Few-shot running

Softmax

Fully connected layer

Gradient

Heuristic (휴리스틱)

Hidden layer

Hinge loss: 수식 이해

Hyperparameter

Hyperplane

Inference

L1 loss

L1 regularization

L2 loss

L2 regularization

Lambda (람다)

Learning rate

Least Squares regression (최소 제곱 회귀)

Linear regression (선형 회귀)

Logistic regression (로지스틱 회귀)

Loss

MSE, Mean Squared Error (평균 제곱 오차)

Metric

Mini-batch

SGD

Momentum

Neural network

Neuron

Node

Weight

## 극소수 학습(few-shot learning)

객체 분류에 자주 사용되는 머신러닝 접근방식으로서 학습 예제 데이터가 적은 경우에서도 분류기를 효과적으로 학습하는 것을 목적으로 합니다.

[**원샷 학습**](https://developers.google.com/machine-learning/glossary/?hl=ko#one-shot_learning)을 참조하세요.

**소프트맥스(softmax)**

[**다중 클래스 분류 모델**](https://developers.google.com/machine-learning/glossary/?hl=ko#multi-class)에서 가능한 각 클래스의 확률을 구하는 함수입니다. 확률의 합은 정확히 1.0입니다. 예를 들어 소프트맥스는 특정 이미지가 강아지일 확률을 0.9로, 고양이일 확률을 0.08로, 말일 확률을 0.02로 판단할 수 있습니다. **전체 소프트맥스**라고도 합니다.

[**후보 샘플링**](https://developers.google.com/machine-learning/glossary/?hl=ko#candidate_sampling)과 대비되는 개념입니다.

## 완전 연결 레이어(fully connected layer)

각 [**노드**](https://developers.google.com/machine-learning/glossary/?hl=ko#node)가 후속 히든 레이어의 모든 노드에 연결된 [**히든 레이어**](https://developers.google.com/machine-learning/glossary/?hl=ko#hidden_layer)입니다.

완전 연결 레이어를 [**밀집 레이어**](https://developers.google.com/machine-learning/glossary/?hl=ko#dense_layer)라고도 합니다.

## 경사(gradient)

모든 독립 변수를 기준으로 한 [**편미분**](https://developers.google.com/machine-learning/glossary/?hl=ko#partial_derivative)의 벡터입니다. 머신러닝에서 경사는 모델 함수의 편미분의 벡터입니다. 경사는 가장 급격한 상승 방향을 가리킵니다.

## 휴리스틱(heuristic)

문제에 대해 이상적이지는 않지만 진전을 이루거나 교훈을 얻기에는 충분한 실용적인 해법입니다.

## 히든 레이어(hidden layer)

[**신경망**](https://developers.google.com/machine-learning/glossary/?hl=ko#neural_network)에서 [**입력 레이어**](https://developers.google.com/machine-learning/glossary/?hl=ko#input_layer)(특성)와 [**출력 레이어**](https://developers.google.com/machine-learning/glossary/?hl=ko#output_layer)(예측) 사이에 위치하는 합성 레이어입니다. 신경망에 하나 이상의 히든 레이어가 포함될 수 있습니다.

**힌지 손실(hinge loss)**

각 학습 예에서 최대한 멀리 떨어진 [**결정 경계**](https://developers.google.com/machine-learning/glossary/?hl=ko#loss)를 구하여 예와 경계 사이의 간격을 최대화하도록 고안된 [**분류**](https://developers.google.com/machine-learning/glossary/?hl=ko#classification_model)의 [**손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#decision_boundary) 함수군입니다. [**KSVM**](https://developers.google.com/machine-learning/glossary/?hl=ko#KSVMs)은 힌지 손실을 사용하거나 제곱 힌지 손실 등의 관련 함수를 사용합니다. 이진 분류에서는 힌지 손실 함수가 다음과 같이 정의됩니다.

손실최대값

여기에서 *y'*는 분류자 모델의 원시 출력입니다.

또한 *y*는 참 라벨(-1 또는 +1)입니다.

따라서 힌지 손실과 (y \* y')로 그래프를 그리면 다음과 같습니다.

## 초매개변수(hyperparameter)

모델 학습을 연속적으로 실행하는 중에 사용자 본인에 의해 조작되는 '손잡이'입니다. 예를 들어 [**학습률**](https://developers.google.com/machine-learning/glossary/?hl=ko#learning_rate)은 초매개변수 중 하나입니다.

[**매개변수**](https://developers.google.com/machine-learning/glossary/?hl=ko#parameter)와 대비되는 개념입니다.

## 초평면(hyperplane)

한 공간을 두 부분공간으로 나누는 경계입니다. 예를 들어 직선은 2차원의 초평면이고 평면은 3차원의 초평면입니다. 머신러닝에서 통용되는 초평면의 의미는 고차원 공간을 나누는 경계입니다. [**커널 서포트 벡터 머신**](https://developers.google.com/machine-learning/glossary/?hl=ko#KSVMs)은 일반적으로 초고차원 공간에서 초평면을 사용하여 포지티브 클래스와 네거티브 클래스를 구분합니다.

