**Deep Learning practice #3-1 report**

컴퓨터소프트웨어학부 2018008768 윤정

**[구현 과정]**

텍스트, 모니터, 화면, 플랫이(가) 표시된 사진

자동 생성된 설명

input X와 X의 코사인 값의 부호에 대응되는 {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 train, predict를 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) | 0.6864483820666747 | 0.5948627149325524 | 0.5324120976808924 | 0.5064719270836949 | 0.497657582533408 |
| Accuracy (with ‘n’ test samples) | 65.6000 | 70.7 | 72.2 | 73.7 | 74.2 |

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) | 0.478997625493468 | 0.504260476713385 | 0.495182470432068 |
| Cost (with ‘n’ test samples) | 0.706387573309050 | 0.521914037184811 | 0.499780661611376 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10, n=1000, K=5000 | m=100, n=1000, K=5000 | m=10000, n=1000, K=5000 |
| Accuracy (with ‘m’ train samples) | 70.0 | 77.0 | 74.2299999 |
| Accuracy (with ‘n’ test samples) | 50.1 | 72.5 | 73.4 |

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) | 0.692105283194714 | 0.681870419711400 | 0.495182470432068 |
| Cost (with ‘n’ test samples) | 0.694080612373324 | 0.681434568464428 | 0.499780661611376 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | m=10000, n=1000, K=10 | m=10000, n=1000, K=100 | m=10000, n=1000, K=5000 |
| Accuracy (with ‘m’ train samples) | 50.84999999 | 68.63 | 74.2299999 |
| Accuracy (with ‘n’ test samples) | 47.09999999 | 68.600000001 | 73.4 |

**[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 m을 변화시켰을 때 train data의 크기가 커질수록 train samples와 test samples간의 cost, accuracy 값 차이가 줄어들었습니다.

또한, 학습 반복 수인 K값을 늘리면 늘릴수록 학습의 성능이 향상되는 것을 볼 수 있었습니다.

이번 과제는 sin함수의 결과값을 바탕으로 한 저번 과제보다 성능이 더 좋지 않게 나왔는데, 이는 0~360도 구간에서 cos함수가 양수, 음수가 나뉘는 구간이 더 많아서 학습이 더 어려웠기 때문이라고 판단되었습니다.