



인공지능설계실습1 텀프로젝트

전자공학부 임베디드시스템 전공

2019146037 홍석영

프로젝트 개요

프로젝트 설명

- ✓ 이번 프로젝트는 **시그모이드 함수**를 활성화함수로 하여 **2계층 인공 신경망**을 설계한 후 0,1,2의 **MNIST 데이터**들을 분류 하는 프로젝트입니다.
- ✓ **시그모이드 함수**
: 하나의 연속된 입력값을 0과 1사이의 값으로 변환하는 함수로, 이진 분류에서의 클래스 확률을 표현하거나 하나의 뉴런에서 활성화 여부를 타나내는데 적합합니다.
- ✓ **2계층 인공 신경망**
: 입력층과 출력층 사이에 은닉층이 추가된 2입력 3계층 분류 시스템으로 계층이 추가되면서 속성 공간 분할능력이 향상된 신경망입니다.
- ✓ **MNIST 데이터**
: 손으로 쓴 글자들로 이루어진 대형 데이터베이스이며, 다양한 화성 처리 시스템을 트레이닝 하거나 기계 학습 분야의 트레이닝 및 테스트에 널리 사용됩니다.

프로젝트 개요

프로젝트 방향(모델 설계 과정 - 외부루프)

모델 훈련 과정
(내부루프)

1. 입력 데이터 후보 10개 중 5개를 택한다.
2. 초기 매개변수 및 학습률을 설정한다.
3. 은닉층 노드의 개수(L)를 선택한다.
4. One-Hot Encoding을 통해 출력 변수를 가공한다. (y_q ($q = 0,1,2$)를 만든다.)
5. 위의 옵션을 토대로 2계층 인공 신경망을 만든 후 epoch만큼 훈련시킨다.

6. 최종 정확도 및 MSE를 확인 후 다른 데이터 후보 5개를 재선택 한다.

7. 최종 결과를 비교한 후 결론을 도출한다.



프로젝트 개요

프로젝트 방향(모델 훈련 과정 - 내부 루프)

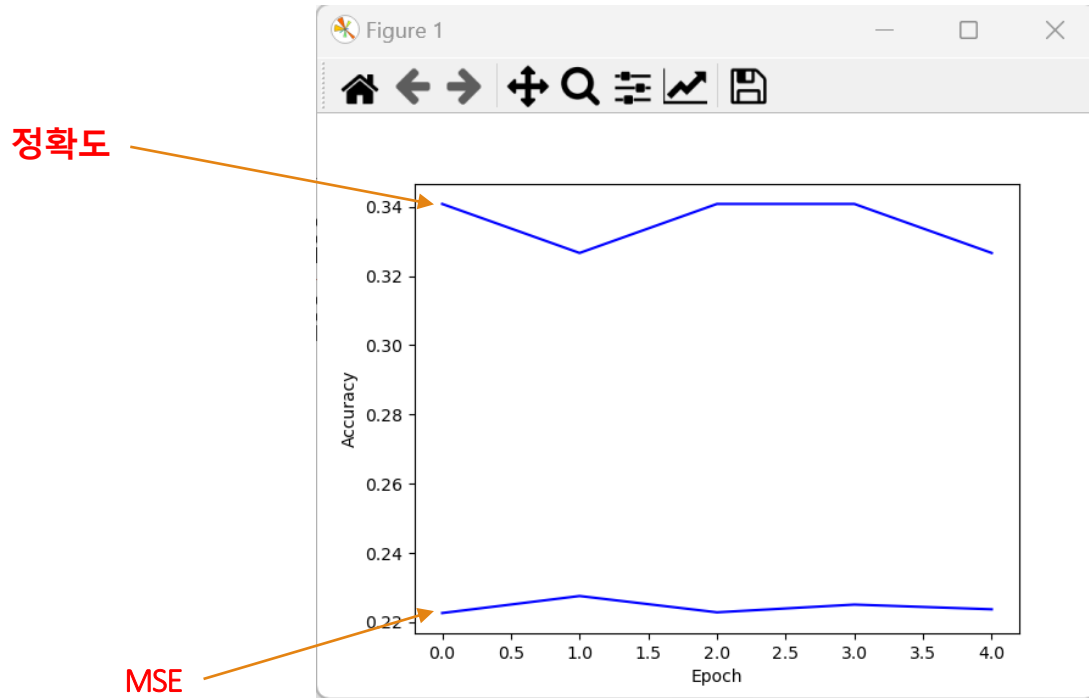
1. (n=0)번째 데이터를 현재 매개변수 값을 바탕으로 예측값 \hat{y}_{qn} 을 구한다.
2. 예측 오차 e_{qn} 을 구한다.
3. 출력층 매개변수에 대한 기울기 $\frac{\partial}{\partial w_{lq}} \in MSE$ 를 계산한다.
4. 은닉층 매개변수에 대한 기울기 $\frac{\partial}{\partial v_{lq}} \in MSE$ 를 계산한다.
5. n 번째 매개변수를 업데이트 한다.
6. n 번째 데이터의 MSE 를 계산한다.
7. 다음 데이터(n = n+1)로 넘어간다.

