

인공지능(딥러닝)개론 HW1 결과 보고서

2020.10.13.(화) 11:59PM

20181202 김수미

1. XOR 연산과 Backpropagation

먼저 **XOR 연산**이란 X1값과 X2값이 있을 때 두 값이 같으면 $Y=0$ 을 반환하고, 두 값이 다르면 $Y=1$ 을 반환하는 연산이다. 이 연산은 linear 한 모양의 관계성을 가지지 않기 때문에 기존의 회귀 연산으로는 구현할 수 없었지만, '**Backpropagation(역전파)**' 알고리즘이 개발되면서 XOR 연산을 수행하는 인공지능을 구현할 수 있게 되었다. Backpropagation 알고리즘이란 출력값의 오차를 계속 체크하면서 이를 피드백 신호로 사용해, 현재 샘플의 손실 점수가 감소되는 방향으로 가중치 값을 조금씩 수정해 나가는 알고리즘이다.

2. 코드에 대한 간단한 설명

해당 코드는 XOR table을 학습하여 XOR연산을 수행하는 Neural Network 모델을 형성한다. python언어로 작성된 코드이며, Google Colab를 통해 작성했다. Output, hidden layer의 Activation function은 모두 Sigmoid 함수($1 / \{ \text{np.exp}(-x)+1 \}$)를 사용했으며, Cost function으로는 Binary Cross entropy 함수($-\{ y\log(\hat{y}) + (1-y)\log(1-\hat{y}) \}$)를 사용했다.

1) 1-layer 모델

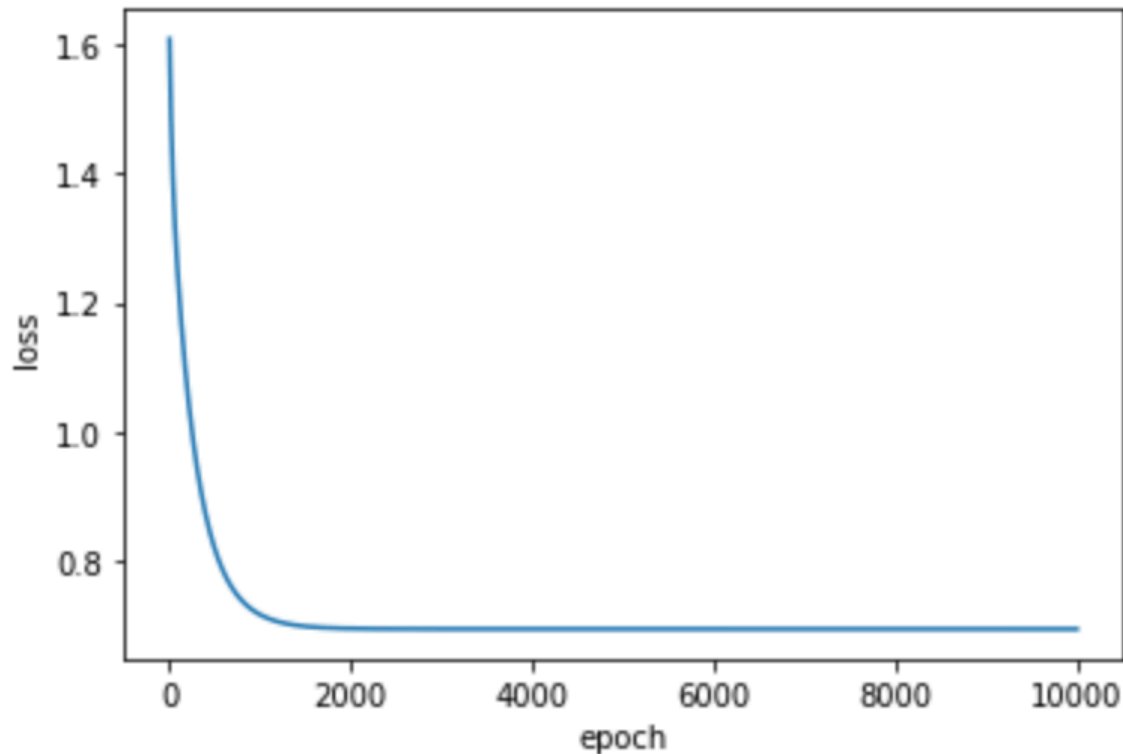
1-layer모델은 Hidden Layer가 없고 오직 input, output layer 만 가진다. 해당 모델은 linearly separable classes만 분류가능하다는 한계를 가지고 있다. 이러한 특성은 아래 XOR연산 결과에서도 확인할 수 있다. Hidden Layer가 없어도 최종 값에 대한 손실 값을 매 반복 마다 체크하며 Backpropagation 알고리즘을 통해 input, output layer의 weight값과 bias 값을 업데이트한다.

