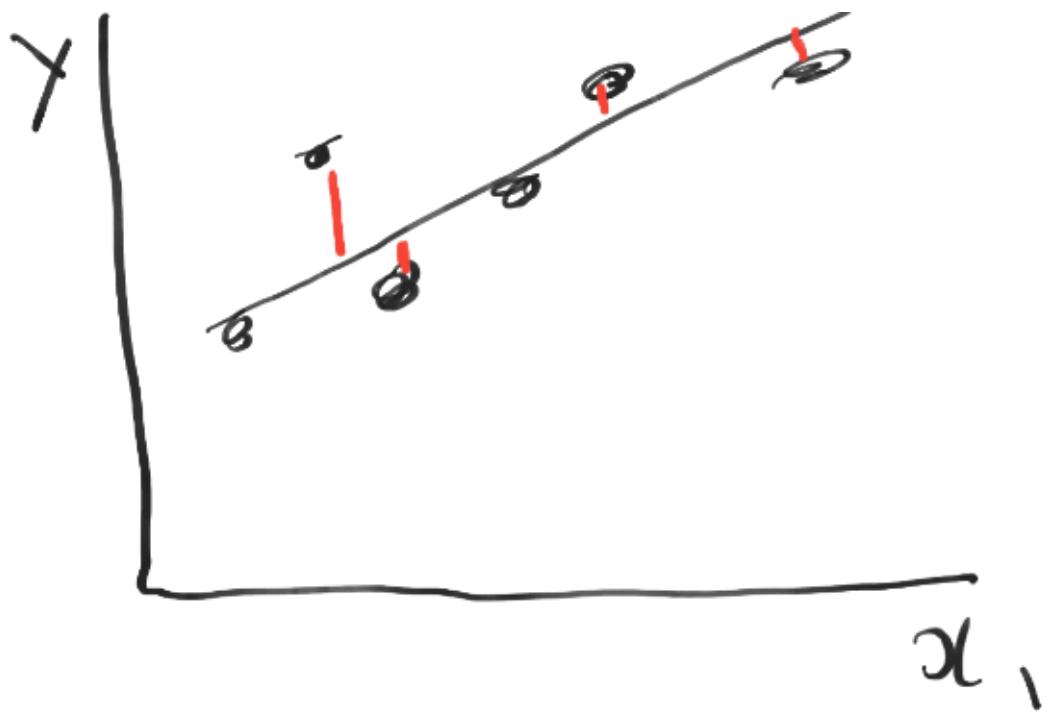


# 머신러닝



$$y = w_1 x_1 + b$$

손실 적을 수록 정확함

LOSS

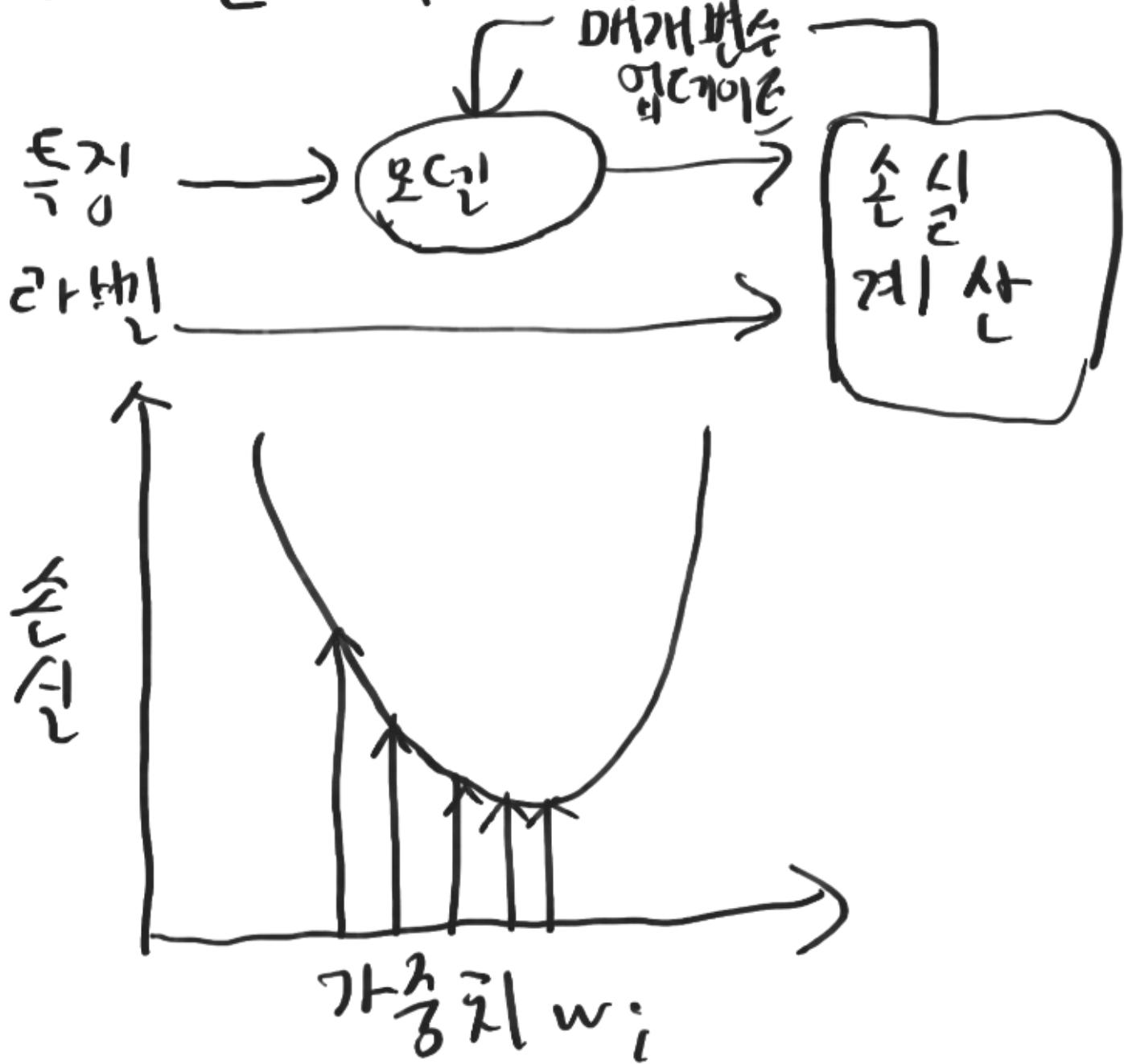
$$= (y - y')^2$$

실제값 - 예측값

전체 손실

$$= (loss_1 + loss_2 + \dots + loss_n) / n$$

# 손실 줄이기 학습법



기울기 하강법은

시작점 중요X

음의 방향으로 감소시킨다.

(기우기의 대체방법이 SGD)

