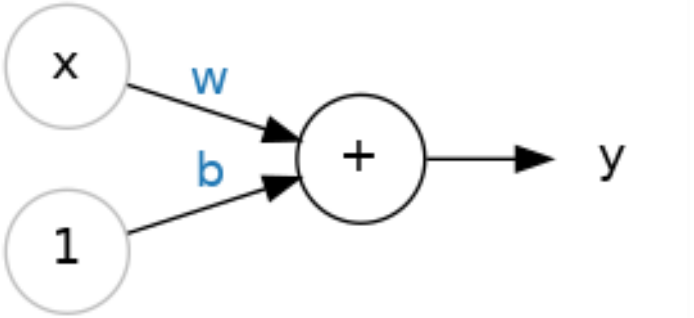
Machine Learning

2022. 4. 1

정 준 수 Ph.D

A Single Neuron

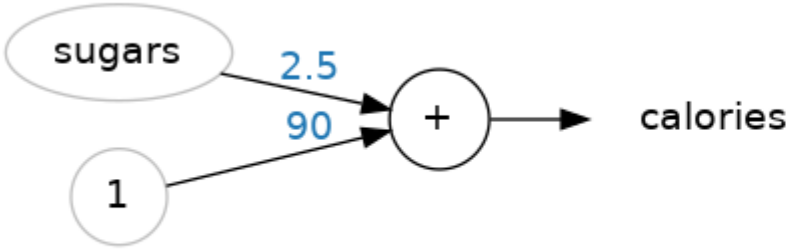
The individual neuron. As a diagram, a neuron (or unit) with one input looks like:



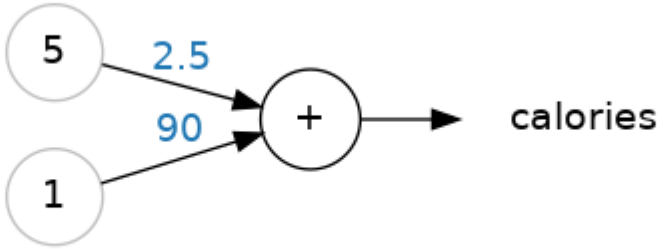
The Linear Unit: **y=wx+b**

선형 모델의 예제

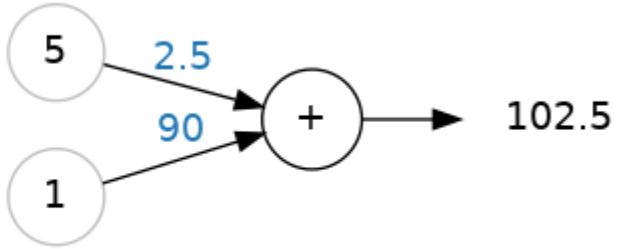
Training a model with 'sugars' (grams of sugars per serving) as input and 'calories' (calories per serving) as output, we might find the bias is b=90 and the weight is w=2.5. We could estimate the calorie content of a cereal with 5 grams of sugar per serving like this:



w=2.5, b=90



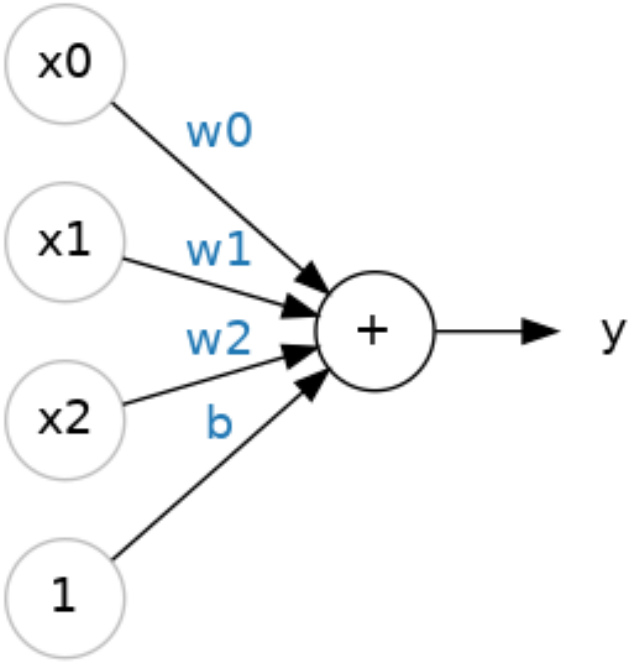
Input sugars=5



Multiply inputs by weights and sum.

calories=2.5×5+90=102.5

Multiple Inputs



this neuron y=w0x0+w1x1+w2x2+b 예를 들어
three features as input ('sugars', 'fiber', and 'protein')

Linear Units을 Keras의 Dense 표현하면

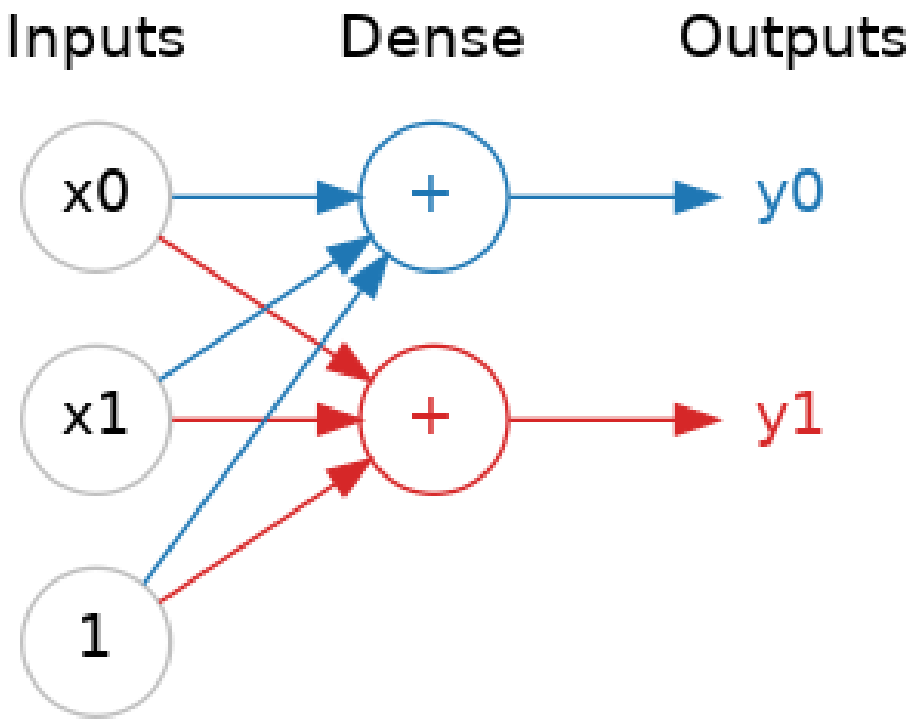
```
from tensorflow import keras
from tensorflow.keras import layers

# Create a network with 1 linear unit
model = keras.Sequential([
    layers.Dense(units=1, input_shape=[3])
])
```

