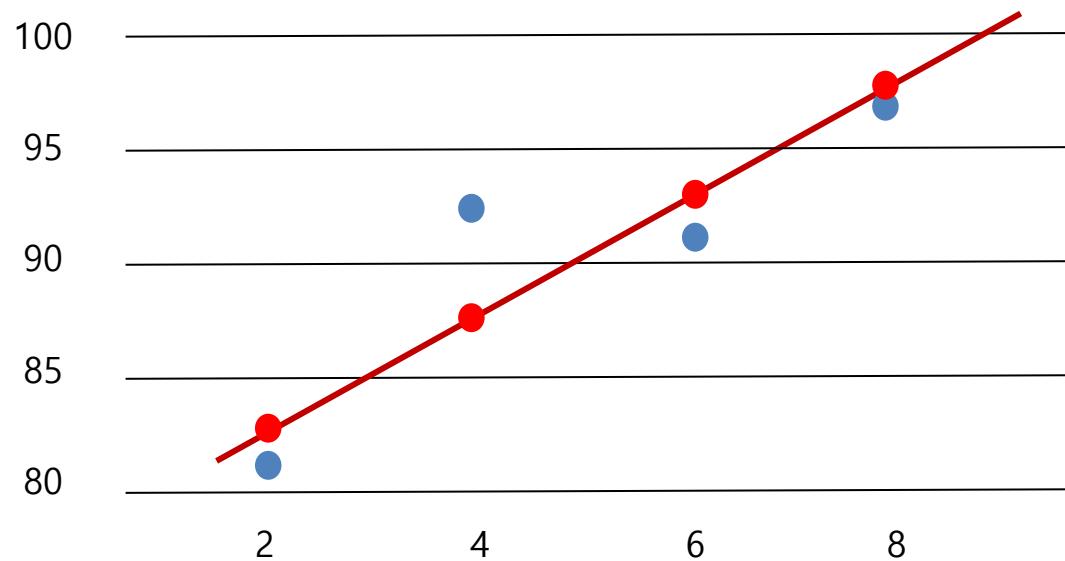
$$y = 2.3x + 79$$

공부 시간	2	4	6	8
성적	81	93	91	97
예측 값	83.6	88.2	92.8	97.4

최소 제곱법

- 공부한 시간의 차이에 따른 시험 성적 예측하기

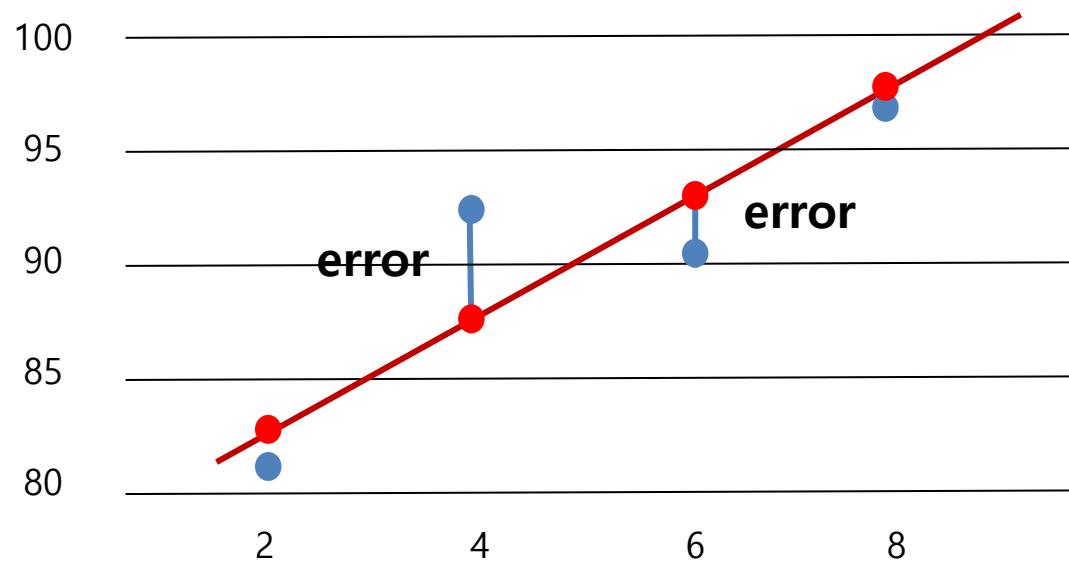
공부 시간	2	4	6	8
성적	81	93	91	97
예측 값	83.6	88.2	92.8	97.4



평균 제곱 오차

▪ 평균 제곱 오차

선형 회귀는 임의의 직선을 그어 이에 대한 평균 제곱 오차를 구하고, 이를 값이 가장 작게 만들어주는 기울기(a)와 절편(b)을 찾아가는 작업이다.



평균 제곱 오차

- 성능 평가 - 평균 제곱 오차 / 평균 절대값 오차

- MSE(Mean Squared Error)

- 평균 제곱오차
 - 오차의 제곱을 평균으로 나눈 것

$$MSE = \frac{\sum_{i=1}^n (y - \hat{y})^2}{n}$$

- MAE(Mean Absolute Error)

