**Deep Learning practice #2-1 report**

컴퓨터소프트웨어학부 2018008768 윤정

**[구현 과정]**

텍스트이(가) 표시된 사진

자동 생성된 설명

Practice 1-2에서 진행했던 것과 동일하게 input X와 {0, 1} 로 구성된 y를 생성하였습니다.

Logistic regression을 Back propagation을 이용하여 W1, b1, W2, b2 를 구하였습니다.

이 과정에서, 계산상의 에러 방지 및 더 좋은 성능을 위해 정규화를 진행하였습니다. degree value인 x값을 radian value값으로 변환한 뒤 실험을 진행하였습니다.

구한 parameter W1, b1, W2, b2 값에 대하여 Cost와 Accuracy를 계산하였습니다.

**[결과]**

m = 10000, n = 1000, K = 5000, alpha = 0.9 일 때 결과값 (매 500회마다 출력, 최종 w와 b)

텍스트이(가) 표시된 사진

자동 생성된 설명텍스트이(가) 표시된 사진

자동 생성된 설명

alpha값을 변화시키며 예측한 결과값입니다. (M = 10000, N = 1000, k = 5000)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | alpha=0.01 | alpha=0.05 | alpha=0.1 | alpha=0.5 | alpha=0.9 |
| Cost (with ‘n’ test samples) | 0.671444826246355 | 0.11022152007795825 | 0.05207759044485469 | 0.03323181182795685 | 0.01662077096127 |
| Accuracy (with ‘n’ test samples) | 72.7 | 99.1 | 99.8 | 98.3 | 100.0 |

가장 작은 cost를 가지는 0.9를 본 실험에서 알파 값으로 사용하였습니다.

Train data m을 변화시키며 예측한 결과값입니다. (alpha=0.9)

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10, n=1000, K=5000 | m=100, n=1000, K=5000 | m=10000, n=1000, K=5000 |
| Cost (with ‘m’ train samples) | 0.00130054369998 | 0.005304403602337 | 0.014578705977754 |
| Cost (with ‘n’ test samples) | 0.092721242010896 | 0.034993528607039 | 0.016844615188822 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10, n=1000, K=5000 | m=100, n=1000, K=5000 | m=10000, n=1000, K=5000 |
| Accuracy (with ‘m’ train samples) | 100.0 | 100.0 | 99.98 |
| Accuracy (with ‘n’ test samples) | 96.39999 | 98.1 | 99.9 |

K를 변화시키며 예측한 결과값입니다. (alpha = 0.9)

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10000, n=1000, K=10 | m=10000, n=1000, K=100 | m=10000, n=1000, K=5000 |
| Cost (with ‘m’ train samples) | 0.691027501062590 | 0.523459438328402 | 0.014578705977754 |
| Cost (with ‘n’ test samples) | 0.689818730212681 | 0.525056743681524 | 0.016844615188822 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10000, n=1000, K=10 | m=10000, n=1000, K=100 | m=10000, n=1000, K=5000 |
| Accuracy (with ‘m’ train samples) | 50.07 | 88.9 | 99.98 |
| Accuracy (with ‘n’ test samples) | 53.300000000000004 | 88.3 | 99.9 |

**[분석]**

학습 진행속도를 위한 alpha 값을 각각 0.01, 0.05, 0.1, 0.5, 0.9로 예측해 본 결과, m=10000, n=1000, k=5000의 환경에서는 0.9가 가장 최적의 알파 값이었음을 알 수 있었습니다.

Train data m, 반복 수 K를 변화시켰을 때 결과는 전반적으로 Practice#1-2와 비슷한 경향을 보였습니다. 레어어 한 개를 가지고 진행했던 저번 실험에서는 m = 10000, n = 1000, K = 5000일 때 Cost는 대략 0.183(test data 기준), 정확도는 95.8999 정도가 나왔습니다.

같은 train data수와 같은 반복 수에서 레이어를 하나 더 증가시킴으로 더 좋은 성능을 낼 수 있었습니다.