Input [3] 개, Output [1] 개

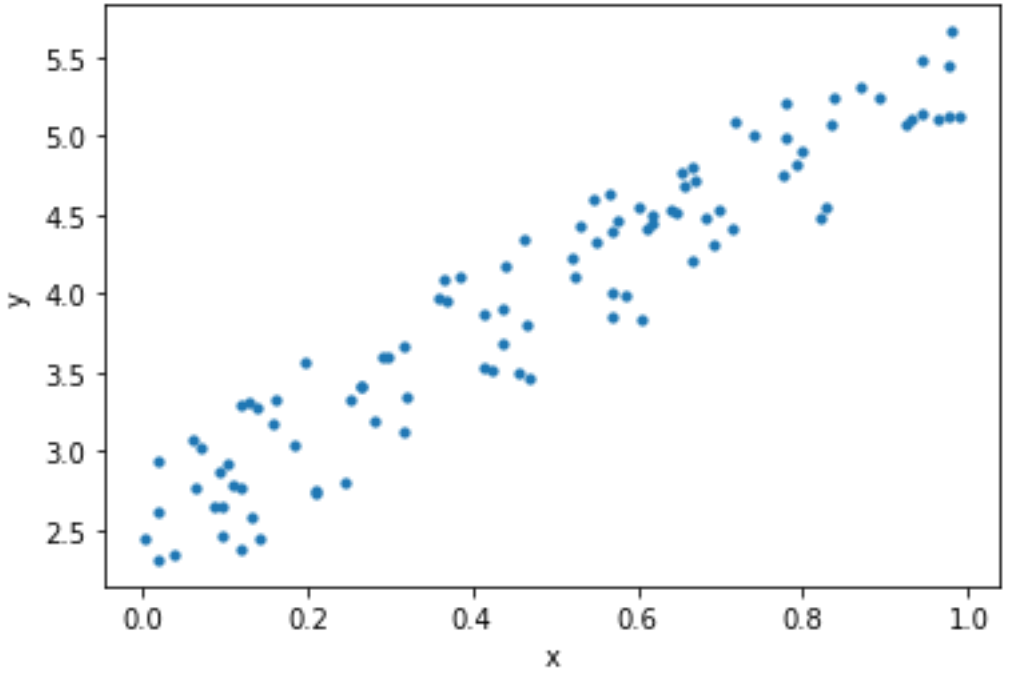
Layers

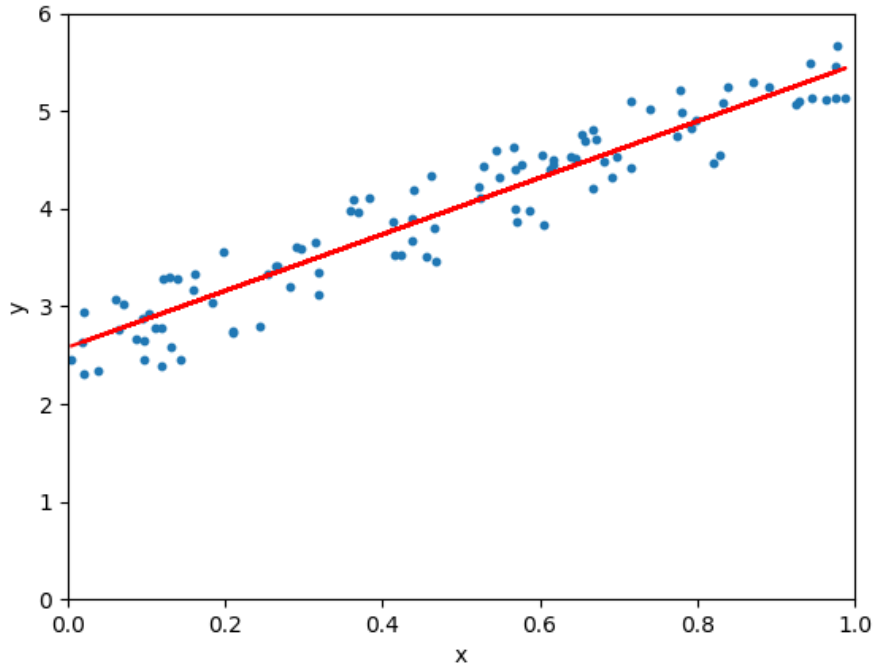
Neural networks typically organize their neurons into layers. When we collect together linear units having a common set of inputs we get a dense layer.



왼쪽 그림은 Layers 가 몇 개일까요?