8. 모든 데이터를 학습시킬 때까지 반복한다.

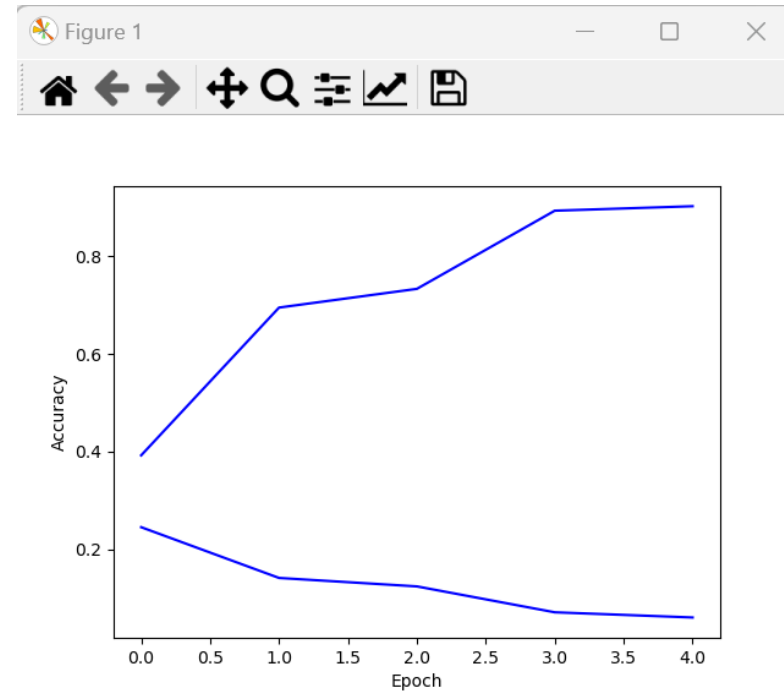
9. 위에서 구한 모든 데이터 MSE 값의 평균을 계산해 최종 MSE값을 구한다.

해결과정

- 매개변수 및 학습률 초기값을 결정하는 단계입니다.
- Epoch를 400번 실행 시키면서 100의 배수 마다 정확도와 MSE를 나타낸 그래프로 아래에 위치한 파란선은 MSE, 위에 위치한 파란 선은 정확도입니다.
- 매개변수와 학습률 초기값에 따라서 **최종 정확도와 MSE 값이 달라짐**을 확인하였습니다.