## 추론(inference)

머신러닝에서는 학습된 모델을 [**라벨이 없는 예**](https://developers.google.com/machine-learning/glossary/?hl=ko#unlabeled_example)에 적용하여 예측을 수행하는 과정을 의미할 때가 많습니다. 통계학에서는 특정한 관찰 데이터에 맞게 분포의 매개변수를 조정하는 과정을 의미합니다. [통계적 추론에 대한 위키백과 문서](https://en.wikipedia.org/wiki/Statistical_inference)를 참조하세요.

## L1 손실(L1 loss)

모델이 예측하는 값과 [**라벨**](https://developers.google.com/machine-learning/glossary/?hl=ko#loss)의 실제 값 차이의 절대값에 기초한 [**손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#label) 함수입니다. L1 손실은 [**L2 손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#squared_loss)보다 이상점에 둔감합니다.

## L1 정규화(L1 regularization)

가중치의 절대값 합에 비례하여 가중치에 페널티를 주는 [**정규화**](https://developers.google.com/machine-learning/glossary/?hl=ko#regularization) 유형입니다. [**희소 특성**](https://developers.google.com/machine-learning/glossary/?hl=ko#sparse_features)에 의존하는 모델에서 L1 정규화는 관련성이 없거나 매우 낮은 특성의 가중치를 정확히 0으로 유도하여 모델에서 해당 특성을 배제하는 데 도움이 됩니다. [**L2 정규화**](https://developers.google.com/machine-learning/glossary/?hl=ko#L2_regularization)와 대비되는 개념입니다.

## L2 손실(L2 loss)

[**선형 회귀**](https://developers.google.com/machine-learning/glossary/?hl=ko#linear_regression)에 사용되는 [**손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#loss) 함수입니다. **L2 손실**이라고도 합니다. 이 함수는 라벨이 있는 [**예**](https://developers.google.com/machine-learning/glossary/?hl=ko#example)에 대한 모델의 예측 값과 [**라벨**](https://developers.google.com/machine-learning/glossary/?hl=ko#label)의 실제 값 차이의 제곱을 계산합니다. 이 손실 함수는 제곱을 구하므로 부정확한 예측에 더 큰 영향을 줍니다. 즉, 제곱 손실은 [**L1 손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#L1_loss)보다 이상점에 민감하게 반응합니다.

## L2 정규화(L2 regularization)

가중치 [**제곱**](https://developers.google.com/machine-learning/glossary/?hl=ko#regularization)의 합에 비례하여 가중치에 페널티를 주는 정규화 유형입니다. L2 정규화는 높은 긍정 값 또는 낮은 부정 값을 갖는 이상점 가중치를 0은 아니지만 0에 가깝게 유도하는 데 도움이 됩니다. [**L1 정규화**](https://developers.google.com/machine-learning/glossary/?hl=ko#L1_regularization)와 대비되는 개념입니다. L2 정규화는 선형 모델의 일반화를 항상 개선합니다.

## 람다(lambda)

람다로 표현되는 스칼라값으로서, 정규화 함수의 상대적 중요도를 지정합니다. 다음은 정규화율의 영향을 보여주는 단순화된 [**손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#loss) 방정식입니다.

최소화(손실함수 + 정규화 함수)

정규화율을 높이면 [**과적합**](https://developers.google.com/machine-learning/glossary/?hl=ko#overfitting)이 감소하지만 모델의 [**정확성**](https://developers.google.com/machine-learning/glossary/?hl=ko#accuracy)이 떨어질 수 있습니다.

## 학습률(learning rate)

경사하강법을 통해 모델을 학습시키는 데 사용되는 스칼라값입니다. 각 반복에서 [**경사하강법**](https://developers.google.com/machine-learning/glossary/?hl=ko#gradient_descent) 알고리즘은 학습률을 경사에 곱합니다. 이 곱셈의 결과를 **경사 스텝**이라고 합니다.

학습률은 핵심적인 [**초매개변수**](https://developers.google.com/machine-learning/glossary/?hl=ko#hyperparameter)입니다.

## 최소 제곱 회귀(least squares regression)

[**L2 손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#L2_loss)을 최소화하면서 학습시킨 선형 회귀 모형입니다.

## 선형 회귀(linear regression)

[**회귀 모형**](https://developers.google.com/machine-learning/glossary/?hl=ko#regression_model)의 한 유형으로서, 입력 특성의 선형 조합으로부터 연속 값을 출력합니다.

## 로지스틱 회귀(logistic regression)

분류 문제에서 선형 예측에 [**시그모이드 함수**](https://developers.google.com/machine-learning/glossary/?hl=ko#sigmoid_function)를 적용하여 가능한 각 불연속 라벨값에 대한 확률을 생성하는 모델입니다. 로지스틱 회귀는 [**이진 분류**](https://developers.google.com/machine-learning/glossary/?hl=ko#binary_classification) 문제에 흔히 사용되지만 [**다중 클래스**](https://developers.google.com/machine-learning/glossary/?hl=ko#multi-class) 분류 문제에도 사용될 수 있습니다. 이러한 경우를 **다중 클래스 로지스틱 회귀** 또는 **다항 회귀**라고 합니다.