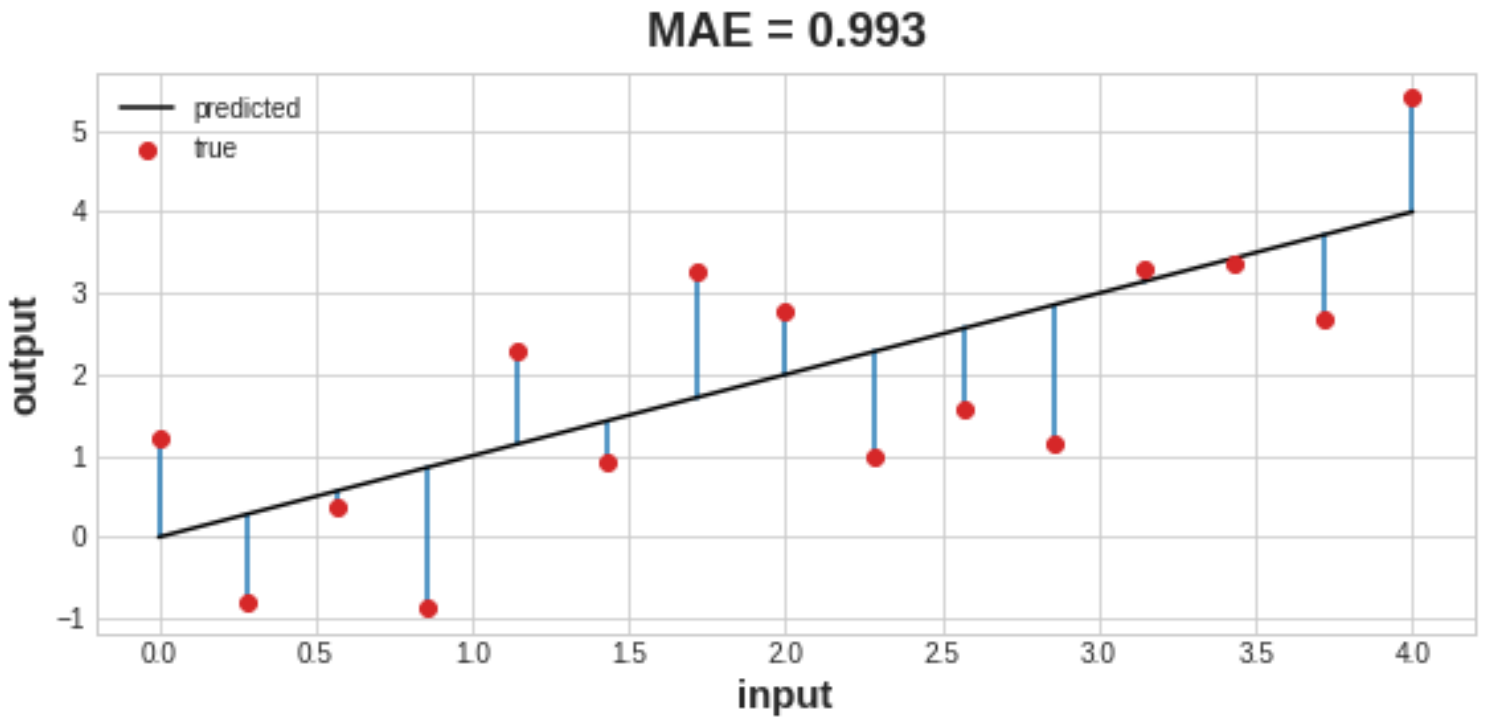
회기(Regression) 예측분석



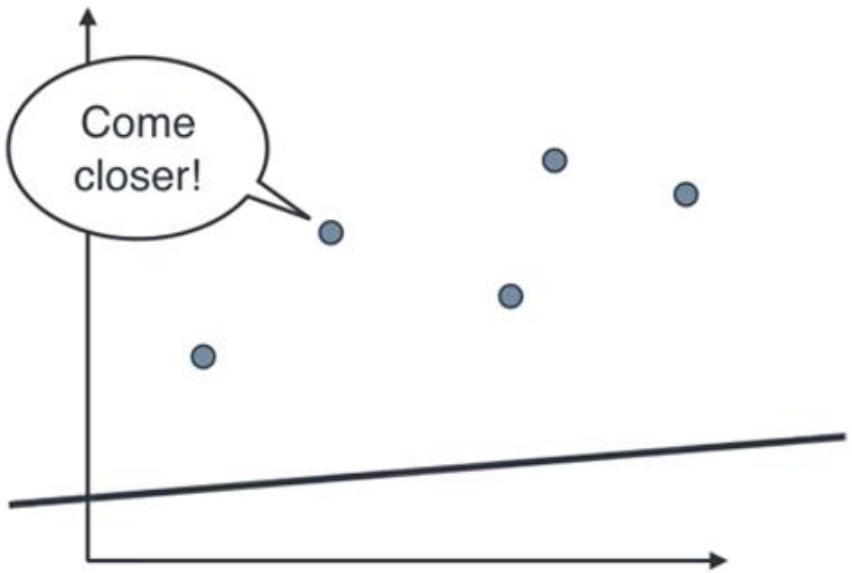


The Loss Function

mean absolute error or MAE



Linear regression algorithm



Step 1: Start with a random line

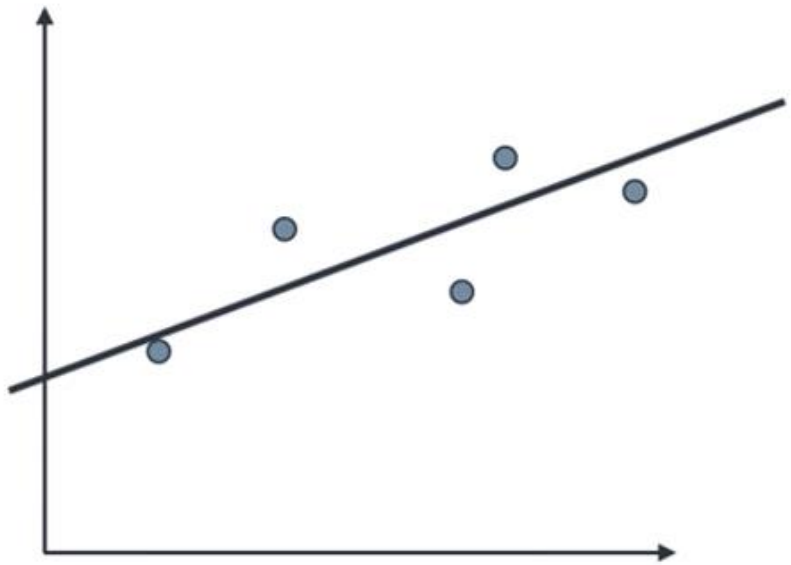
Step 2: Pick a large number. 1000
(number of repetitions, or epochs)

Step 3: Pick a small number. 0.01
(learning rate)

Step 4: (repeat 1000 times)
-Pick random point
- Add (learning rate)x(vertical distance) x (horizontal distance) to **slope**
- Add (learning rate)x(vertical distance) to **y-intercept**