- 평균 절대값 오차
 - 오차의 차이를 절대값으로 변환한 뒤 합산

$$MAE = \frac{\sum |y - \hat{y}|}{n}$$

평균 제곱 오차

- 오차 = 실제값 - 예측값

공부 시간	2	4	6	8
성적	81	93	91	97
예측 값	82	88	94	100
오차	1	-5	3	3

$$\text{분모} : 1 + 25 + 9 + 9 = 44$$

$$\text{MSE} = 44 / 4 = 11$$

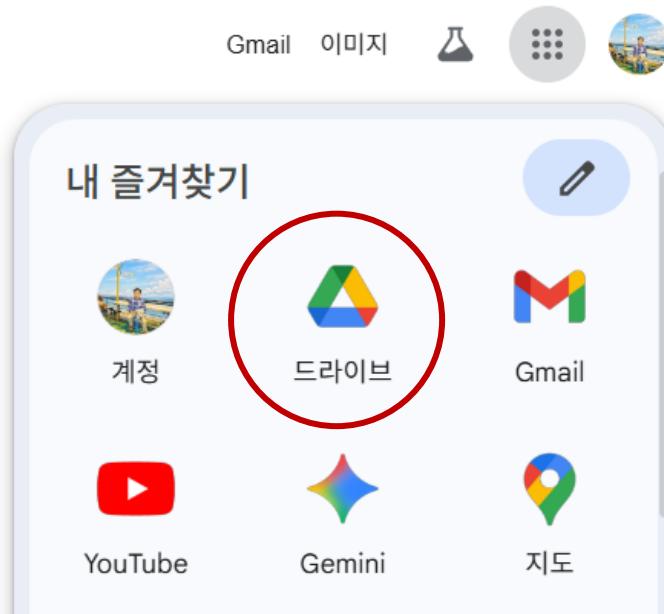
$$\text{MAE} = 1+5+3+3 = 12$$

구글 코랩

❖ Google Colab 사용하기

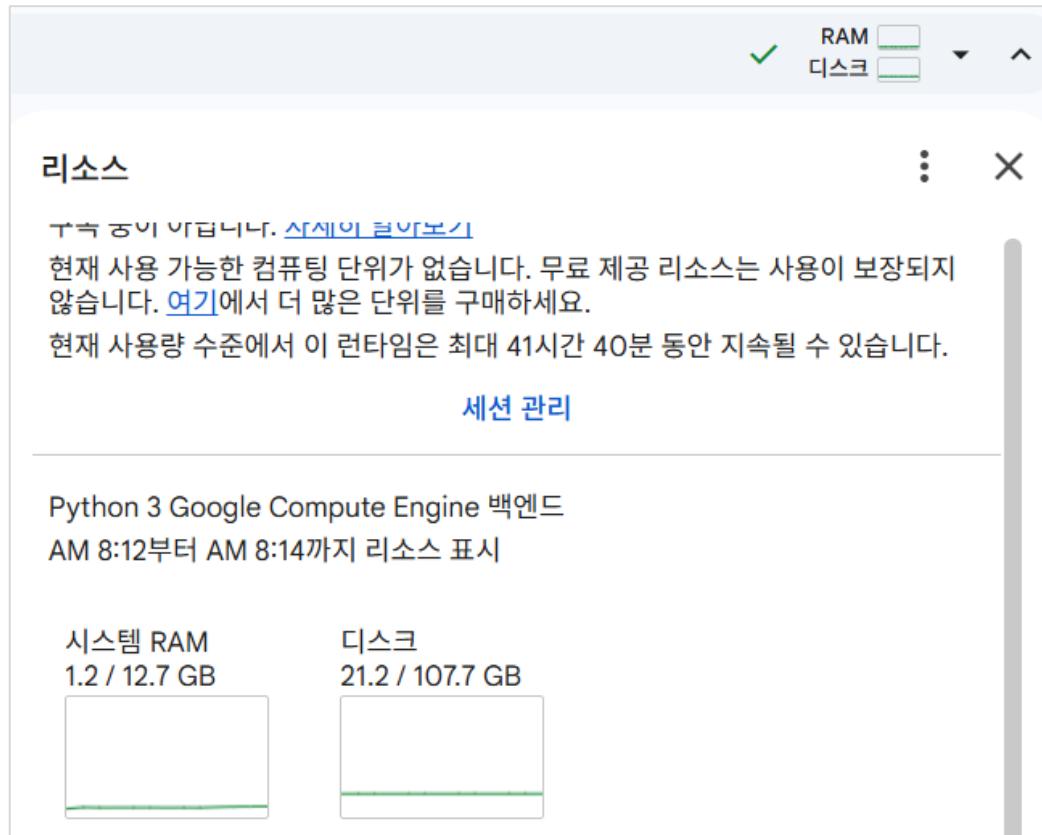
Colaboratory는 구글의 Jupyter Notebook 기반 클라우드 서비스로, TPU나 GPU, CPU를 무료 제공하며 클라우드 IDE의 일부이다

- 계정 로그인 > 메뉴(선택) > 드라이브



구글 코랩

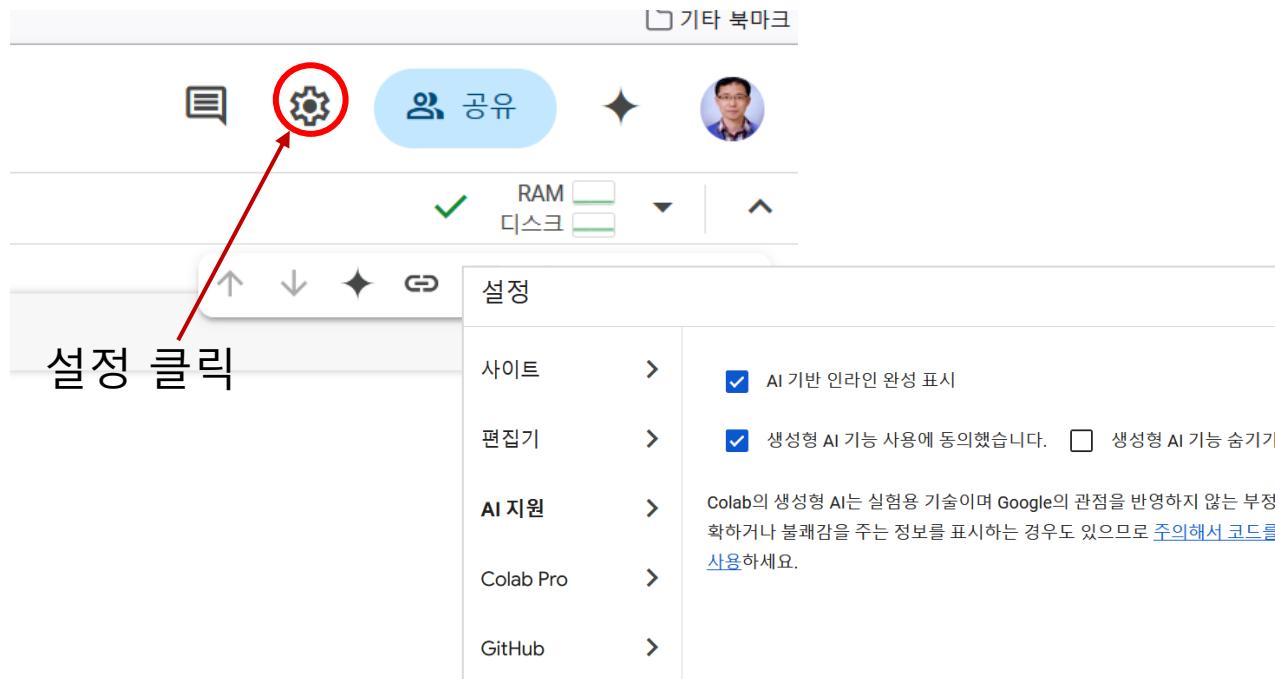
❖ 리소스 사용 – 클라우드 컴퓨터 사용



구글 코랩

❖ 구글 코랩(Colaboratory)

