

- **인공 지능 (Artificial Intelligence)**
 - 인간이 가진 **지적 능력**을 **컴퓨터**를 통해 구현 하는 기술
- **머신 러닝 (Machine Learning)**
 - 데이터와 **알고리즘**을 통해 컴퓨터를 **학습** 시켜 인공 지능의 성능을 향상 시키는 기술
- **딥 러닝 (Deep Learning)**
 - **신경망 (Neural Network)** 을 이용한 머신 러닝 방법.
 - 여러 층으로 이루어진 신경망을 사용.



Machine Learning: A Primer to Laboratory Applications (thermofisher.com)



<https://m.blog.naver.com/pwj6971/221614497987>

머신러닝과 프로그래밍의 차이점

전통적인 프로그래밍 방식은 문제 해결 알고리즘을 만들어서 처리하지만 머신러닝은 많은 수의 데이터를 시스템에 제공하고 어떤 것이 답인지 알려줘서 학습시킨다.

Chat GPT vs Deep Seek

	Chat GPT	Deep Seek
모델 version	GPT-4o, o1	DeepSeek-V3, DeepSeek-R1
사용한 모델	transformer	Mixture of expert (MOE)
Parameter 수	1.8 조	370 억
Open source?	no	yes
focus	광범위한 주제	Technical Mathematics, coding ,,

과적합 : 학습 데이터에만 익숙해져서 새로운 데이터에서는 성능이 떨어지는 것

과소적합 : 모델 성능이 떨어져서 학습 데이터에서도 예측 성능이 나오지 않는 것

Fit : 내부적으로 학습 데이터와 정답 데이터를 가지고 가중치와 편향을 조정한다.

정리

순서

1. 예측값 계산
2. 실제값 - 예측값 해서 오차 계산
3. 오차 \times 입력값 다 더해서 평균 \rightarrow 기울기 바꿀 값
4. 오차만 다 더해서 평균 \rightarrow 편향 바꿀 값
5. 학습률 \times 계산값 해서 기울기, 편향 바꾸기
6. 다시 예측값 계산
7. 반복

꼭 지켜야 하는 순서

계산할 때	순서
오차 계산할 때	예측값 - 실제값
기울기 계산할 때	오차 \times 입력값
편향 계산할 때	오차만 더하기
값 갱신할 때	기존 값 - (학습률 \times 계산값)



기울기와 편향 업데이트 공식

기울기

1. MSE 과정인 예측값 구한 후 오차값을 구해 입력값과 곱한 후 그 수들을 더하여 평균을 구한다.
2. 기존 기울기 - 학습률 * 오차값을 통해 구한 답에 해당하는 값으로 구한 평균이 공식이 기울기를 구하는 공식이다. (부호 주의하기)
3. $W = w - \text{학습률} * (\text{오차값} * \text{입력값의 평균})$

1. 예측 값 계산:

먼저, 예측 값을 계산합니다. 예측 값 y_{pred} 는 가중치 w 와 입력 값 X 에 편향 b 를 더하여 구합니다.

$$y_{\text{pred}} = w \times X + b$$

초기 상태에서 계산한 예측 값은:

- $w = 0.5, b = 1.0$
- 예측 값 y_{pred} 는:

$$y_{\text{pred}} = 0.5 \times [1, 2, 3, 4] + 1.0 = [1.5, 2.0, 2.5, 3.0]$$

2. 오차 계산:

오차는 예측 값과 실제 값의 차이를 의미합니다. 이를 "오차"라고 하며, 각 예측 값에서 실제 값을 빼면 됩니다.

$$\text{오차} = y_{\text{pred}} - y_{\text{true}}$$

실제 값 y_{true} 는 주어진 출력 값 $[3, 6, 7, 8]$ 입니다. 따라서 오차는 다음과 같습니다:

- $y_{\text{pred}} = [1.5, 2.0, 2.5, 3.0]$
- $y_{\text{true}} = [3, 6, 7, 8]$

각각의 오차는:

- $1.5 - 3 = -1.5$
- $2.0 - 6 = -4.0$
- $2.5 - 7 = -4.5$
- $3.0 - 8 = -5.0$

3-1. 가중치 w 업데이트:

가중치 w 는 오차와 입력 값 X 를 곱한 후 평균을 구한 값을 사용하여 업데이트합니다. 이를 경사하강법 공식을 통해 업데이트합니다.

경사하강법 공식은:

$$w = w - \alpha \times \frac{1}{n} \sum_{i=1}^n (\text{오차}) \times X_i$$

오차와 입력 값을 곱한 값을 모두 더해 평균을 구한 후, 이를 학습률 α 에 곱해서 기존 가중치에서 빼주면 됩니다.

먼저 오차와 입력 값을 곱합니다:

$$\begin{aligned} \text{오차} \times X &= [-1.5 \times 1, -4.0 \times 2, -4.5 \times 3, -5.0 \times 4] \\ &= [-1.5, -8.0, -13.5, -20.0] \end{aligned}$$

이 값을 모두 더한 후 평균을 구합니다:

$$\sum = -1.5 - 8.0 - 13.5 - 20.0 = -43.0$$

평균은:

$$\frac{-43.0}{4} = -10.75$$

가중치 w 를 업데이트합니다:

$$w = 0.5 - 0.01 \times (-10.75) = 0.5 + 0.1075 = 0.6075$$

편향

1. 기울기와 동일하게 오차값을 구하지만 입력값을 곱한 후 평균을 구하는 것이 아닌 오차값의 평균을 구한다.
2. $B = b - \text{학습률} \times (\text{오차값의 평균})$