https://www.youtube.com/watch?v=wYPUhge9w5c

Linear regression algorithm



Step 1: Start with a random line

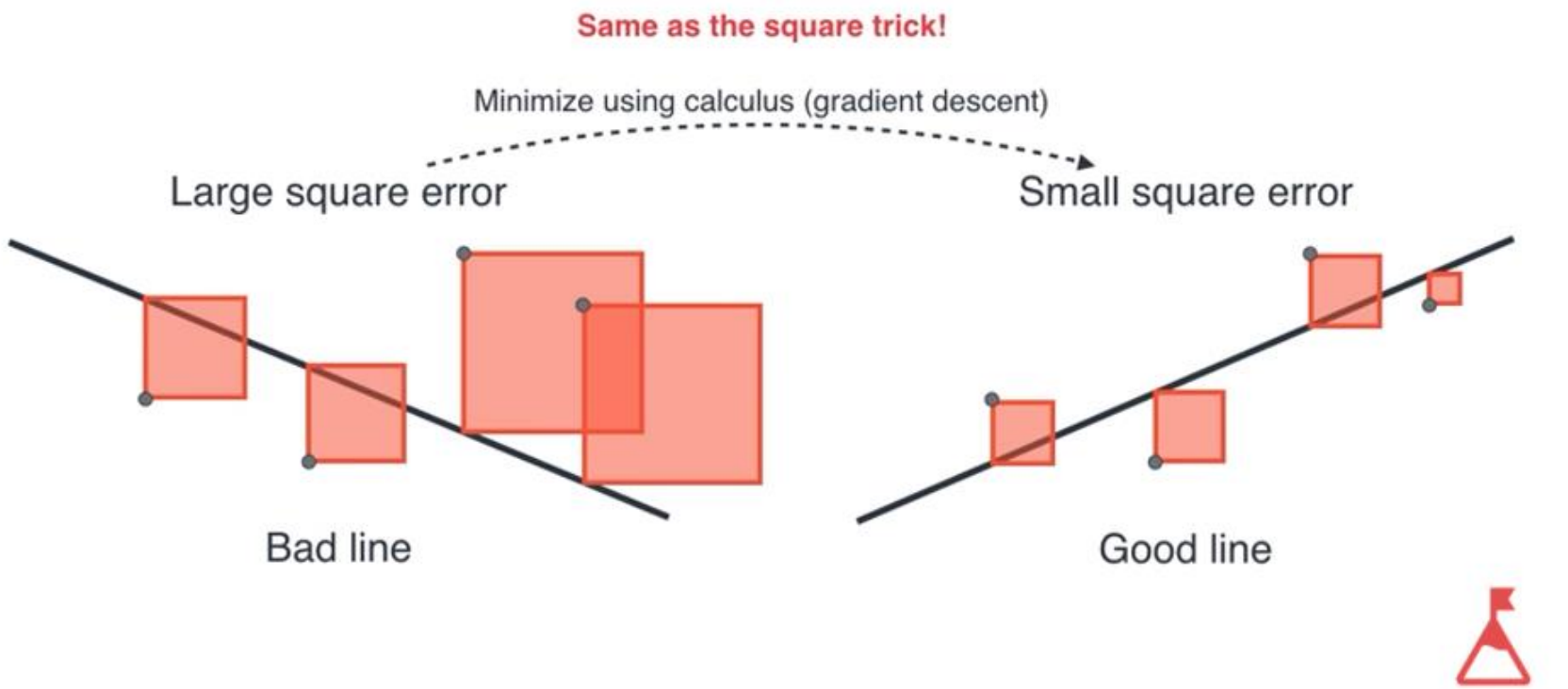
Step 2: Pick a large number. 1000
(number of repetitions, or epochs)

Step 3: Pick a small number. 0.01
(learning rate)

Step 4: (repeat 1000 times)
-Pick random point
- Add (learning rate)x(vertical distance) x (horizontal distance) to **slope**
- Add (learning rate)x(vertical distance) to **y-intercept**

https://www.youtube.com/watch?v=wYPUhge9w5c

Square error

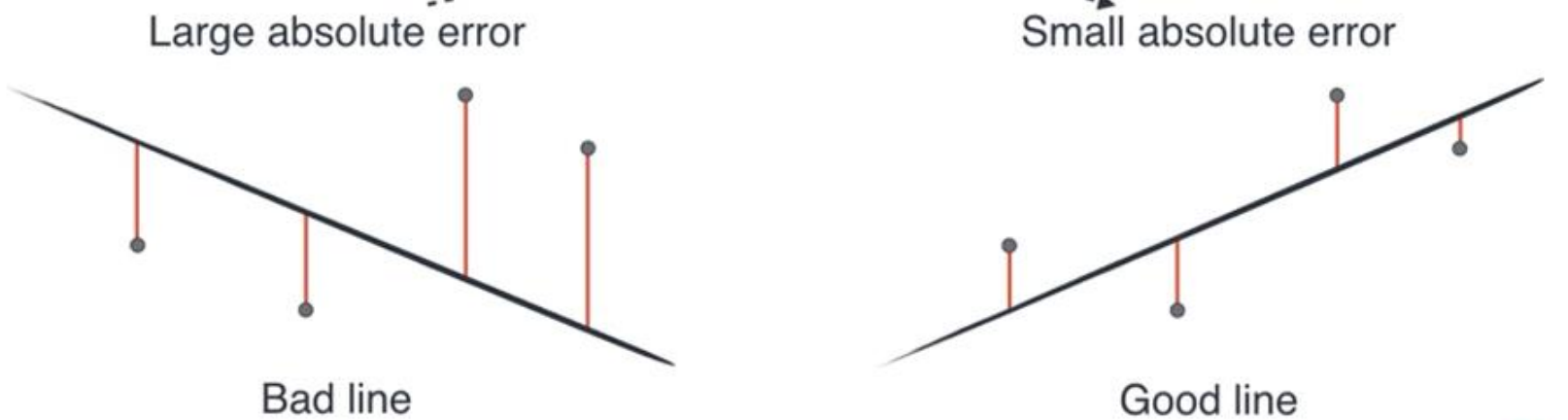


https://www.youtube.com/watch?v=wYPUhge9w5c

Absolute error

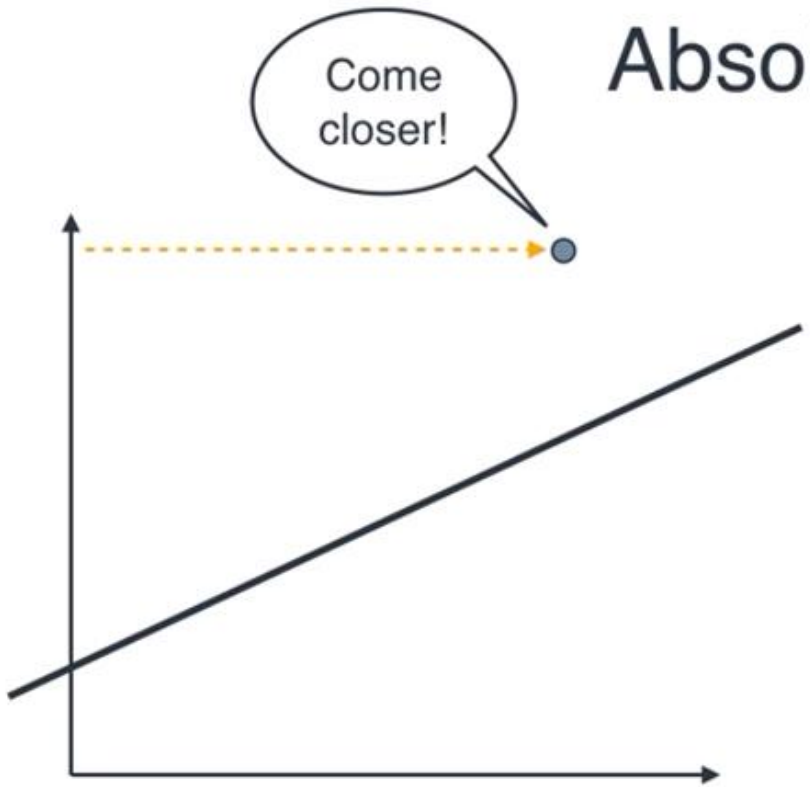
Develop an absolute trick!