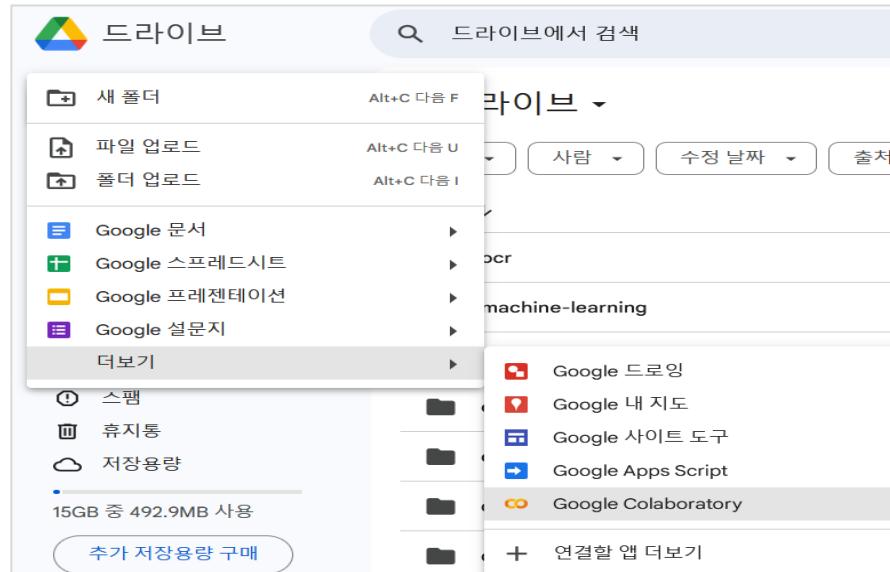
AI(Gemini) 지원 : 설정 > AI 지원 > 동의



구글 코랩

❖ 새 노트 만들기

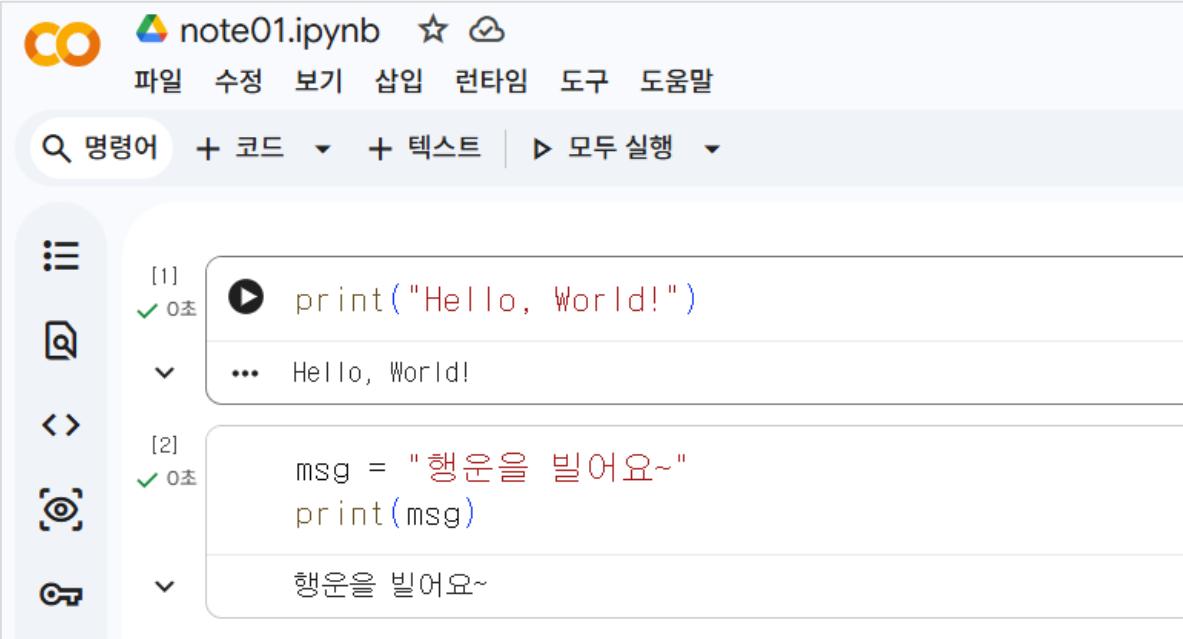
- 드라이브 > 홈
- + 신규 > 더보기 > Google Colaboratory



구글 코랩

❖ 새 노트 만들기

- 파일이름 > note01.ipynb



```
[1]: print("Hello, World!")
... Hello, World!

[2]: msg = "행운을 빌어요~"
print(msg)
행운을 빌어요~
```

구글 코랩

❖ 새 노트 만들기

- 파일이름 > note01.ipynb

The screenshot shows the Google Colab interface. At the top, there is a code cell containing Python code to create a DataFrame:

```
▶ import pandas as pd

df = pd.DataFrame({
    "name": ['마우스', '모니터'],
    "price": [20000, 80000]
})

df
```

Below the code cell is the output of the DataFrame:

	name	price
...		
0	마우스	20000
1	모니터	80000

At the bottom of the interface, there are two buttons: "다음 단계:" and "df 변수로 코드 생성". There is also a link to "New interactive sheet".

구글 코랩

❖ 새 폴더 만들기

- 드라이브 > 내 드라이브
- + 신규 > 새폴더 > 머신러닝

내 드라이브 ▾

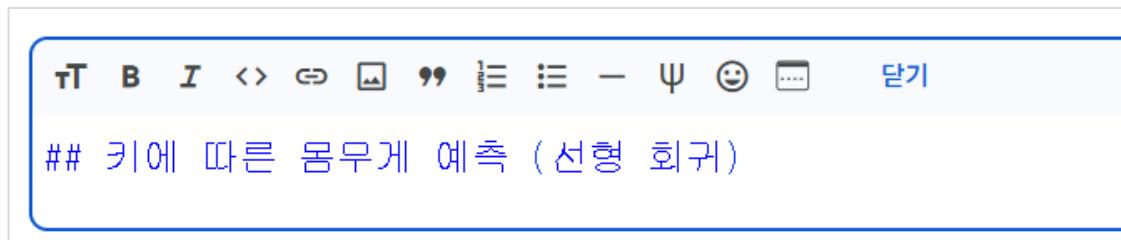
유형 ▾ 사람 ▾ 수정 날짜 ▾ 출처 ▾

이름	▲	소유자
머신러닝		나
note01.ipynb		나

구글 코랩

❖ 텍스트 사용하기

- 파일 이름 : 선형회귀.ipynb
- 텍스트(버튼) > 문서의 제목을 만들때 사용



<편집화면>
shift + enter

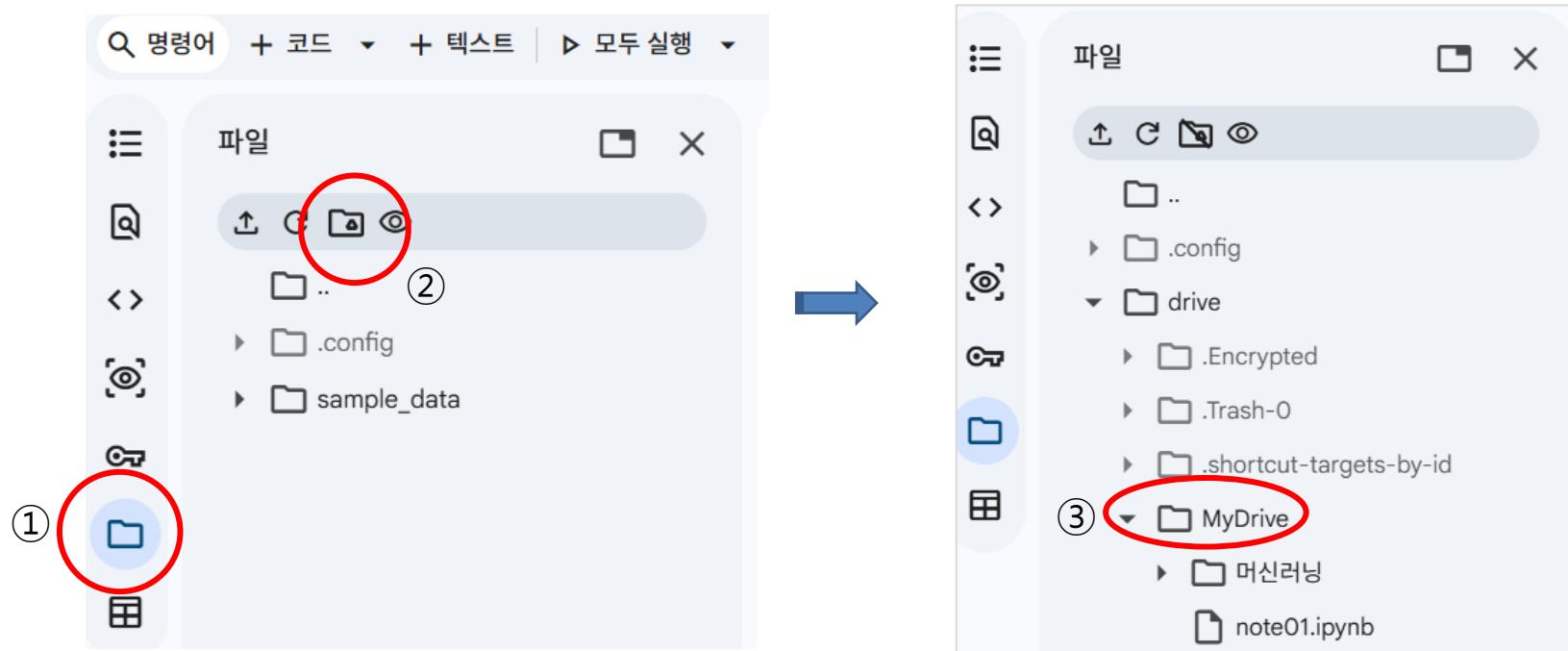
키에 따른 몸무게 예측 (선형 회귀)