초기값(매개변수, 학습률)이 나쁜 예



초기값(매개변수, 학습률)이 좋은 예

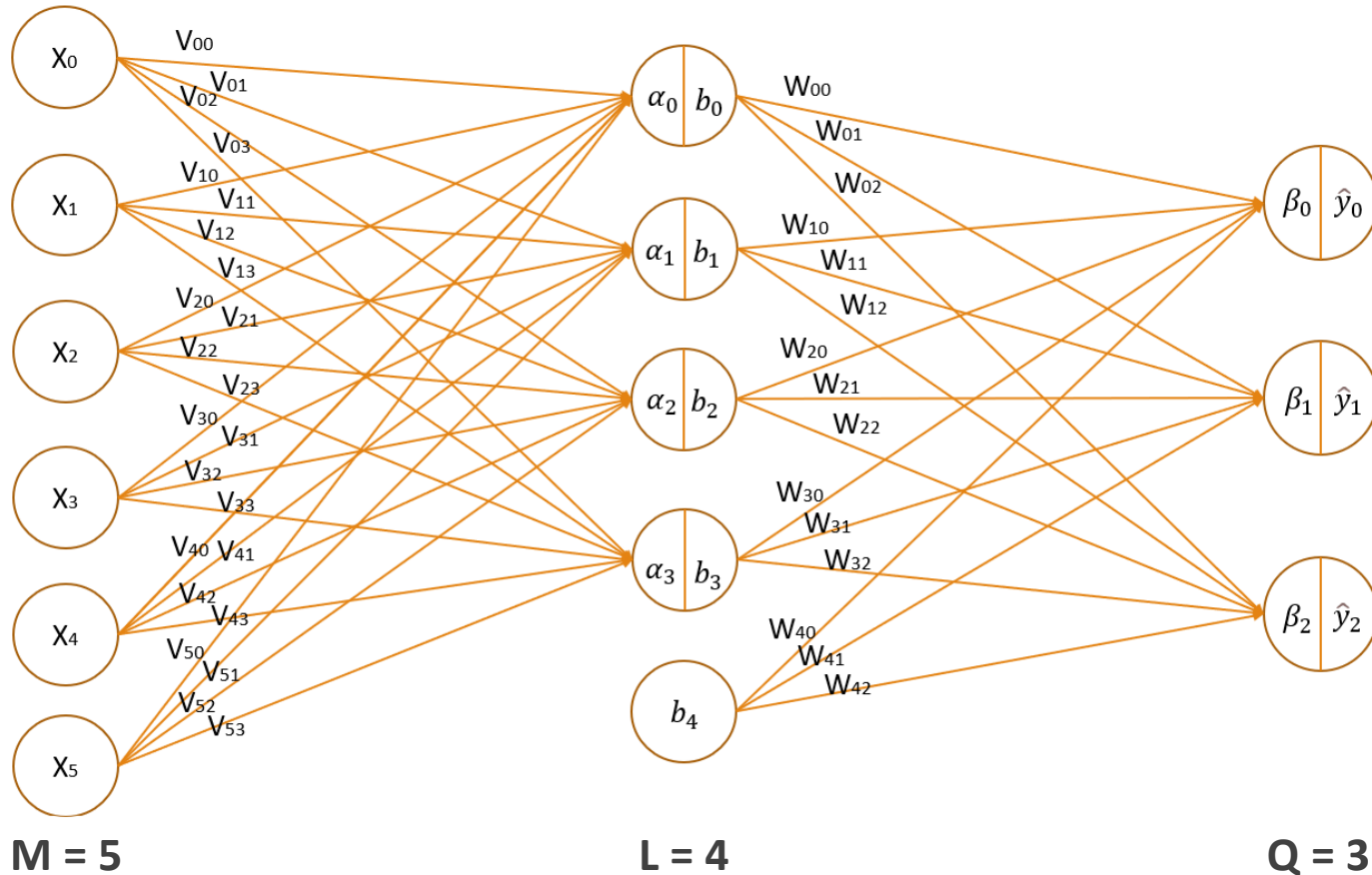
해결과정

프로그램 구조 결정하기

- 은닉층 노드가 많아질 수록 **정확도가 높아지는** 장점이 있지만, 계산량이 많아져서 **복잡해지고**, **과적합**의 문제가 생길 수 있으므로 적절한 은닉층의 개수를 찾는 것이 중요합니다.

그 결과,

- 6개의 입력층 노드(5개 입력노드 + 1개 더미노드)
5개의 은닉층 노드(4개 은닉노드 + 1개 더미노드)
3개의 출력층 노드로 구성된 2계층 인공신경망 구조를 선택하였습니다.



해결과정

최종 모델로 선택



- 입력 후보 10개 중 5개를 택하는데 가장 잘 훈련되는 조합을 찾는 과정입니다.
- 매개변수는 랜덤값으로 10번 돌려서 가장 학습이 잘된 매개변수, 학습률은 0.0001, epoch는 400번씩 훈련하여 모델을 성능을 측정하였습니다.
- 초기 매개변수 값에 따라 정확도가 달라질 수도 있지만, 실습 결과를 통해 추론해본 바로는 클래스 마다 각각의 값 차이가 많이 날수록 즉, '**클래스 마다의 특징이 두드러지는 입력 후보를 사용할 수록 MSE값이 감소된다**' 라는 결론이 나왔습니다.