Minimize using calculus (gradient descent)



https://www.youtube.com/watch?v=wYPUhge9w5c

Absolute trick



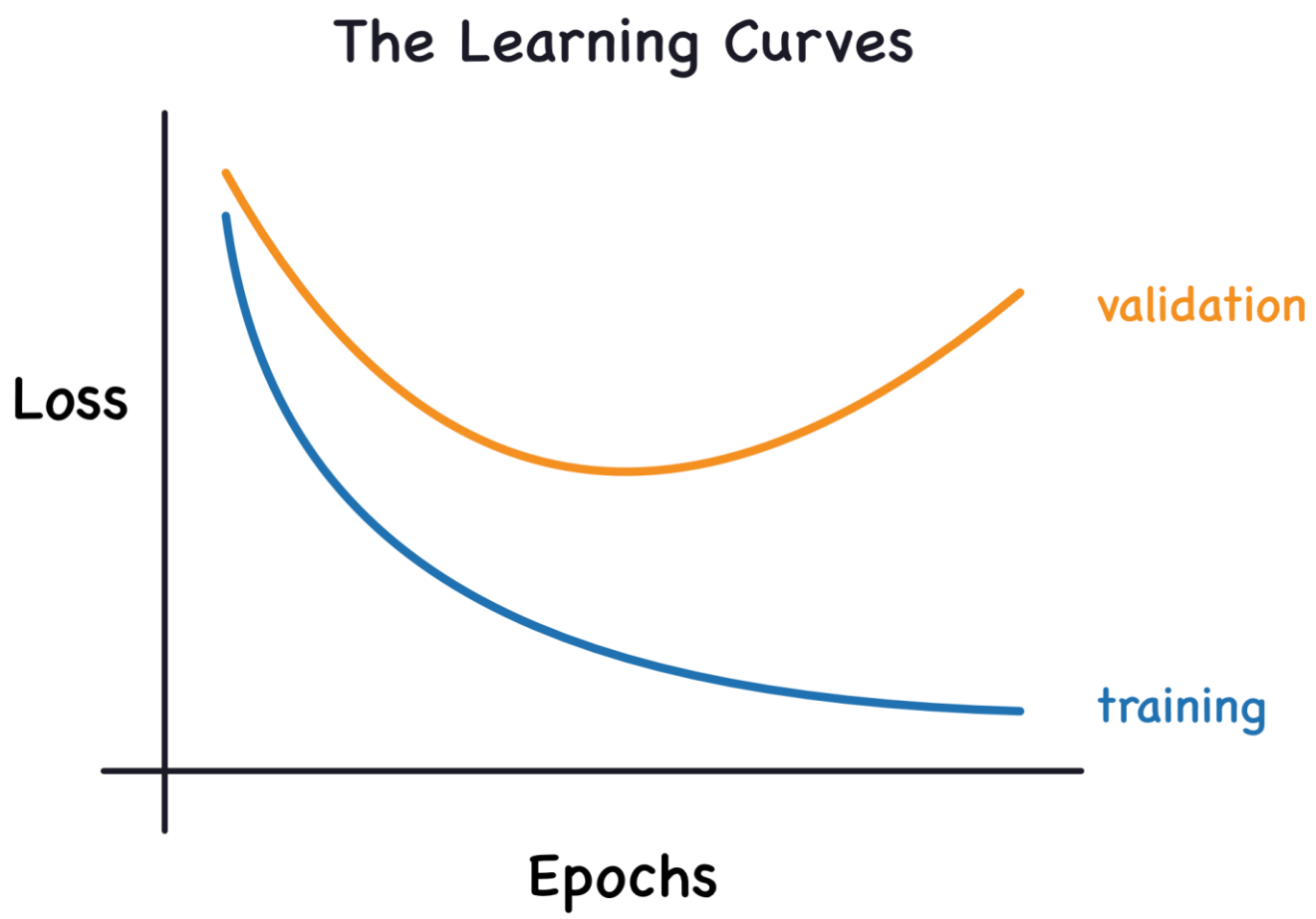
Step 1:
Pick a small number (learning rate)

Step 2:
- If point is above the line:
  - Add (learning rate) x (horizontal distance) to **slope**
  - Add (learning rate) to **y-intercept**
- If point is below the line:
  - Subtract (learning rate) x (horizontal distance) to **slope**
  - Subtract (learning rate) to **y-intercept**

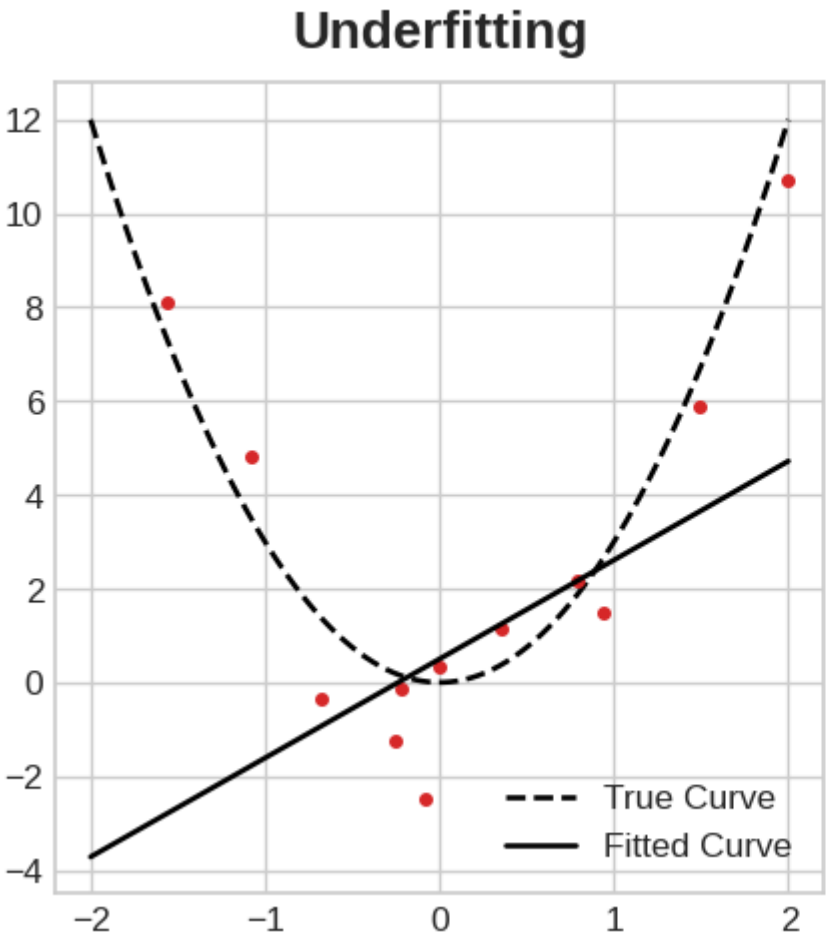
https://www.youtube.com/watch?v=wYPUhge9w5c