1995-2-5

## 경사하강법

기울기가 학습률을 높하여  
다음지점 결정

학습률이 작으면 속도↓

// 크면 초기 손실률↑

### 1. 학률적 경사하강법

데이터가 많을 때

계산 속도↑

캡슐으로 데이터 하나를

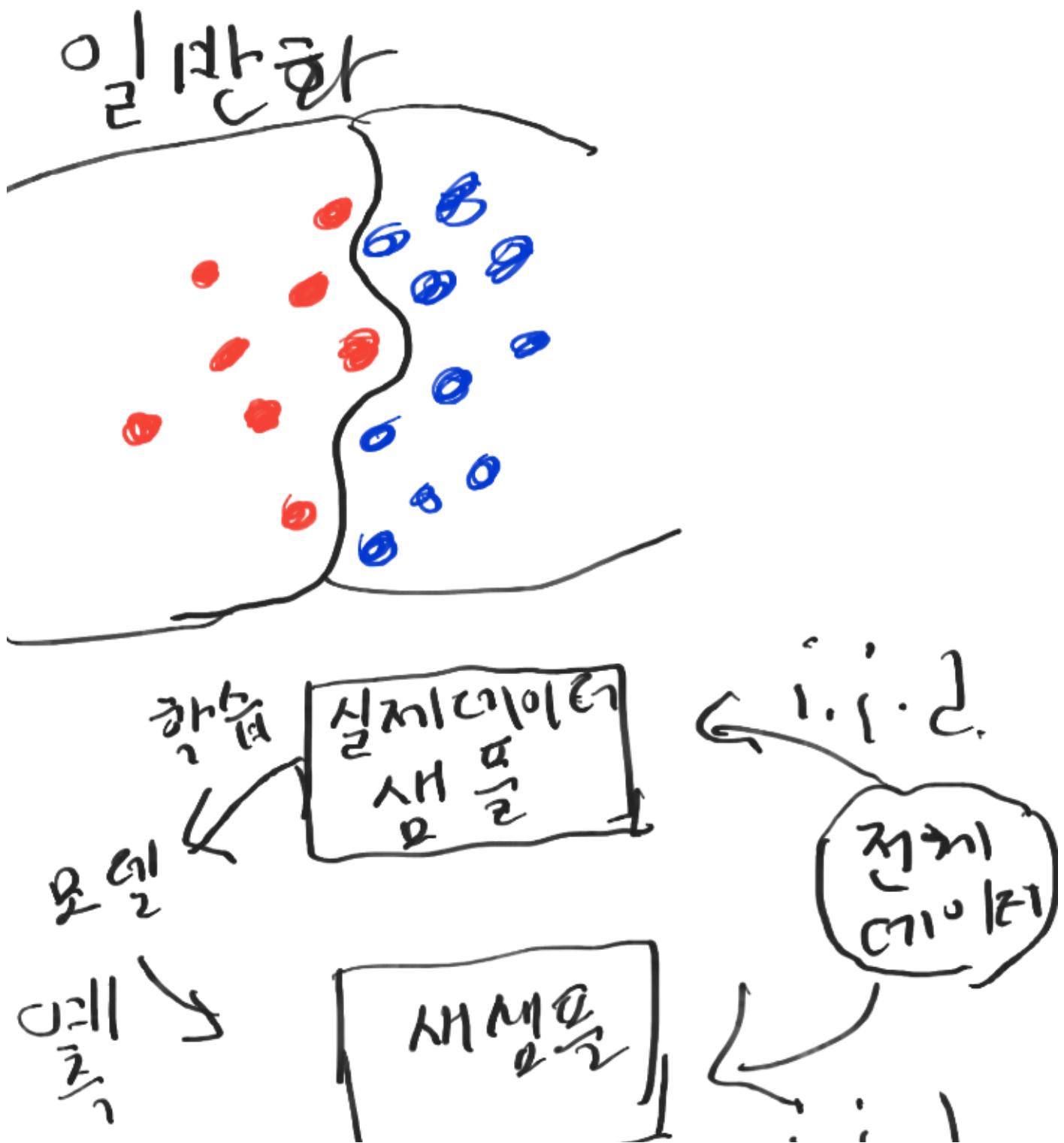
별도로 사용 경험적으로

증명은 됐지만 높아가 크다.

### 2. 비나 배치학률적 경사하강법

· 비가 배치학률적 경사하강법

1. 단세포 1회 미리법적 운동은  
1~10~100번의  
시도/시도는 2단계으로 나뉨



→ 경험적인 접근 방식  
 기준에 부합하는  
 새로운 데이터를 뽑아내야  
 테스트 시험 선호

### 주의점

1. 독립적 개별적으로  
 데이터를 뽑아야 한다.
2. 짧은 시간이 지나도  
 변하지 않아야 한다.
3. 항상 동일한 데이터를  
 사용 새로운 데이터  
 추가하면 안된다.  
 학습, 테스트 시험

## 오점의 연도는 뱀해

C기이터는 단순할수록 좋다.

### 과적합의 발생

학습하는 동안 손실은 적지만 잘 예측하지 못 한다.

↳ ~~X~~ 데이터 적합도<sup>2</sup> 유지하되 최대한

단순하게 바꾸어야 한다.