<출력화면>

구글 코랩

❖ 드라이브 마운트

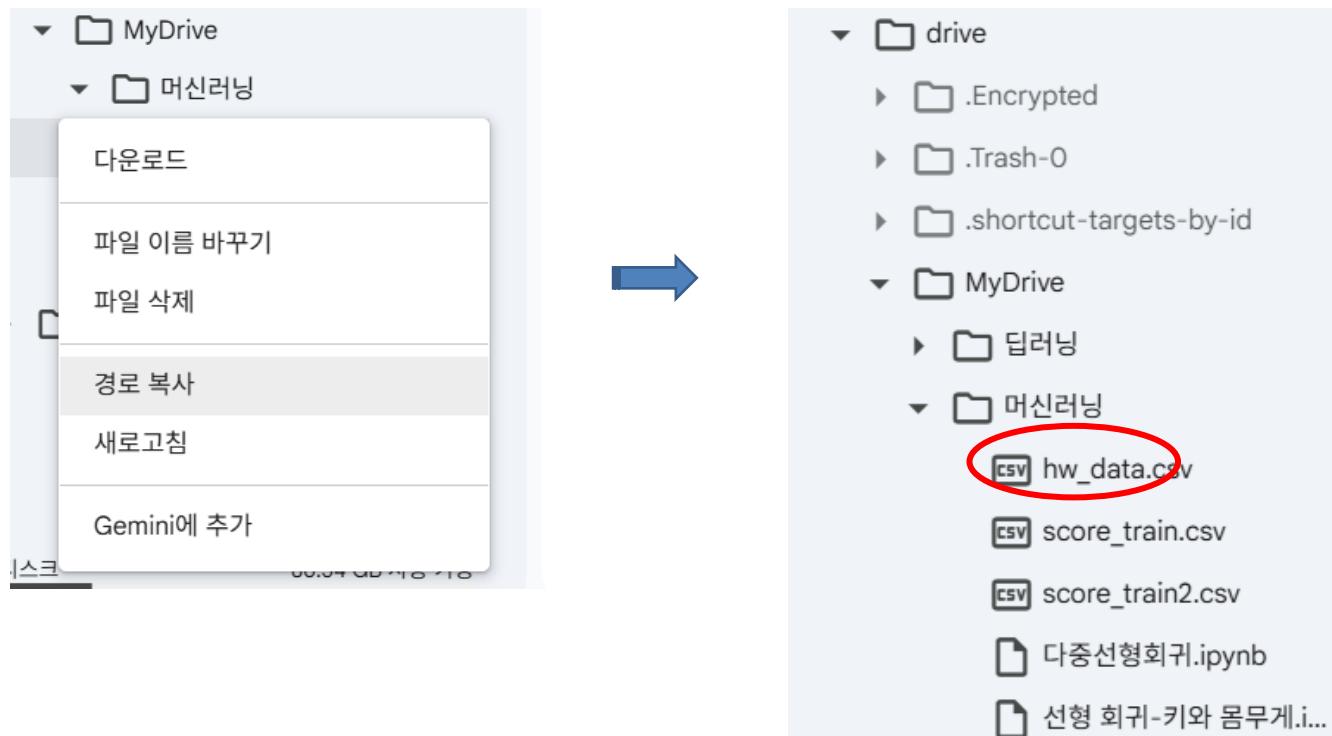
- ▶ 폴더 > 드라이브 마운트 > 드라이브 연결



구글 코랩

❖ 훈련 데이터 업로드

- (...) 메뉴 > 업로드 > data_hw.csv



키에 따른 몸무게 예측

● 데이터 준비

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# 가상의 키와 몸무게 데이터 생성
data_hw = {
    '키': [150, 160, 170, 175, 165, 155, 172, 168, 174, 158,
           162, 173, 185, 159, 167, 163, 171, 169, 180, 161],
    '몸무게': [42, 50, 70, 65, 60, 48, 68, 60, 65, 40,
               54, 67, 89, 51, 65, 60, 69, 71, 85, 53]
}
df_hw = pd.DataFrame(data_hw)

# csv 파일로 저장하기
df_hw.to_csv("/content/drive/MyDrive/머신러닝/hw_data.csv", index=False)

df_hw.head()
```

키에 따른 몸무게 예측

● 변수 선택

변수 선택

- 독립변수 (X): 키
- 종속변수 (y): 몸무게

```
▶ # csv 파일 읽기  
df_hw = pd.read_csv("/content/drive/MyDrive/머신러닝/hw_data.csv")  
print(df_hw)  
  
# 독립 변수 '키'  
X_hw = df_hw[['키']]  
# 종속 변수 '몸무게'  
y_hw = df_hw['몸무게']  
  
print("X (키):\n", X_hw.head())  
print("y (몸무게):\n", y_hw.head())
```

키에 따른 몸무게 예측

● 모델 학습

모델 학습

```
# 선형 회귀 모델 초기화  
model_hw = LinearRegression()  
  
# 모델 학습  
model_hw.fit(X_hw, y_hw)  
  
print("키-몸무게 모델 학습이 완료되었습니다.")
```

키-몸무게 모델 학습이 완료되었습니다.

키에 따른 몸무게 예측

● 모델 학습 및 예측 수행

▼ 모델 학습

[3]
✓ 0초

```
# 선형 회귀 모델 초기화  
model_hw = LinearRegression()  
  
# 모델 학습  
model_hw.fit(X_hw, y_hw)  
  
print("키-몸무게 모델 학습이 완료되었습니다.")
```

▼ 키-몸무게 모델 학습이 완료되었습니다.

▼ 예측 수행

[4]
✓ 0초

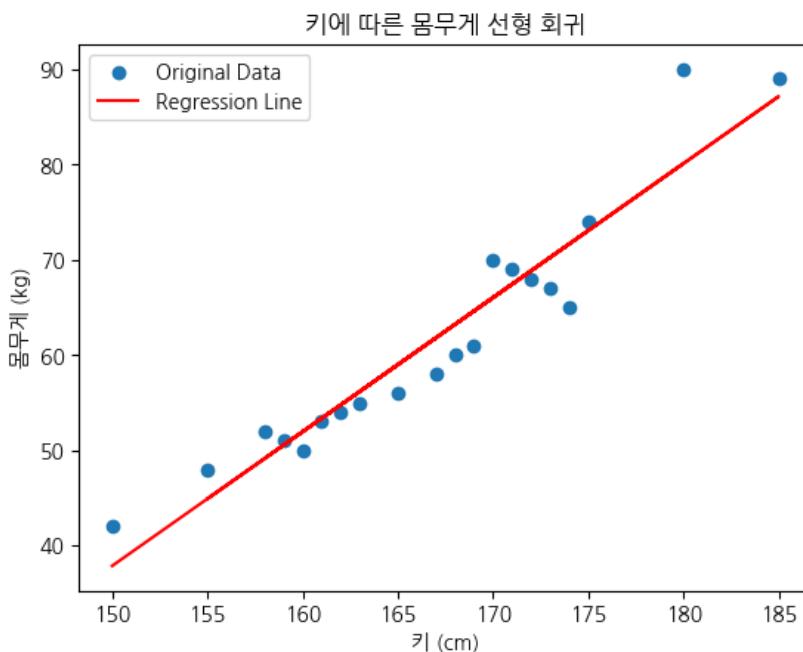
```
pred_hw = model_hw.predict(X_hw)  
  
print("예측 결과 (일부):\n", pred_hw[:5])
```

▼ 예측 결과 (일부):
[37.87287743 51.95425581 66.03563419 73.07632338 58.994945]

키에 따른 몸무게 예측

● 모델 시각화