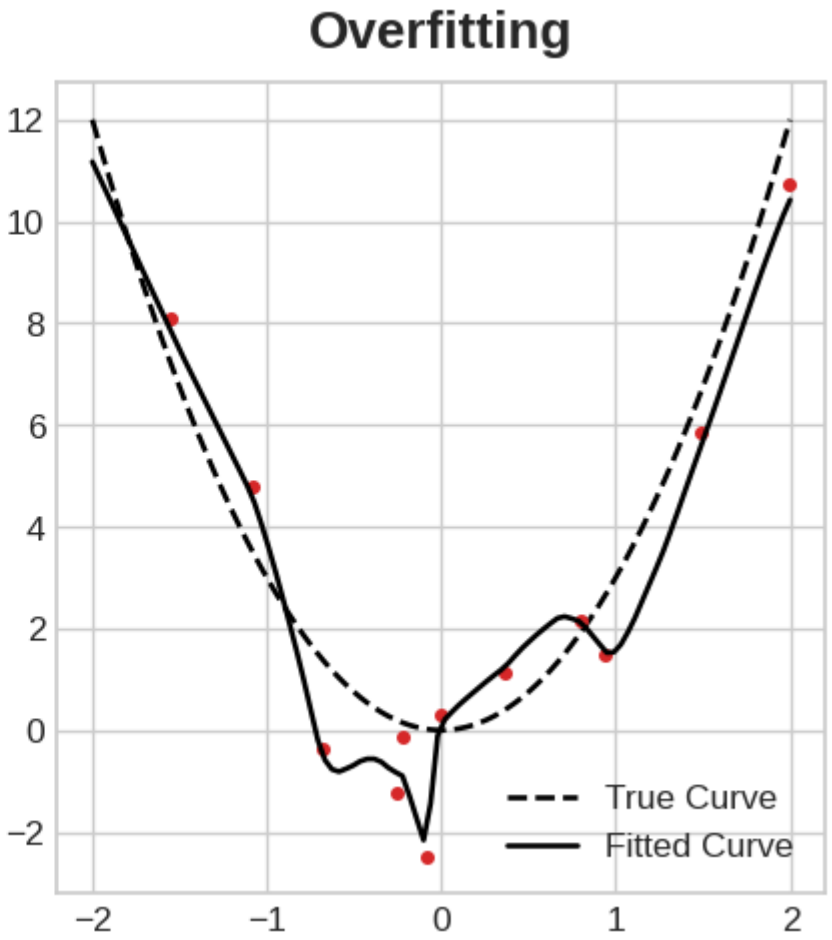
Stochastic Gradient Descent (The Optimizer)