### 학습 세트

(테스트 세트) 성능이 높아야:  
하는

1. 테스트 세트가 충분히 크고

2. 같은 세트를 반복 사용 X

학습 세트 -

드는 힘을 주거나 -

많이

주기

하나로 흡수되지 않기하기 ~~X~~

Tensorflow  
Estimators

← 높은 수준의  
기계학적 학습

TF. layers, losses  
metrics

일반 모델  
구성 요소를 재사용  
가능성이 높다

tf Python

커널을  
재사용하는

TF C++

오픈 소스 API

CPU

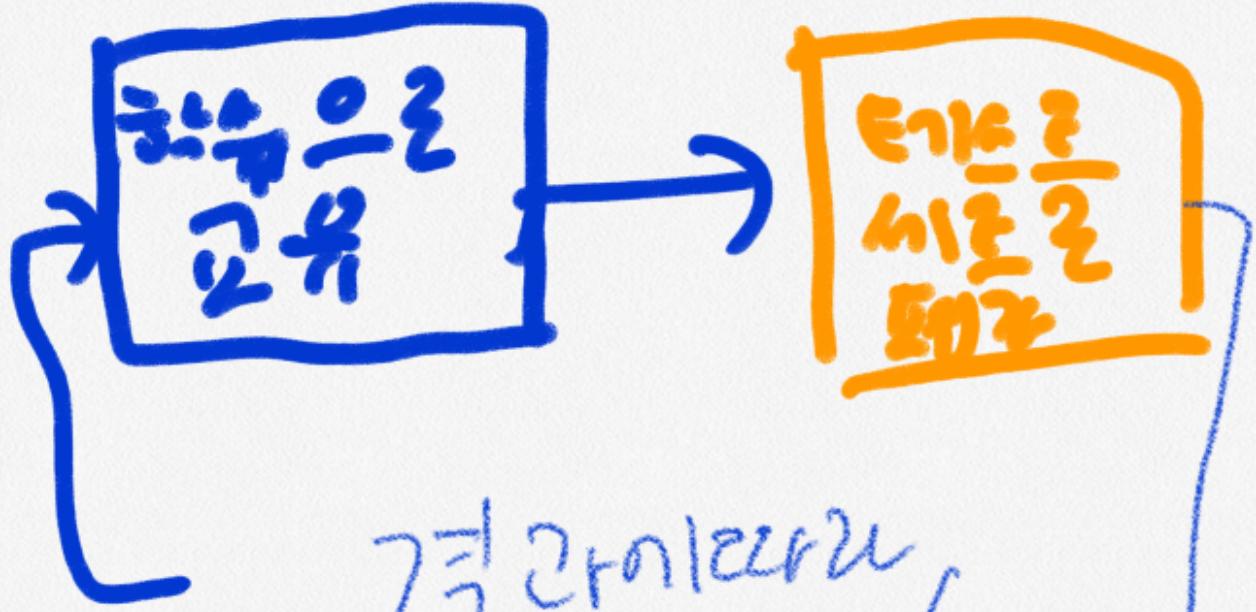
GPU

TPU

커널은  
하나의 종류의  
플랫폼에서  
작동

TensorFlow는  
그래픽 프로토콜 버퍼  
분산된 그래프를 실행하는 런타임  
TVM과 달리 속도가 느려  
모든 CPU(CPU 등)에서  
작동 가능해졌으나,  
초고수준의 API 주석은  
사용될로 유연성이  
적은  
초고수준의 API 주석은  
나중에 수준 낮은 편이다.

검증



결과에 따라  
모델간증



가장 우수한 모델

선택

고객은 모델이나 온수로 있다  
해결책



학습 검증 E가이

E가이스터 데이터 활용

검증데이터를 통해 민족드라마  
제작자가 나온 데 까지 터미널

등을 바탕으로 가며 검증

검증데이터를 통해 확인 후

E가이스터 데이터를 통해 재차 확인

표현

특성화

숫자

↓

숫자

문자

↓ 원핫인코딩

숫자

좁은 특성

1. 0이 아닌 데 IEEE-754

연속적으로 나와야 한다.

2. 부정하고 부호를 같이 하면

3. 특성은 특수값을 포함해 낸다

4. 특성의 정의는 시간이 지나도  
변하지 않아야 합니다.

5. 분단이 극단적인 이상점이  
없어야 한다.