```
plt.scatter(X_hw, y_hw, label='Original Data')
plt.plot(X_hw, pred_hw, color='red', label='Regression Line')
plt.rc('font', family='NanumGothic') # 한글 글꼴 적용
plt.title('키에 따른 몸무게 선형 회귀')
plt.xlabel('키 (cm)')
plt.ylabel('몸무게 (kg)')
plt.legend()
plt.show()
```



키에 따른 봄무게 예측

● 모델 시각화 – 한글 글꼴 설정

```
import matplotlib.pyplot as plt
import matplotlib.font_manager as fm

# Install NanumGothic font
!sudo apt-get install -y fonts-nanum
!sudo fc-cache -fv

# Add font to font manager
fm.fontManager.addfont('/usr/share/fonts/truetype/nanum/NanumGothic.ttf')

# Set the default font to NanumGothic
plt.rc('font', family='NanumGothic')

print("NanumGothic font installed and configured.")
```

키에 따른 몸무게 예측

● 성능 평가 및 몸무게 예측

직선의 방정식 (기울기 및 y절편)

```
print(f"기울기 (Coefficient): {model_hw.coef_[0]:.4f}")
print(f"절편 (Intercept): {model_hw.intercept_:.4f}")
```

```
기울기 (Coefficient): 1.4081
절편 (Intercept): -173.3478
```

성능 평가

```
# 키-몸무게 모델 (model_hw)에 대한 성능 평가 (전체 데이터)
print(f"키-몸무게 모델 (전체 데이터) R2 Score: {model_hw.score(X_hw, y_hw):.4f}")
```

```
키-몸무게 모델 (전체 데이터) R2 Score: 0.9135
```

키가 170cm일 때 예상 몸무게

```
print(model_hw.predict([[170]]))
```

```
[66.03563419]
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names.
warnings.warn(
```

키에 따른 몸무게 예측

● 데이터 세트 분리

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# csv 파일 읽기
df_hw = pd.read_csv("/content/drive/MyDrive/머신러닝/hw_data.csv")
print(df_hw.head())

# 변수 선택
# X = df_hw[['키']]
X = df_hw.drop("몸무게", axis=1)
y = df_hw['몸무게']

print("X (키):\n", X.head())
print("y (몸무게):\n", y.head())
```

키 몸무게

0	150	42
1	160	50
2	170	70
3	175	65
4	165	60

X (키):

키

0	150
1	160
2	170
3	175
4	165

키에 따른 몸무게 예측

● 데이터 분할

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y,
                                                    test_size=0.2, random_state=12) # 테스트 세트 크기를 20%로 변경

print(X_train.shape, X_test.shape, y_train.shape, y_test.shape)

# 훈련 세트와 테스트 세트 확인
print("훈련 세트 X_train:", X_train.head())
print("\n훈련 세트 y_train:", y_train.head())
print("\n테스트 세트 X_test:", X_test.head())
print("\n테스트 세트 y_test:", y_test.head())

(16, 1) (4, 1) (16,) (4,)
훈련 세트 X_train:
    몸무게
14    58
16    69
10    54
8     65
5     48

훈련 세트 y_train:
    14    58
    16    69
    10    54
    8     65
    5     48
Name: 몸무게, dtype: int64
```

키에 따른 몸무게 예측

● 모델 학습

분리된 데이터를 통한 모델 학습

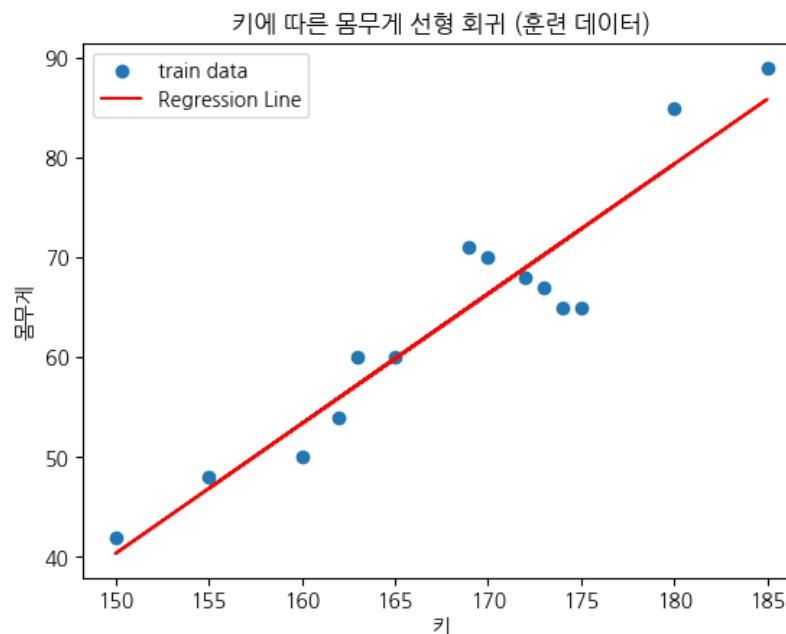
```
from sklearn.linear_model import LinearRegression  
  
# 선형 회귀 모델 객체 생성  
model = LinearRegression()  
  
# 모델 학습  
model.fit(X_train, y_train)  
  
print("모델 학습이 완료되었습니다.")
```

모델 학습이 완료되었습니다.

키에 따른 몸무게 예측

● 모델 시각화 – 훈련 데이터

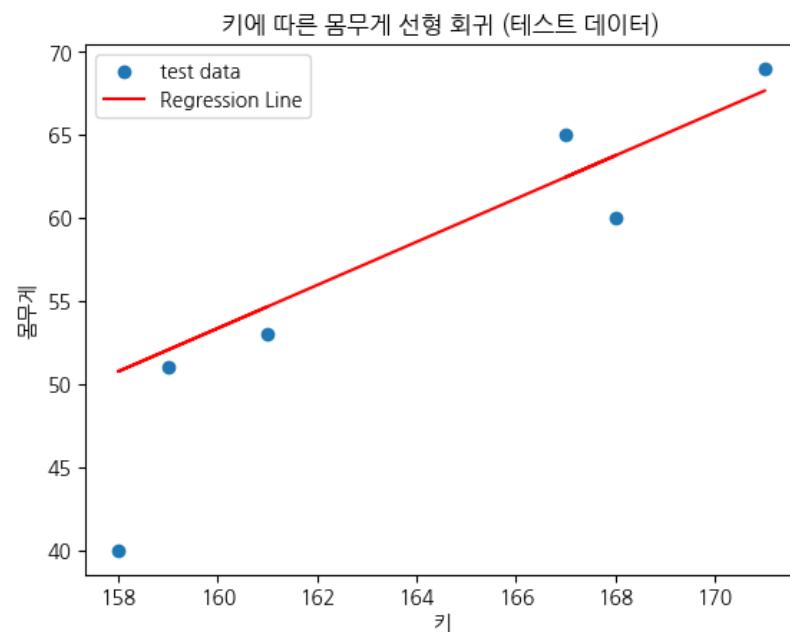
```
plt.scatter(X_train, y_train, label='train data')
plt.plot(X_train, model.predict(X_train), color='red', label='Regression Line')
plt.rc('font', family='NanumGothic') # 한글 글꼴 적용
plt.title('키에 따른 몸무게 선형 회귀 (훈련 데이터)')
plt.xlabel('키')
plt.ylabel('몸무게')
plt.legend()
plt.show()
```



키에 따른 몸무게 예측

● 모델 시각화 – 테스트 데이터