2) 2-layer 모델

1-layer모델은 Hidden Layer를 1개 가지는 Neural Network모델이다. Hidden Layer 의 unit개수는 3개로 잡았고, 최종 값에 대한 손실 값을 매 반복 마다 체크하며 Backpropagation 알고리즘을 통해 hidden, input, output layer의 weight값과 bias값을 손실 값을 줄여 나가는 방향으로 업데이트한다.

3. 각 모델에서의 Loss 그래프 및 Testing Step 출력 결과

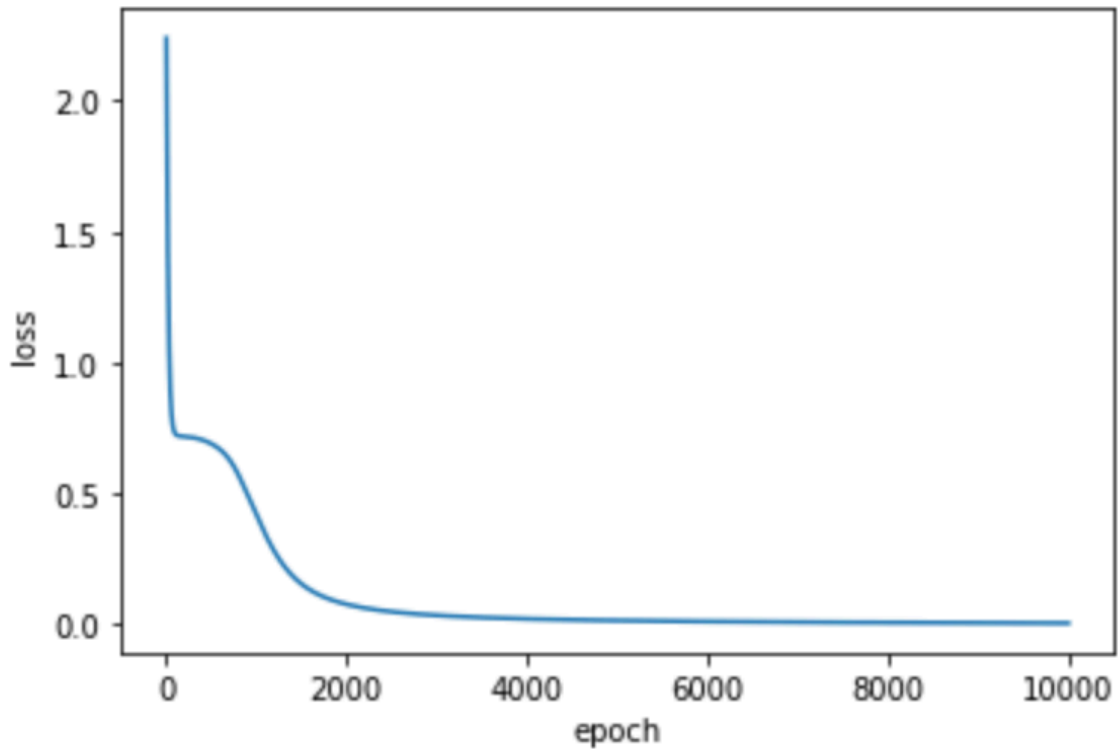
1) 1-layer 모델



```
↳ input:  [1 1] , answer:  0 , pred: 0.5000
   input:  [1 0] , answer:  1 , pred: 0.5000
   input:  [0 1] , answer:  1 , pred: 0.5000
   input:  [0 0] , answer:  0 , pred: 0.5000
```

1-layer모델의 학습 결과를 반복 횟수에 따른 loss값으로 그래프를 그려 나타낸 것과, 학습 후의 weight, bias값으로 XOR연산을 수행한 결과이다. 학습이 반복되어 수행될수록 loss값은 감소하지만, 1-layer 모델은 linearly separable classes만 분류가 가능하다는 한계로 인해 linear한 관계성을 가지지 않는 XOR연산에 대해, 모든 input에 대한 예측값을 학습 후에도 정답과 오답 사이의 중앙값인 0.5를 일정하게(linearly) 출력하고 있음을 확인할 수 있다.

2) 2-layer 모델



```
[> input:  [1 1] , answer:  0 , pred: 0.0053
      input:  [1 0] , answer:  1 , pred: 0.9951
      input:  [0 1] , answer:  1 , pred: 0.9909
      input:  [0 0] , answer:  0 , pred: 0.0053
```

2-layer모델의 학습 결과를 반복 횟수에 따른 loss값으로 그래프를 그려 나타낸 것과, 학습 후의 weight, bias값으로 XOR연산을 수행한 결과이다. 1-layer모델과 동일하게 학습 횟수가 늘어날수록, loss값이 감소하지만 1-layer모델과는 달리 input값에 대한 예측치가 정답에 가깝게 수렴하는 것을 확인할 수 있다.