비교기법



특정지역이 비율 나타내거나  
지역 나누기는 가능

가드

부록

시각화 출연원으로  $\rightarrow$  하드로그램

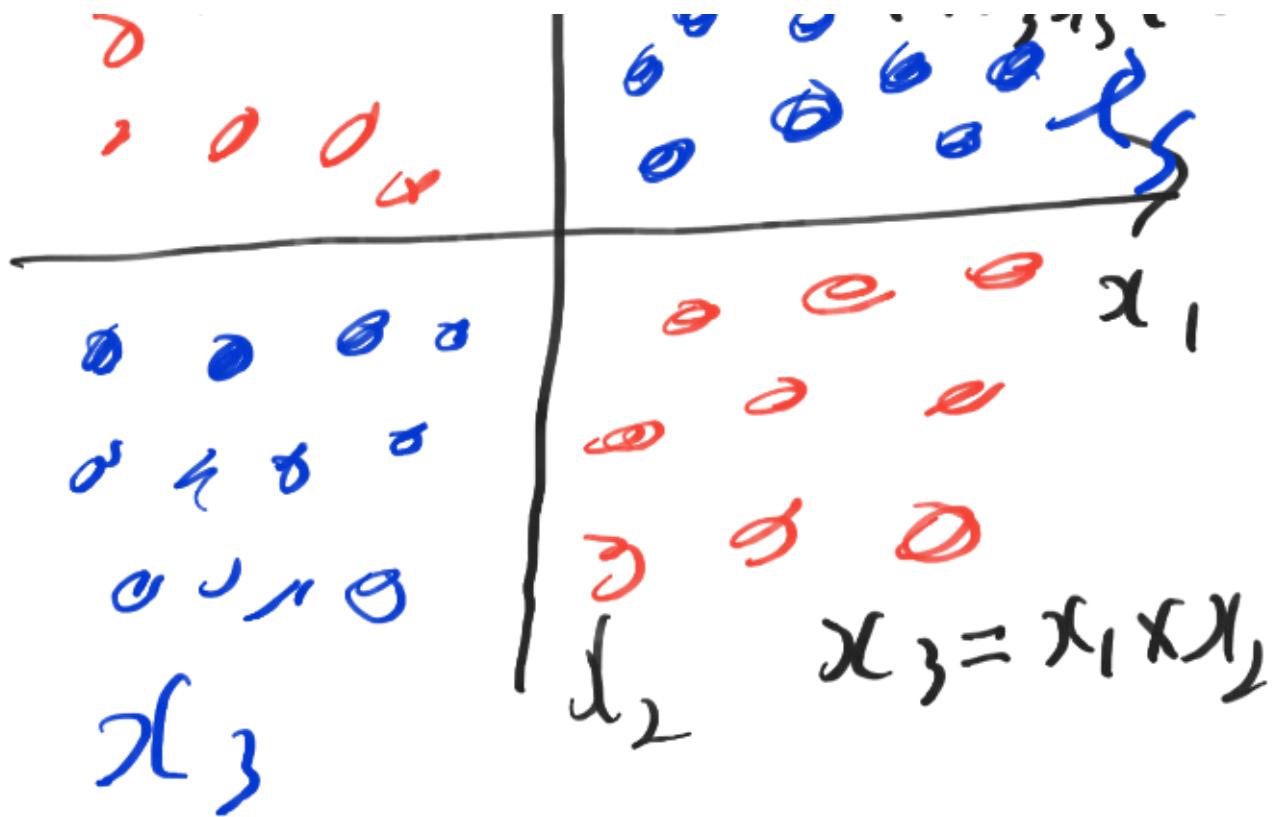
디버그 잘 웃된거 이미지가 없어  
모니터에

특성 분위 및 시간마다 다른  
예

특성화

1. 양식  $A \times B$ 의 템플릿을 찾을  
정의합니다.
2. 더 넓은  
3. A와 B가 빈 가능성을 찾으  
곱의 결과 범위가 매우  
최소화하기 나눌 수 있는

$$\begin{array}{c} 0 4 0 0 \\ - - 0 \\ \hline 0 4 0 0 \end{array} \quad \begin{array}{c} \uparrow \\ 0 0 \\ 0 0 \\ 0 0 \end{array} \quad \begin{array}{l} Y = w_1x_1 + w_2x_2 \\ \quad + b \end{array}$$