Learning Curves

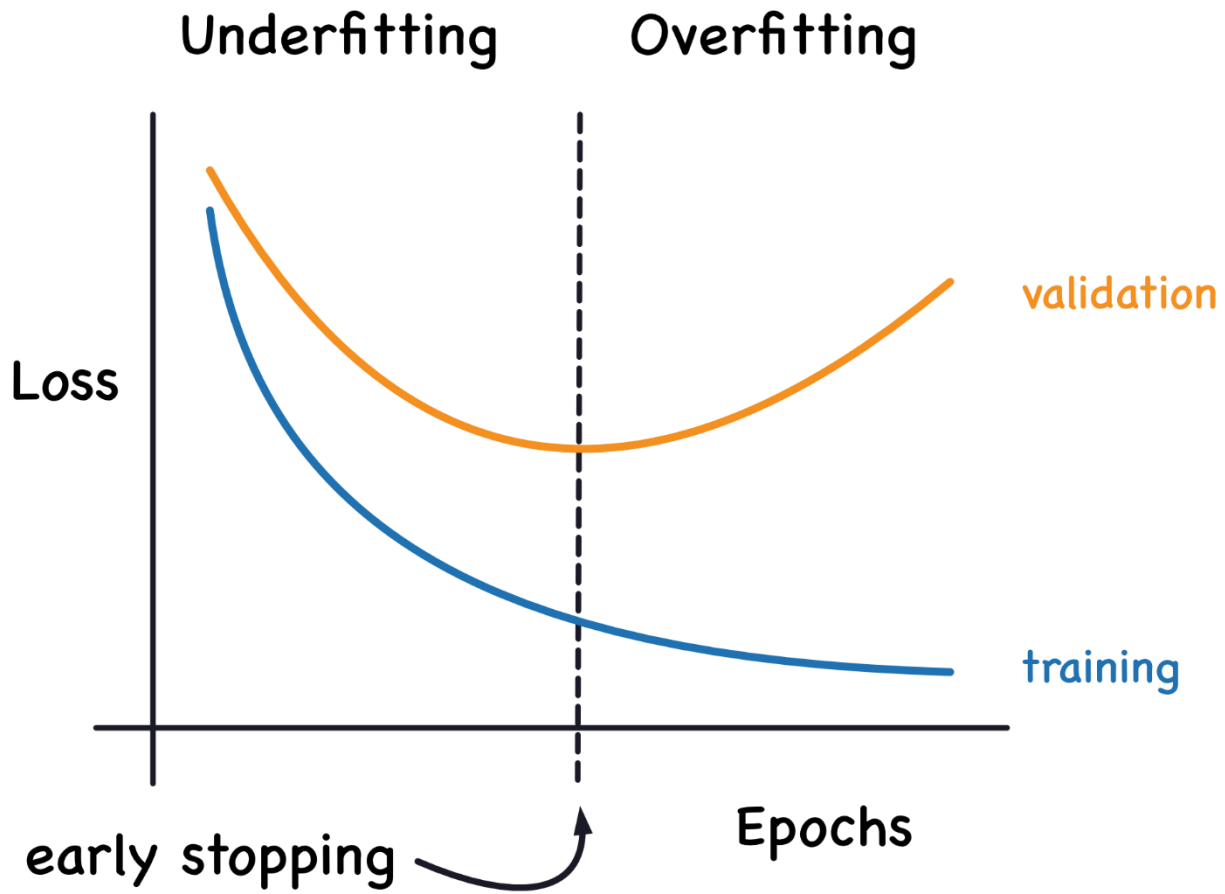


Overfitting and Underfitting

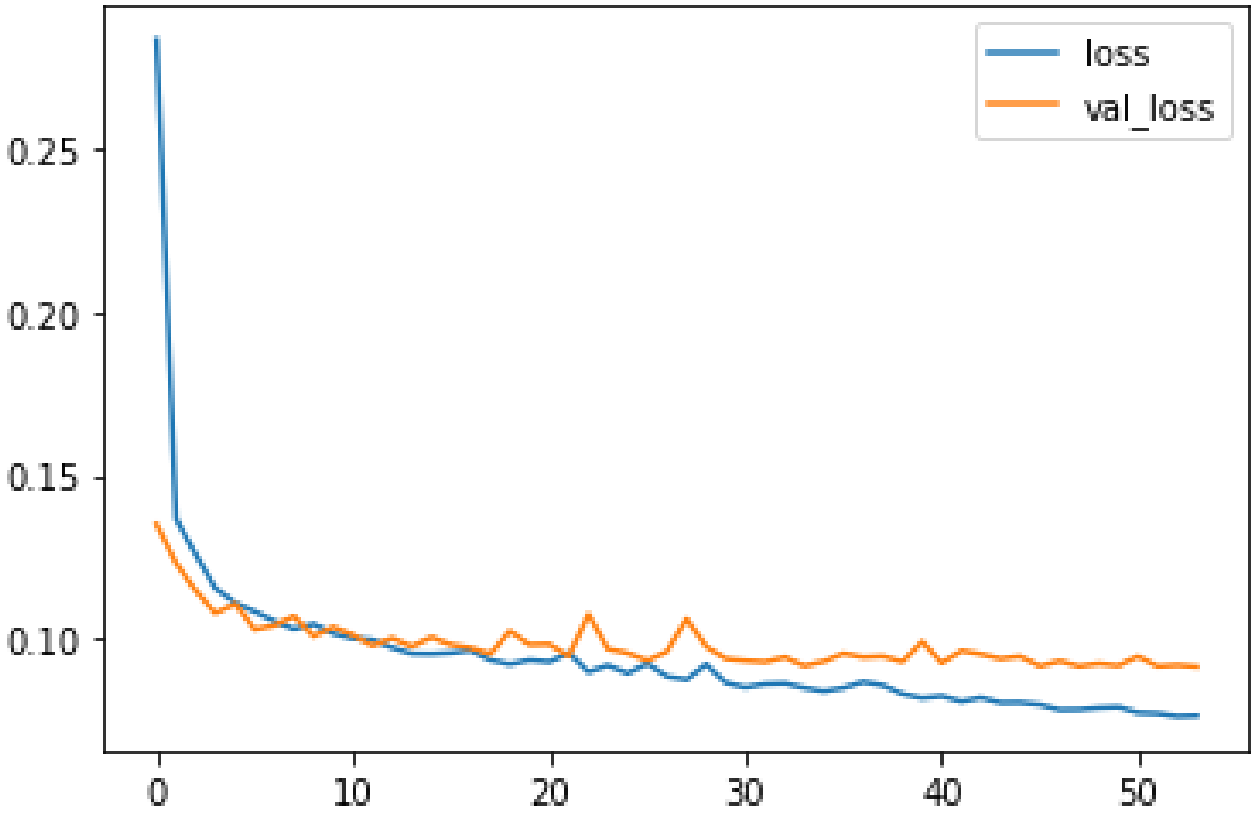




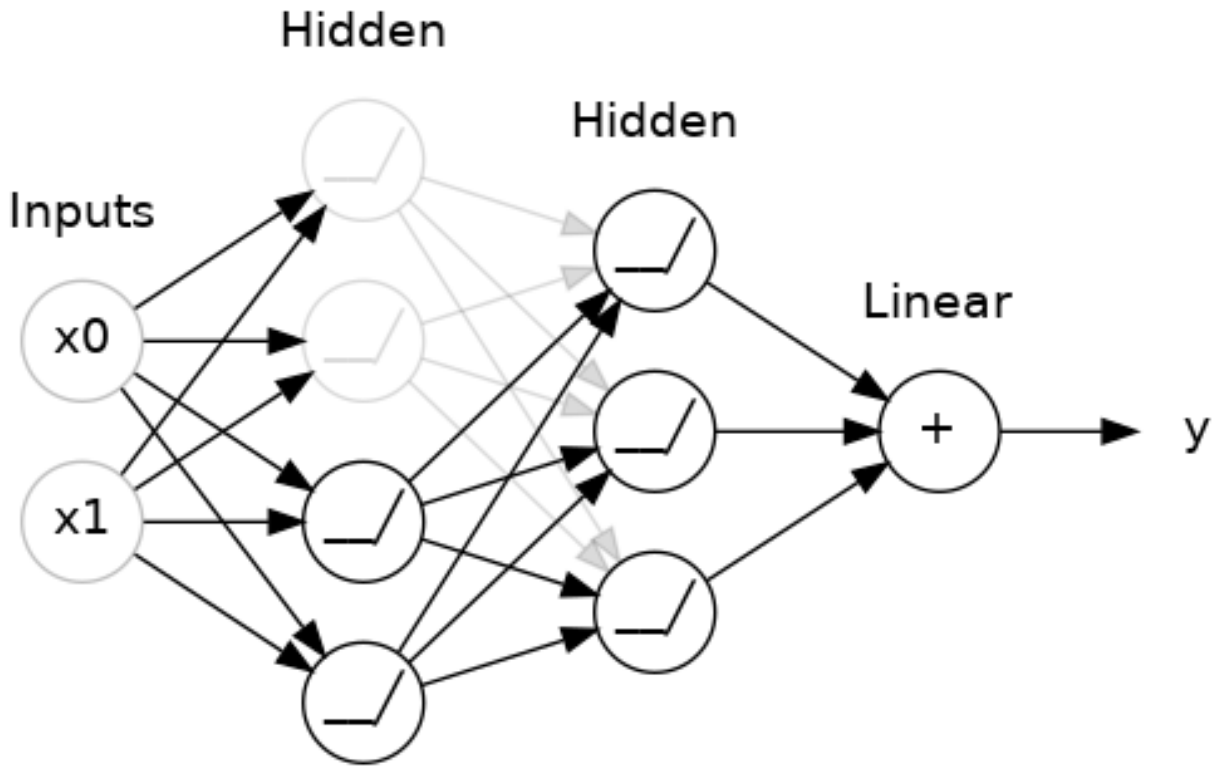
모델 학습은 언제 멈추어야 할까요? (Early Stopping)