입력 종류	1,3,4,7,10	2,3,4,7,10	2,4,6,7,9	2,4,6,9,10
매개변수	Random값	Random값	Random값	Random값
Learning rate	0.0001	0.0001	0.0001	0.0001
Epoch	400	400	400	400
정확도	74.8%	67.8667%	89.25%	89.4667%
MSE	0.366562	0.420763	0.198916	0.178346

결과

- 입력 속성

- 2. 가로축 Projection => 확률밀도함수로 변환 => 분산
- 4. 세로축 Projection => 확률밀도함수로 변환 => 분산
- 6. Diagonal 원소 배열 추출 => 확률밀도함수로 변환 => 분산
- 9. Anti-Diagonal 원소 배열 추출 => 확률밀도함수로 변환 => 분산
- 10. Anti-Diagonal 원소 배열 추출 => 0의 개수

- 프로젝트 결과

위에서 선택한 옵션들을 바탕으로 2계층 인공신경망 모델을 학습시킨 결과

최종 정확도는 **89.4667%** , 최종 MSE는 **0.178346**의 좋은 결과가 나왔습니다.

따라서 우리는 모델이 잘 훈련되었다고 볼 수 있고, 이를 통해서 훈련되지 않은 다른 0,1,2의 MNIST 데이터들을 모델에 적용시켜도 분류할 수 있을 것입니다.

- 은닉층 개수

5개

- Epoch와 학습률

Epoch = 400 / 학습률 = 0.0001

분석

2계층 인공신경망 모델에 영향을 주는 조건은 무엇일까?

1. 초기 매개변수와 학습률

- 홀드아웃, K겹 교차검증과 같이 훈련데이터와 검증데이터를 통하여 모델을 검증해 좋은 매개변수와 학습률을 찾습니다.
- 값들을 변경시키며 수 없이 많은 테스트 결과를 바탕으로 좋은 매개변수와 학습률을 찾습니다.

2. 2계층 인공 신경망 구조 중 은닉층 노드의 개수

- 은닉층 노드의 개수가 많아지면 정확도는 높아질 수 있지만, 과적합의 문제가 생겨 훈련데이터에서의 정확도만 올라갈 수 있습니다. 따라서 너무 많지도 않고, 너무 적지도 않은 적절한 개수를 선택하여야 합니다.

3. 2계층 인공 신경망 구조 중 입력 속성의 특징

- 입력속성을 통해 인공 신경망 모델의 매개변수가 결정됩니다. 따라서 분류가 잘 되는 입력속성을 선택하여야 합니다.
즉, 각 클래스마다 특징이 뚜렷하게 나타나 입력 속성 데이터를 통해 클래스 분류가 잘 될 경우 MSE값이 작은 정확한 모델을 얻을 수 있습니다.

결론

2계층 인공신경망은 선형분할 뿐만 아니라 비선형 분할도 가능하여 유용한 구조입니다. 하지만, 단순히 2계층 인공신경망 모델을 만들고, 훈련 시켰을 때 항상 학습이 잘 되어 분류가 잘 되는 것은 아닙니다.

‘epoch 및 학습 반복 횟수가 과도할 때 혹은 적을 때 과적합 및 과소적합이 생기는 문제’,
‘매개변수 및 학습률 초기화 선택을 잘못 하였을 경우 제대로 분류가 되지 않는 문제’,
‘입력 데이터의 속성 값이 출력 클래스마다 비슷한 값이 나와 제대로 분류가 되지 않는 문제’
등 여러가지 문제가 발생할 수 있습니다.

이를 해결하기 위해서는 ‘위의 조건 값들을 변경시키며 수 많은 테스트 결과를 바탕으로 좋은 초깃값을 얻는 방법’, ‘홀드 아웃, k겹 교차검증 과 같이 분류데이터와 훈련데이터를 나누어 모델을 검증하여 얻는 방법’ 등이 존재 합니다.

위의 방법을 사용하여 모든 문제를 해결 하였을 때 우리는 분류가 잘 되는 최상의 모델을 얻을 수 있습니다.