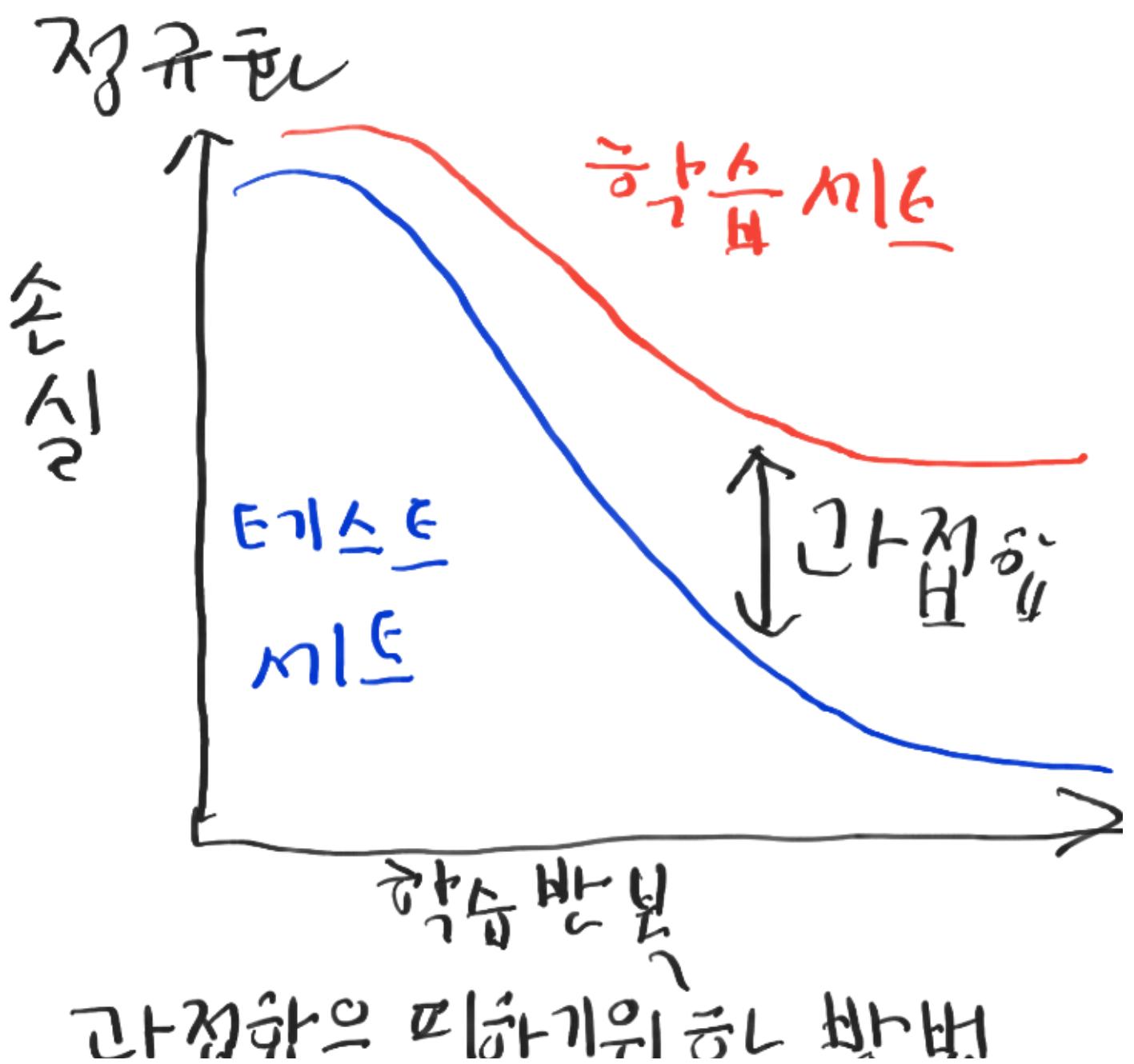


↳ 선형학습자와  
 비선형학습자를  
 합하여 표현 가능하다.  
 선형  $\rightarrow$  데량의 차이 더  
 많을 때  
 두 차차 차이 많으면  
 표현이 복잡하니

|

↳ 특정교차 와대량의 데미

$E^2$ 을 사용하면 복잡한 물질을  
효율적 표현 가능하다.



# 정규화

## 1. 조기중단

학습사이트가 상승하기 전에

멈춘다.

2. 모델 복잡도가 매번 틀리면 정

모델 복잡성이 모델 틀리를 준다

모델 정의하는 방식

1. 더 작은 가중치 선호

학습 과정을 쭉 풀어나면 매개변수

수작기만든다.

## 2. 정규화

복잡도 = 가중치의 제곱 합

아주큰 가중치에 대한 오기널 틀리게

서로 다른 데이터는 서로 비례되며

신영 노년기 2017-2018학년  
기울기선과

비이자인 사전학습

가중치는 0을 중심으로

가중치는 정규분포되어야

수식

수식

$$\underline{L(w, b)} + \cancel{\lambda w^T w}$$

주제, 점을  
이어 나누어진다

$L$  학습으로 낮추는 것이 목표

$\lambda$  가중치의 균형을 유지하는

방법을 제어하는 스칼라값

$w$  복잡도를 낮출 수 있다

2: 다음의 계급:  $L_2$  가중치 회

정규화

. . .

2

0

|| w ||\_2 = w\_1 + w\_2 + \dots  
 $w_n^2$

가장이 끝으면 모던한 눈  
데이터 관리 체계

↳ 충분한 학습 X  
x ↳ 모던 학습 관리  
특수성을 너무 많이 학습  
하지 않고 새로운 데이터 인식 X

X