```
plt.scatter(X_test, y_test, label='test data') # 레이블 설정  
plt.plot(X_test, model.predict(X_test), color='red', label='Regression Line')  
plt.rc('font', family='NanumGothic') # 한글 글꼴 적용  
plt.title('키에 따른 몸무게 선형 회귀 (테스트 데이터)') # 제목 설정  
plt.xlabel('키') # 라벨 설정  
plt.ylabel('몸무게') # 라벨 설정  
plt.legend()  
plt.show()
```



키에 따른 몸무게 예측

● 모델 성능 평가 및 예측

모델의 성능 평가

```
model.score(X_train, y_train) #훈련세트를 통한 모델 평가
```

```
0.896152355234782
```

```
model.score(X_test, y_test) #테스트 세트를 통한 모델 평가
```

```
0.7439908761659375
```

키가 170cm일때 예상 몸무게

```
print(model.predict([[170]]))
```

```
[66.36418161]
/usr/local/lib/python3.12/dist-packages/sklearn/utils/validation.py:2739: UserWarning: X does not have valid feature names - [None]
  warnings.warn("X does not have valid feature names - %s" % X_.get_feature_names_out(), UserWarning)
```

딥러닝(Deep Learning)

● 딥러닝이란?

딥러닝(Deep Learning)은 인공 신경망을 기반으로 하는 머신러닝의 한 분야입니다.

사람의 뇌가 작동하는 방식과 유사하게, 여러 계층(layer)의 신경망을 통해 데이터를 학습하고 복잡한 패턴을 인식하여 예측이나 분류 등의 작업을 수행합니다.

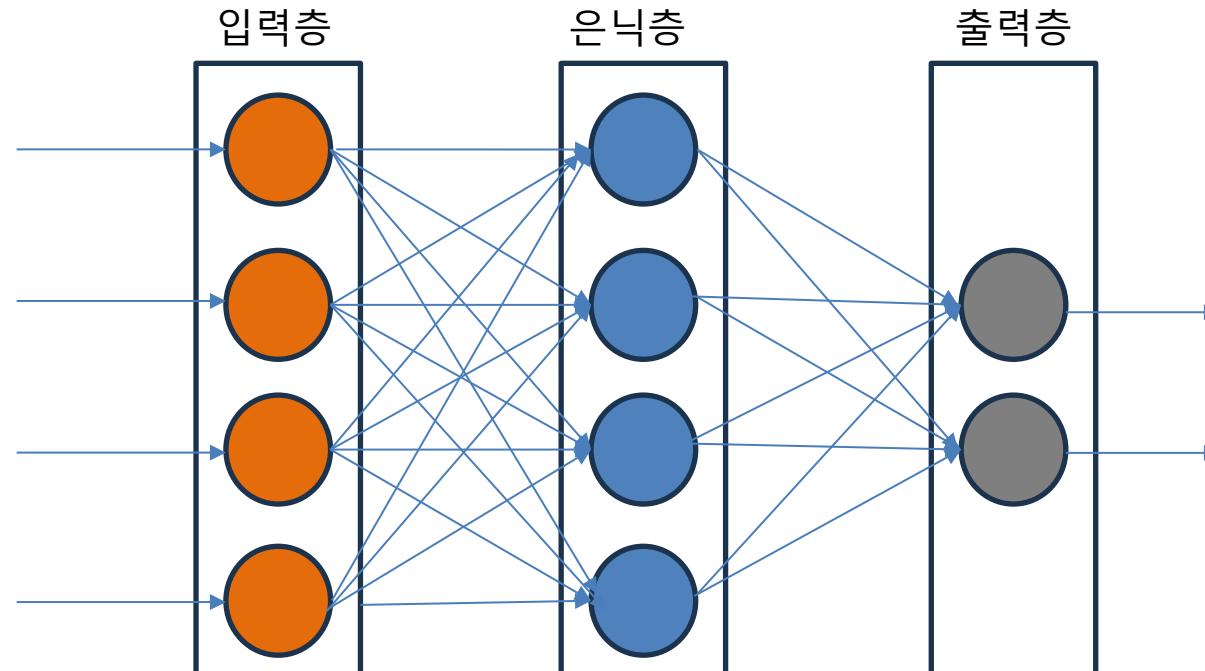
입력층 → 은닉층 → 은닉층 → 출력층



딥러닝(Deep Learning)

● 인공 신경망

- **입력층(input layer)** – 데이터를 입력받는 층
- **은닉층(hidden layer)** – 입력층에서 들어온 데이터가 여러 신호로 바뀌어 출력층까지 전달됨
- **출력층(output layer)** – 출력층에 어떤 값이 전달되었는냐에 따라서 예측 값이 결정됨



딥러닝(Deep Learning)

● 신호 전달 원리

인공신경망에서는 단순하게 신호를 전달해 주는 것이 아니라 **신호 세기를 변경해서 전달**합니다.

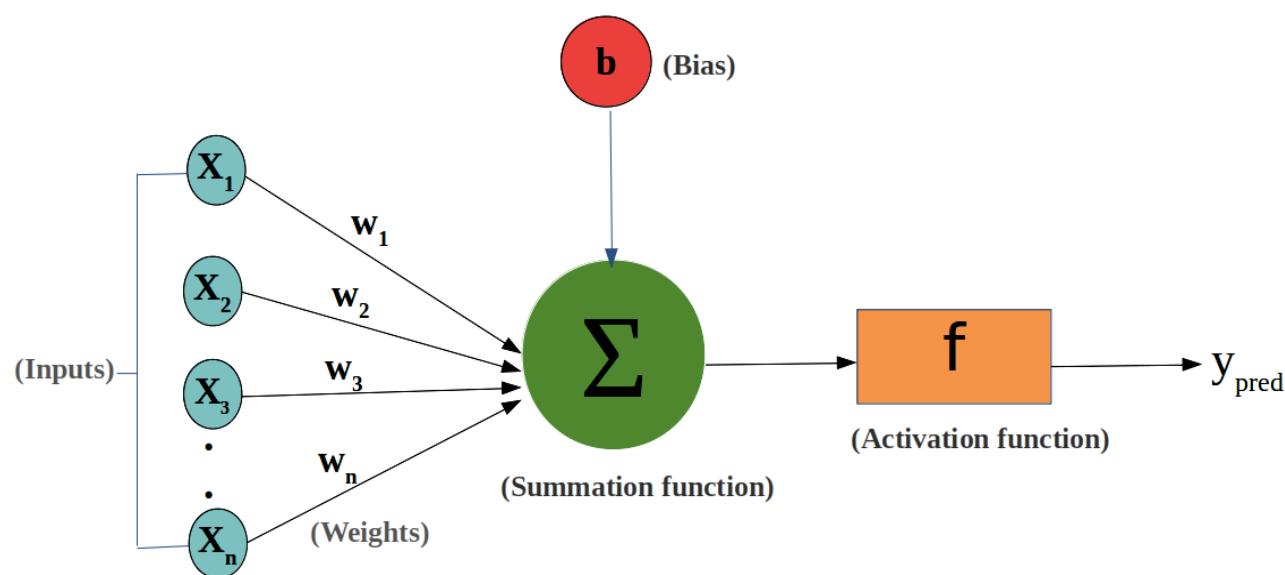
신경망에서 중요한 용어 중 **가중치(weight)**와 **편향(bias)**이 있는데 이들이 바로 신호 세기를 변경하는 데 사용된다.

한 뉴런에서 다음 뉴런으로 전달되는 신호 세기는 가중치와 편향에 의해 결정됨

가중치란 그 값이 얼마나 중요한지 그렇지 않은지를 표현하기 위한 도구이고, **편향**은 한쪽으로 치우치는 값을 더할때 사용함

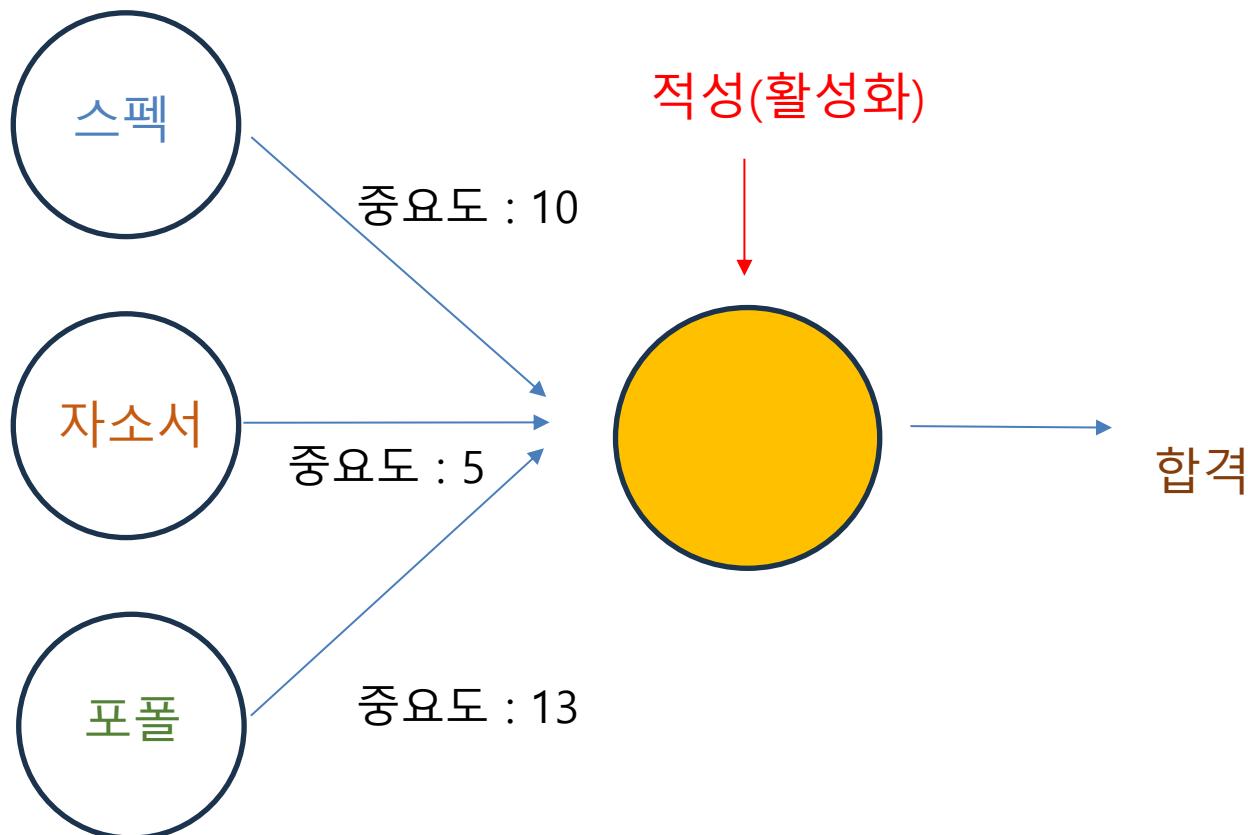
딥러닝(Deep Learning)

● 가중치와 편향



딥러닝(Deep Learning)

- 가중치와 편향



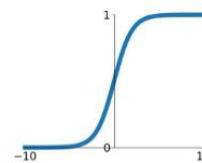
딥러닝(Deep Learning)

● 활성화 함수

- 뉴런을 활성화 할 필요가 있는지 없는지 결정하는데 도움을 주는 함수
- 비선형 문제를 해결하는데 중요한 역할을 한다