학습 데이터와 검증 데이터 셋의 Loss 변환

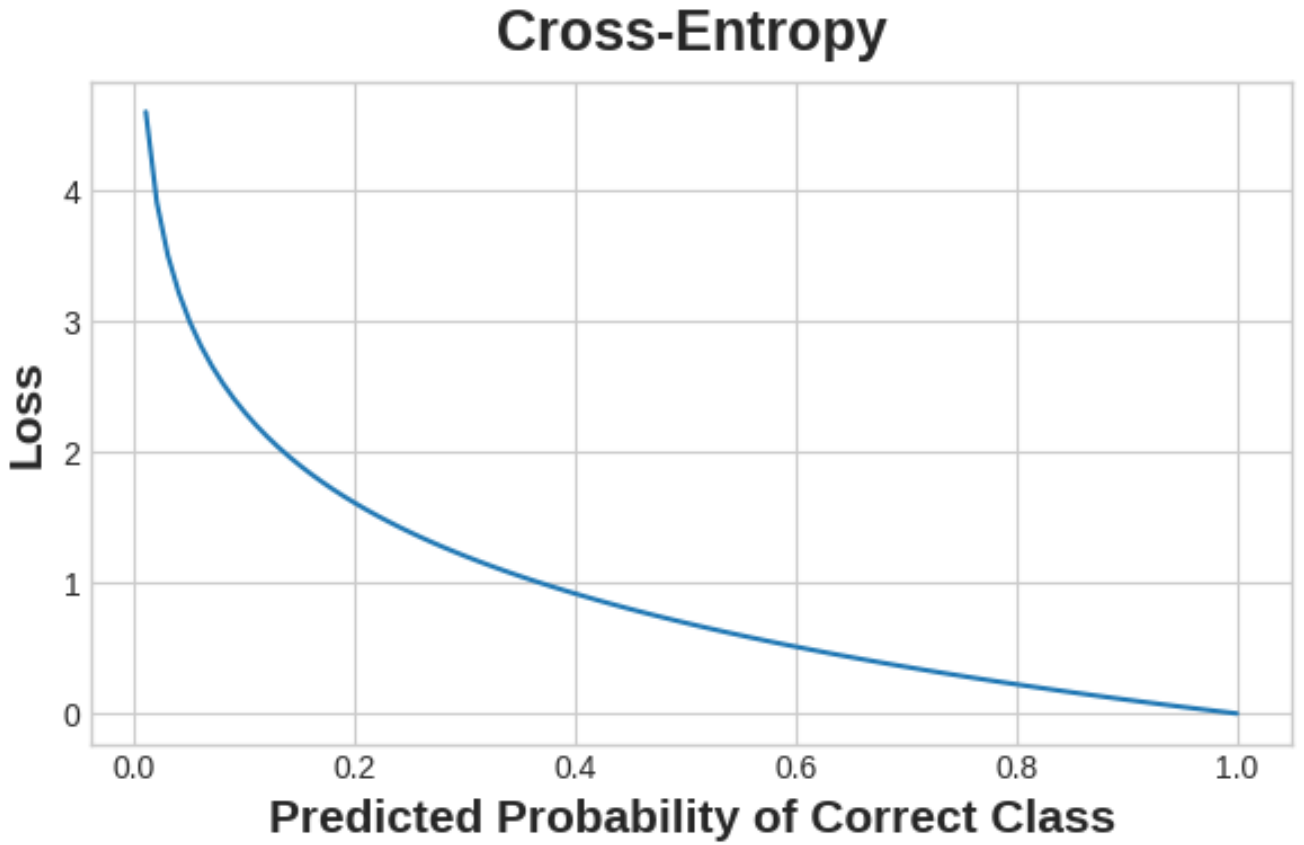


과적합을 방지하는 방법 : Dropout 과 Batch Normalization



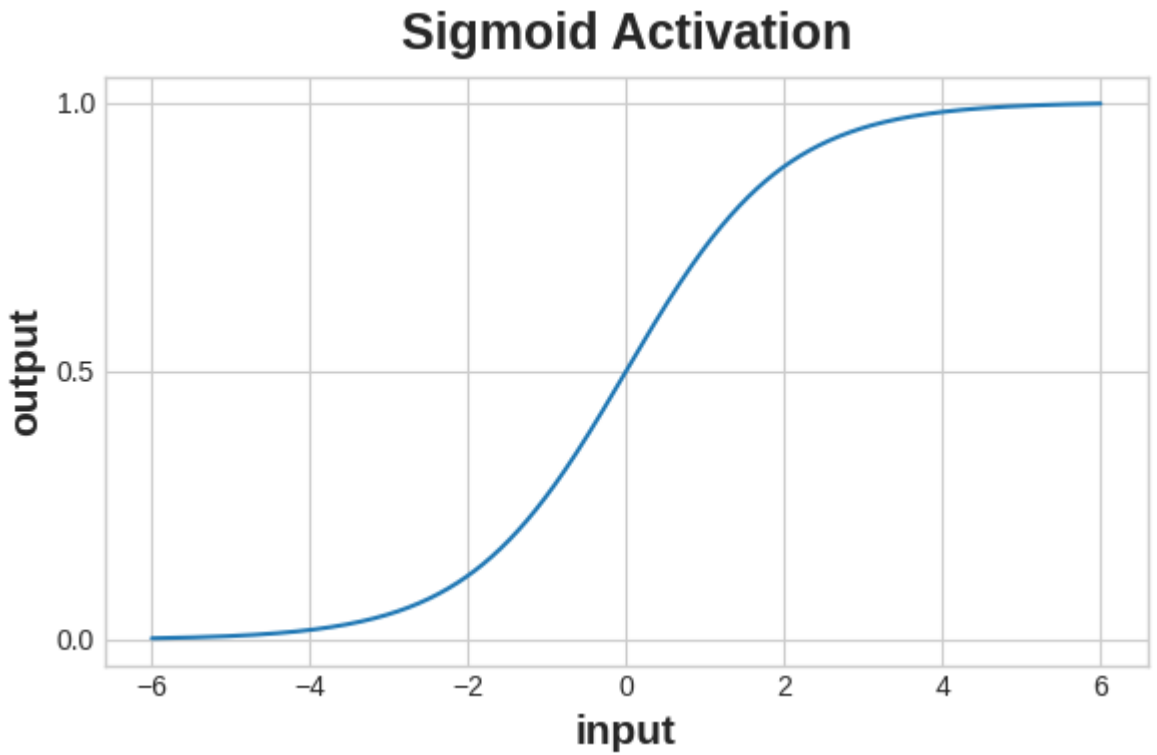
Cross-entropy

For classification, what we want instead is a distance between probabilities, and this is what cross-entropy provides. Cross-entropy is a sort of measure for the distance from one probability distribution to another.



Making Probabilities with the Sigmoid Function

The cross-entropy and accuracy functions both require probabilities as inputs, meaning, numbers from 0 to 1.



The sigmoid function maps real numbers into the interval [0,1] .

정 준 수 / Ph.D （ jsjeong@hansung.ac.kr ）

- 前) 삼성전자 연구원
- 前) 삼성의료원 (삼성생명과학연구소)
- 前) 삼성SDS (정보기술연구소)
- 現) (사)한국인공지능협회, AI, 머신러닝 강의
- 現) 한국소프트웨어산업협회, AI, 머신러닝 강의
- 現) 서울디지털재단, AI 자문위원
- 現) 한성대학교 교수(겸)
- 전문분야: Computer Vision, 머신러닝(ML), RPA
- https://github.com/JSJeong-me/