지능형 플랫폼 프로젝트 보고서

20201040 손가현

**1. 데이터 추출 및 분리**

텍스트이(가) 표시된 사진

자동 생성된 설명

Boston dataset을 사용하여 선형회귀를 진행한다.

해당 데이터의 속성을 data에 저장하고 클래스값을 label에 저장한다.

텍스트이(가) 표시된 사진

자동 생성된 설명

넘파이 배열을 데이터 프레임으로 변환한다.

W2X2 + W1X1 + W0 형태로 입력값을 설정한다. 따라서 2개의 독립변수값 x1, x2가 필요하다.

LSTAT(급여가 낮은 직종에 종사하는 비율)을 x1으로, RM(집의 평균 방 수)을 x2로 추출한다. y는 주택 가격이다. x, y의 훈련셋과 테스트셋을 각각 400개, 106개로 나눈다.

**2. 파라미터 리스트, 오차 리스트 생성 및 초기화**

텍스트이(가) 표시된 사진

자동 생성된 설명

파라미터 W0, W1, W2 을 길이가 10인 일차원 리스트로 선언하고 0으로 초기화한다.

cost를 3차원 리스트로 선언하고 0으로 초기화한다. cost는 W0,W1, W2 조합에 따른 평균 오차 제곱을 저장한다. W0[0],W1[0], W2[0]을 0~9사이의 랜덤한 숫자로 초기화한다.

초기화된 W0[0], W1[0], W2[0]을 출력하면 위와 같다.

**3. 훈련셋으로 파라미터 학습 ( W2[1]~W2[10] -> W1[1]~W1[10] -> W0[1]~W0[10] )**

3.1 W2업데이트

텍스트이(가) 표시된 사진

자동 생성된 설명

학습을 통해 W2값을 업데이트한다. W2[1]~ W2[9]까지 업데이트하므로 바깥의 for문은 9번

반복된다. 훈련셋은 총 400개이므로 안쪽 for문은 400번 반복된다. MSE를 W2에 대해 미분한

식에 훈련셋 x,y를 대입한다. 미분한 식은 다음과 같다.



W0, W1은 아직 학습되지 않았으므로 각각W0[0], W1[0]을 사용한다. 400번 더한 총합을 diff\_w2에

저장한다. 해당 총합의 평균을 구하여 diff\_w2에 다시 대입한다. diff\_w2 구하는 과정을 수식으로

정리하면 다음과 같다.



3.2 W1 업데이트

텍스트이(가) 표시된 사진

자동 생성된 설명

학습을 통해 W1값을 업데이트한다. W1[1]~ W1[9]까지 업데이트하므로 바깥의 for문은 9번

반복된다. 훈련셋은 총 400개이므로 안쪽 for문은 400번 반복된다. MSE를 W1에 대해 미분한

식에서 훈련셋 x\_train, y\_train를 대입한다. 미분한 식은 다음과 같다.



W0는 아직 학습되지 않았으므로 W0[0]을 사용한다. 총합은 diff\_w1에 저장한다. 해당 총합의

평균을 구하여 diff\_w1에 다시 대입한다. diff\_w1을 구하는 과정을 수식으로 정리하면 다음과

같다. (W2는 W2[9], W0은 W0[0]을 사용한다.)



기존의 W1에서 학습률\*diff\_w1의 값을 빼서 W1을 업데이트한다. 초기 학습률은 0.01이고,

W1값을 한 번 업데이트할 때마다 학습률이 0.001씩 감소한다.

3.3 W0 업데이트

텍스트이(가) 표시된 사진

자동 생성된 설명

학습을 통해 마지막 파라미터인 W0값을 업데이트한다. W0[1]~ W0[9]까지 업데이트한다.

MSE를 W0에 대해 미분한 식에 훈련셋 x,y를 대입한다. 미분한 식은 다음과 같다.



W1, W2는 업데이트된 W1[9], W2[9]을 사용한다. 400번 더한 총합을 diff\_w0에 저장한다. 해당

총합의 평균을 구하여 diff\_w0에 다시 대입한다. diff\_w0를 수식으로 정리하면 다음과 같다.

(W2는 W2[9], W1은 W1[9]를 사용한다)



기존의 W0에서 학습률\*diff\_w0의 값을 빼서 업데이트한다. 초기 학습률은 0.01이고, W0값을 한

번 업데이트할 때마다 학습률이 0.001씩 감소한다.

텍스트이(가) 표시된 사진

자동 생성된 설명

각각의 파라미터들이 어떻게 업데이트 되는지 출력한다.

**4. Cost 최소로 하는 최적의 파라미터 찾기**

텍스트이(가) 표시된 사진

자동 생성된 설명

평균 제곱 오차를 구해 3차원리스트에 저장하는 부분이다.

5~6번째 줄은 평균 제곱 오차를 구하는 것이다. 해당 공식은 다음과 같다.

텍스트이(가) 표시된 사진

자동 생성된 설명

3차원 리스트를 사용하므로 리스트값에 접근하기 위해 3개의 for문이 필요하다.

0~399개까지의 trainset을 사용하므로 안쪽 for문이 필요하다.

인덱스 i는 W0, k는W1, v는W2리스트에 각각 접근한다.

출력문을 통해 인덱스 번호에 따른 평균 오차 제곱을 확인할 수 있다.

텍스트이(가) 표시된 사진

자동 생성된 설명

평균 오차 제곱을 가장 최소화하는 cost값의 인덱스를 찾는다. W0 W1, W2 에 대응되는 세 개의 인덱스를 구해야 하므로 삼중 반복문이 된다. cost리스트에 모두 접근하여 저장한 최소값보다 작을 때마다 최소값을 갖는 인덱스를 갱신한다. 각각 인덱스들을 W0, W1, W2 리스트에 대입해 최적의 파라미터 W0 W1, W2를 구한다.

테이블이(가) 표시된 사진

자동 생성된 설명

W0[9], W1[1], W2[8] 일 때 cost가 최소화된다. 최적의 파라미터 값들은 위와 같다.

**5. 테스트셋으로 테스트 후 오차 구하기**

텍스트이(가) 표시된 사진

자동 생성된 설명

예측값 -1.434…\*X2+2.544…\*X1+2.8837…와 실제 y값을 뺀 값의 제곱합을 구한다. 총 106개의 테스트셋이므로 106번 반복된다. 그 후 평균을 낸 값을 test\_cost에 다시 대입한다. 테스트셋을 돌려본 결과 테스트cost는 955.8007…이 나온다.