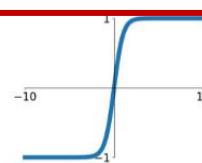
Sigmoid

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



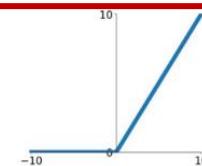
tanh

$$\tanh(x)$$



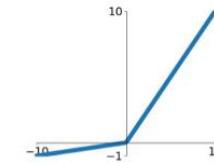
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

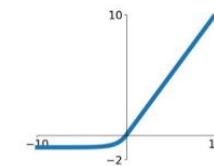


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



딥러닝(Deep Learning)

● Softmax(소프트 맥스) 함수

소프트 맥스 함수는 사실 활성화 함수는 아니고, 인공신경망의 마지막 부분 출력층에서 주로 사용되는 함수이다.

최종 결과값을 정규화하는데 사용하는 함수가 소프트 맥스 함수이다.

이 소프트 맥스 함수를 사용하면 예를 들어 남자와 여자를 구분할 때 2개의 노드가 100% 확률. 60%, 40%라면 총합이 1(0.6, 0.4)이 되도록 보여준다.

머신러닝과 딥러닝 비교

⚙️ 머신러닝 vs 딥러닝 한눈에 비교

구분	머신러닝	딥러닝
데이터	적어도 가능	많이 필요
특징 추출	사람이 설계	자동
계산량	적음	매우 큼
성능	보통	매우 높음

🎯 한 줄 요약

- **머신러닝:** 데이터로 규칙을 배우는 AI
- **딥러닝:** 신경망을 깊게 쌓아 더 똑똑하게 만든 머신러닝

딥러닝(Deep Learning)

● 손글씨 예측 문제

MNIST 데이터셋의 이미지 하나를 시각화해줘

이미지 데이터를 신경망에 맞게 전처리해줘

첫 번째 딥러닝 모델을 만들어줘

MNIST에 대해 설명해 줘

+

Gemini 2.5 Flash ▾ ▶

MNIST

미국 국립표준기술연구소에서 만든 손글씨 숫자 이미지를 혼합해서 만든 데이터 셋입니다.
머신러닝 및 딥러닝 분야에서 가장 널리 사용되는 데이터셋 중 하나입니다. 주로 손으로 쓴 숫자 이미지로 구성되어 있으며, 이미지 분류 모델의 성능을 평가하는 데 사용됩니다.

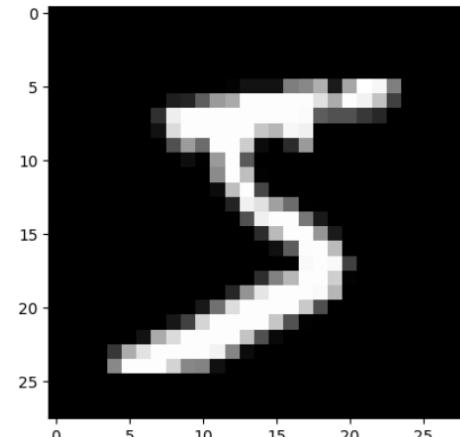
A 4x4 grid of handwritten digits from the MNIST dataset. The digits are arranged in a 4x4 pattern. The digits are mostly black on a white background, with some variations in stroke thickness and orientation. The digits include 0s, 1s, 2s, 3s, 4s, 5s, 6s, 7s, 8s, and 9s.

0	0	0	0
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9

딥러닝(Deep Learning)

● MNIST의 주요 특징

- **구성:** 0부터 9까지의 손글씨 숫자 이미지로 이루어져 있습니다.
- **데이터 양:** 훈련 세트에 60,000개의 이미지, 테스트 세트에 10,000개의 이미지가 포함되어 있습니다.
- **이미지 크기:** 각 이미지는 28x28 픽셀의 흑백(회색조) 이미지입니다.
- **용도:** 주로 이미지 분류, 특히 숫자 인식 분야에서 딥러닝 모델의 기본 성능 테스트 및 벤치마킹을 위해 활용됩니다. 간단하면서도 충분히 복잡하여 초보자들이 딥러닝을 학습하고 다양한 모델을 시도해보기 좋은 데이터셋입니다.



딥러닝(Deep Learning)

● 라이브러리 불러오기 및 데이터 준비

