

허준혁 불참

1. 기계학습에서 학습이란 훈련 데이터로부터 손실 함수를 줄이기 위한 가중치 매개 변수의 최적값을 자동으로 획득하는 것이다. 이때 가중치는 모델의 파라미터로, 각 신호의 영향력을 조절하는 매개변수이다. 가중치 값을 조정하여 모델의 성능을 최적화해준다. 손실 함수란 모델의 출력과 실제 정답 사이의 차이를 측정하는 함수이다. 모델이 예측한 결과와 실제 결과 간의 오차를 정량화하며, 이 오차를 최소화하기 위해 모델의 가중치를 조절한다. 일반적으로 손실 함수는 모델의 예측값과 정답 사이의 차이를 계산하며, 이 차이가 작을수록 함수의 값은 작아진다.

2. 확률적 경사 하강법 코드 분석

* `n_epochs = 50`은 학습을 반복하는 횟수를 50으로 설정해준다.

* `+0, +1 = 5`, 50은 학습을 반복하는 횟수를 50으로 설정해준다. 이러한 매개변수는 학습률 스케줄링에 사용된다.

* `def learning_schedule(t)`는 학습 스케줄링 함수를 정의한다. `t`는 현재 학습 스텝을 나타낸다. 이 함수는 현재 스텝을 기반으로 학습률을 반환한다.

* `theta = np.random.randn(2, 1)`은 무작위로 초기화된 가중치 `theta`를 생성한다. 이 모델은 선형 회귀 모델로 가정되며 `2x1` 크기의 가중치 벡터이다.

* `for epoch in range(n_epochs)`: 주요 학습 루프에서 `epoch`를 반복한다. 에포크는 전체 데이터셋을 한번 통화하는 것을 나타낸다.

* `for l in range(m)`: 각 에포크에서 데이터 포인트를 반복한다. `m`은 전체 데이터 포인트의 수이다.

* `random_index = np.random.randint(m)`: 무작위로 데이터 포인트를 선택하기 위해 0부터 `m`까지의 무작위 인덱스 `random_index`를 생성한다.

* $xi = X_b[random_index_index + 1]$ 및 $yi = y[random_index:random_index + 1]$: 무작위로 선택된 데이터 포인트와 해당 정답을 xi 와 yi 로 가져온다. 이것은 확률적 경사 하강법의 핵심 아이디어 중 하나이다.

* $gradients = 2 * xi.T.dot(xi.dot(theta) - yi)$ 는 현재 데이터포인트를 사용하여 그래디언트를 계산한다. 이것은 손실 함수의 그래디언트를 나타내며 선형 회귀 문제에서는 평균 제곱 오차의 그래디언트를 계산한다.

* $eta = learning_schedule(epoch*m + i)$: 학습 스케줄링 함수를 사용하여 현재 스텝에 해당하는 학습률 eta 를 얻는다. $epoch*m + i$ 는 현재 학습 스텝을 나타내며 스케줄링 함수를 통해 학습률이 동적으로 조절된다.

* $theta = theta - eta * gradients$ 현재 가중치를 그래디언트와 학습률을 사용하여 업데이트 한다. 이것은 모델의 파라미터를 경사하강법을 통해 조정하는 단계이다.