## 손실(loss)

모델의 [**예측**](https://developers.google.com/machine-learning/glossary/?hl=ko#prediction)이 [**라벨**](https://developers.google.com/machine-learning/glossary/?hl=ko#label)과 얼마나 차이가 나는지를 측정합니다. 다르게 표현하면, 모델이 얼마나 부정확한지를 나타냅니다. 이 값을 판단하려면 모델에서 손실 함수를 정의해야 합니다. 예를 들어 선형 회귀 모형은 일반적으로 [**평균 제곱 오차**](https://developers.google.com/machine-learning/glossary/?hl=ko#MSE)를, 로지스틱 회귀 모형은 [**로그 손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#Log_Loss)을 손실 함수로 사용합니다.

## 평균 제곱 오차(MSE, Mean Squared Error)

예시당 평균 제곱 손실입니다. MSE는 [**제곱 손실**](https://developers.google.com/machine-learning/glossary/?hl=ko#squared_loss)을 [**예시**](https://developers.google.com/machine-learning/glossary/?hl=ko#example)의 개수로 나누어 계산합니다. [**텐서플로우 플레이그라운드**](https://developers.google.com/machine-learning/glossary/?hl=ko#TensorFlow_Playground)에서 '학습 손실' 및 '테스트 손실'로 표시하는 값이 MSE입니다.

## 측정항목(metric)

중요한 의미가 있는 수치입니다. 머신러닝 시스템에서 직접 최적화될 수도 있고, 그렇지 않을 수도 있습니다. 시스템에서 최적화를 시도하는 측정항목을 [**목표**](https://developers.google.com/machine-learning/glossary/?hl=ko#objective)라고 합니다

## 미니 배치(mini-batch)

학습 또는 추론의 단일 반복에서 함께 실행되는 [**예**](https://developers.google.com/machine-learning/glossary/?hl=ko#example)의 전체 배치 중에서 무작위로 선택한 소규모 부분집합입니다. 미니 배치의 [**배치 크기**](https://developers.google.com/machine-learning/glossary/?hl=ko#batch_size)는 일반적으로 10~1,000입니다. 전체 학습 데이터가 아닌 미니 배치의 손실을 계산하면 효율성이 크게 향상됩니다.

## 미니 배치 확률적 경사하강법(SGD, mini-batch stochastic gradient descent)

[**미니 배치**](https://developers.google.com/machine-learning/glossary/?hl=ko#mini-batch)를 사용하는 [**경사하강법**](https://developers.google.com/machine-learning/glossary/?hl=ko#gradient_descent) 알고리즘입니다. 즉, 미니 배치 SGD는 학습 데이터 중 작은 부분집합을 기반으로 경사를 예측합니다. [**기본적인 SGD**](https://developers.google.com/machine-learning/glossary/?hl=ko#SGD)에서는 크기가 1인 미니 배치를 사용합니다.

## 모멘텀(Momentum)

학습 단계가 현재 단계의 도함수뿐 아니라 바로 앞 단계의 도함수에도 의존하는 정교한 경사하강법 알고리즘입니다. 모멘텀에서는 물리학의 모멘텀(운동량) 계산과 마찬가지로 시간에 따른 경사의 지수가중이동평균을 계산합니다. 모멘텀은 학습이 국소 최저점에서 정체되는 현상을 방지하는 데 도움이 될 수 있습니다.

## 신경망(neural network)

사람의 두뇌를 본뜬 모델로서, 단순 연결 유닛 또는 [**뉴런**](https://developers.google.com/machine-learning/glossary/?hl=ko#neuron)으로 이루어지며 비선형성을 갖는 여러 레이어로 구성됩니다. 하나 이상의 레이어는 [**히든 레이어**](https://developers.google.com/machine-learning/glossary/?hl=ko#hidden_layer)입니다.

## 뉴런(neuron)

[**신경망**](https://developers.google.com/machine-learning/glossary/?hl=ko#neural_network)의 노드로서, 일반적으로 여러 입력 값을 취하여 하나의 출력 값을 생성합니다. 뉴런은 입력 값의 가중 합에 [**활성화 함수**](https://developers.google.com/machine-learning/glossary/?hl=ko#activation_function)(비선형 변환)를 적용하여 출력 값을 계산합니다.

## 노드(node)

중복으로 정의된 용어로서 다음 중 하나를 의미합니다.

* [**히든 레이어**](https://developers.google.com/machine-learning/glossary/?hl=ko#hidden_layer)의 뉴런
* 텐서플로우 [**그래프**](https://developers.google.com/machine-learning/glossary/?hl=ko#graph)의 연산

## 가중치(weight)

선형 모델에서 [**특성**](https://developers.google.com/machine-learning/glossary/?hl=ko#feature)의 계수 또는 심층 네트워크의 엣지입니다. 선형 모델 학습의 목표는 각 특성의 이상적인 가중치를 결정하는 것입니다. 가중치가 0인 특성은 모델에 영향을 주지 못합니다.