```
from keras.datasets import mnist
from keras.models import Sequential
from keras.layers import Dense
import numpy as np
import matplotlib.pyplot as plt

# 데이터 준비
(X_train, y_train), (X_test, y_test) = mnist.load_data()
```

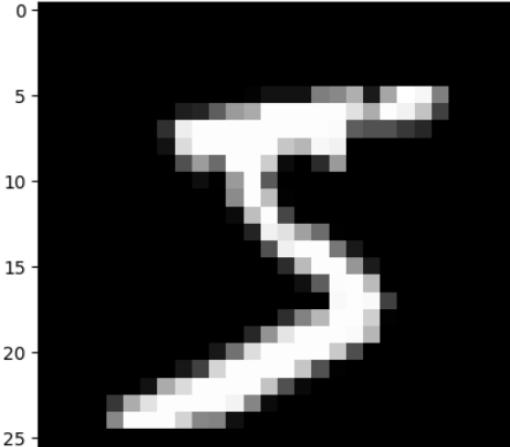
X_train.shape

(60000, 28, 28)

딥러닝(Deep Learning)

● 학습용 첫 데이터

학습용 첫 데이터 보기

```
X_train[0]  
ndarray (28, 28) show data  
  
  
print(y_train[0])  
5  
  
plt.imshow(X_train[0], cmap="gray") # 숫자를 이미지로 출력  
... <matplotlib.image.AxesImage at 0x7d006afc9160>  

```

딥러닝(Deep Learning)

● 데이터 전처리

mnist 데이터셋에서 데이터 X의 형태 바꾸기

(60000, 28, 28) -> (60000, 784)

인공 신경망의 입력층에 데이터를 넣을 때는 한 줄(행)로 만들어 넣어야 함
reshape() 사용

데이터 정규화

0~255(색상) 사이의 값을 0~1 사이의 값으로 바꾸기

RGB(255,255,255) - 흰색

RGB(0,0,0) - 검정색

RGB(255, 0, 0) - 빨간색

데이터를 255로 나눔 - 실수형으로 바뀜

```
X_train = X_train.reshape(60000, 784) #1차원 리스트 형태  
X_test = X_test.reshape(10000, 784)
```

```
X_train = X_train / 255  
X_test = X_test / 255
```

딥러닝(Deep Learning)

● 데이터 전처리

▶ X_train[0]
...
0., 0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0.01176471, 0.07058824, 0.07058824,
0.07058824, 0.49411765, 0.53333333, 0.68627451, 0.10196078,
0.65098039, 1., 0.96862745, 0.49803922, 0., ,
0., 0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., 0., ,
0., 0.11764706, 0.14117647, 0.36862745, 0.60392157,
0.66666667, 0.99215686, 0.99215686, 0.99215686, 0.99215686,
0.99215686, 0.88235294, 0.6745098, 0.99215686, 0.94901961,
0.76470588, 0.25098039, 0., 0., 0., ,
0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0.19215686, 0.93333333,
0.99215686, 0.99215686, 0.99215686, 0.99215686, 0.99215686,
0.99215686, 0.99215686, 0.99215686, 0.98431373, 0.36470588,
0.32156863, 0.32156863, 0.21960784, 0.15294118, 0., ,
0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., ,
0., 0.07058824, 0.85882353, 0.99215686, 0.99215686,
0.99215686, 0.99215686, 0.99215686, 0.77647059, 0.71372549,
0.96862745, 0.94509804, 0., 0., 0., ,
0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., ,
0.31372549, 0.61176471, 0.41960784, 0.99215686, 0.99215686,
0.80392157, 0.04313725, 0., 0.16862745, 0.60392157,
0., 0., 0., 0., 0., ,
0., 0., 0., 0., 0., ,
0.05400105, 0., 0., 0., 0., ,

딥러닝(Deep Learning)

● 모델 만들기

모델 만들기

입력층 - 은닉층(256) - 은닉층2(128) - 은닉층3(64) - 출력층(10)

```
# 모델 객체 생성
model = Sequential()

model.add(Dense(units=256, input_dim=784, activation="relu")) # 입력층
model.add(Dense(units=128, activation="relu")) # 은닉층1
model.add(Dense(units=64, activation="relu")) # 은닉층2
model.add(Dense(units=10, activation="softmax")) # 출력층
model.summary() # 요약
```

```
/usr/local/lib/python3.12/dist-packages/keras/src/layers/core/dense.py:93: UserWarning: Do
super().__init__(activity_regularizer=activity_regularizer, **kwargs)
Model: "sequential"
```

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 256)	200,960
dense_1 (Dense)	(None, 128)	32,896
dense_2 (Dense)	(None, 64)	8,256
dense_3 (Dense)	(None, 10)	650

```
Total params: 242,762 (948.29 KB)
Trainable params: 242,762 (948.29 KB)
Non-trainable params: 0 (0.00 B)
```

딥러닝(Deep Learning)

● 모델 학습 및 평가

```
# 모델 컴파일(최적화 함수, 손실함수, 평가 지표)
model.compile(optimizer='adam', loss='sparse_categorical_crossentropy',
               metrics=['accuracy'])
```

```
# 모델 학습(훈련데이터, 정답, 반복횟수, 범위(100개 단위 실행))
model.fit(X_train, y_train, epochs=5, batch_size=100)
```

```
Epoch 1/5
600/600 - 6s 7ms/step - accuracy: 0.8641 - loss: 0.4873
Epoch 2/5
600/600 - 6s 9ms/step - accuracy: 0.9671 - loss: 0.1065
Epoch 3/5
600/600 - 4s 7ms/step - accuracy: 0.9801 - loss: 0.0649
Epoch 4/5
600/600 - 5s 9ms/step - accuracy: 0.9860 - loss: 0.0438
Epoch 5/5
600/600 - 9s 7ms/step - accuracy: 0.9898 - loss: 0.0323
<keras.src.callbacks.history.History at 0x7cc555510770>
```

```
# 모델 평가(테스트 데이터, 정답)
model.evaluate(X_test, y_test)
# [0.08058327436447144, 0.9757999777793884] #[오차, 정확도]
```

```
313/313 - 1s 3ms/step - accuracy: 0.9722 - loss: 0.0931
[0.08058327436447144, 0.9757999777793884]
```

딥러닝(Deep Learning)

● 모델 예측

```
# 숫자 예측  
model.predict(X_test[0].reshape(1, 784))  
# 9.9990523e-01 : 확률이 가장 높으므로 예측한 숫자는 7임
```

```
1/1 ━━━━━━━━ 0s 115ms/step  
array([[2.8114528e-06, 2.9534567e-07, 7.4817581e-06, 2.4367240e-05,  
       1.1021378e-09, 4.5816659e-08, 9.4153353e-13, 9.9990523e-01,  
       1.3505142e-07, 5.9734855e-05]], dtype=float32)
```

```
▶ print(np.argmax(model.predict(X_test[0].reshape(1, 784)))) #7
```

```
# 출력시엔 이차원 형태로 변환  
plt.imshow(X_test[0].reshape(28, 28), cmap='gray')
```

```
... 1/1 ━━━━━━━━ 0s 38ms/step  
7  
<matplotlib.image.AxesImage at 0x7cc552162000>
```

