

다층 퍼셉트론 신경망을 이용한 숫자 분류기 설계 방식 분석 및 비교

김세송, 김동욱, 정승원
동국대학교 멀티미디어공학과
e-mail : shak@dongguk.ac.kr

Analysis and Comparison of Numeral Classifiers Based on the Multilayer Perceptron

Se-Song Kim, Dong-Wook Kim, and Seung-Won Jung
Dept. of Multimedia Engineering, Dongguk University

요 약

숫자 인식 분야는 인식 분야에서도 오래된 분야이며 다양한 방법이 제시되어 있는데, 그 중 다층 퍼셉트론 신경망을 이용한 숫자 분류기에 대한 비교 분석을 수행한다. 특히 복잡한 문제를 여러 개의 단순한 문제로 나누는 방식의, 각 숫자에 대한 독립적인 분류기를 설계하는 방식에 대하여 분석을 수행한다. 일반적인 하나의 분류기로 전체 숫자를 분류하는 방식과의 비교를 통하여 숫자 분류에는 각 숫자에 대한 독립적인 분류기를 이용하는 것이 적합하다는 사실을 실험적으로 확인하였다.

1. 서론

이미지 인식 분야에서 숫자 이미지를 인식하는 것은 오랜 시간 동안 연구돼 온 분야이며 인식기를 설계하는 방법도 여러 가지가 존재한다. 예를 들어, 픽셀 값 기반 분류[1]이나 패턴을 통한 분류, 신경망 학습을 이용한 분류 등이 있다. 특히 신경망 이론 중 하나인 다층 퍼셉트론 신경망은 최근 화제가 되고 있으며 다양한 문제를 해결하기 위한 효과적인 신경망 이론이라는 평가를 받고 있다.

다층 퍼셉트론이란 인공 신경망 이론인 퍼셉트론이 가진 비선형분리의 한계를 은닉층과 오류 역 전파 알고리즘으로 해결한 인공 신경망이다[3][4][5]. 다층 퍼셉트론은 3 개 이상의 층으로 구성되며 각 층은 하나 이상의 노드를 가진다. 이것은 입력 층, 은닉 층, 출력 층을 가지며 입력 층과 출력 층은 각각 1 개의 층을 가진다. 은닉 층은 여러 층을 가질 수 있다. 본고에서는 다층 퍼셉트론을 이용한 숫자 분류기에 대한 실험 분석을 수행한다.

2. 분류기의 설계

일반적인 다층 퍼셉트론을 이용한 숫자 분류기의 경우 1 개 이상의 은닉 층을 가지며 출력 층은 10 개의 출력 노드를 가진다. 각 출력 노드의 시그모이드 함수 값이 분류를 나타낸다[2]. 이러한 방식으로 분류기를 설계할 경우 높은 분류 정확도를 위해서 많은

은닉 층이 필요하며 학습의 복잡도가 크다.

다층 퍼셉트론은 단층 퍼셉트론이 가지는 비선형분리의 한계를 해결한 인공신경망이다. 이러한 비선형분리는 은닉 층을 통해 가능해지기 때문에 복잡한 분류일수록 더 많은 은닉 층을 필요로 한다. 그러나 은닉 층이 추가될 때마다 은닉 층의 노드의 개수 배 만큼 계산복잡도가 증가하기 때문에 은닉 층이 많아질수록 계산 속도는 크게 떨어진다. 또한 출력 노드가 많은 경우 오류 역 전파를 이용한 학습이 다른 요인에 영향을 미쳐 학습 속도를 저하시킨다.

이 두 가지 문제를 해결하기 위하여 은닉 층을 증가시키는 대신 문제 자체의 복잡도를 분산시킨다면 그 분류기는 상대적으로 적은 계산량을 가지게 된다. 또한 각각의 분류기가 하나의 출력 노드를 가지게 한다면 학습의 효율을 높일 수 있다. 그러므로 본고에서는 아래의 두 가지 방식의 숫자 분류기 설계에 대한 비교를 수행한다.

- 10 개의 출력 노드를 비교하여 숫자를 분류하는 하나의 분류기
- 각각 1 개의 출력 노드를 가져 각각의 출력 값을 비교하여 분류하는 10 개의 분류기

3. 실험

3.1 한 개의 분류기를 이용한 설계

은닉 층이 하나로 입력 층 - 은닉 층 - 출력층 층

3 개의 층을 가지는 다층 퍼셉트론 신경망을 구현하였다. 숫자 학습을 위하여 mnist number dataset 을 이용하였다(mnist number dataset 은 28*28 픽셀 사이즈의 255 단계의 gray scale 을 가진 숫자 학습용 dataset 으로 숫자 인식 학습 및 테스트 분야에 쓰이는 dataset 이다). 입력 층은 mnist dataset 을 입력 받기 위한 784(=28*28)개의 노드와 학습을 위한 bias 노드 1 개를 추가로 가져 785 개의 노드를 가진다. 은닉 층의 경우에는 [6]에서 제시한 노드의 개수(300 개 + bias 노드 1 개)를 지정하였다. 출력 층은 10 개의 노드를 가진다. 각각의 출력노드는 0~9 의 값을 의미한다. 이 중 단 한 개의 노드의 결과 값이 1 이 나오고 나머지는 모두 0 이 나오는 경우를 label(학습을 위한 정답)으로 주었다.

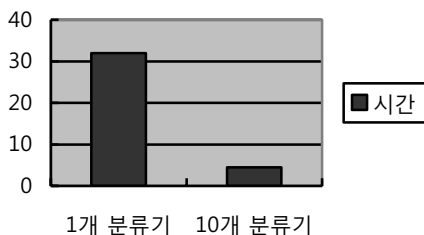
이렇게 설계된 분류기를 mnist training dataset 60000 장의 숫자 이미지를 1000 회 학습시킨 경우와 10000 회 학습시킨 경우 mnist test dataset 10000 장의 이미지를 분류했을 때 각각 81%(학습 시간: 약 4 시간)과 96.15%(학습 시간: 약 36 시간)의 정확도를 얻을 수 있었다

3.2 숫자별로 독립적인 분류기를 이용한 설계

입력 층 노드의 개수와 은닉 층의 수는 3.1 과 같다. 10 개의 분류기는 실험적인 결과를 통해 각각 101 개 (bias 노드 1 개)의 은닉 노드를 가지게 하였다. 10 개의 분류기의 출력 층이 가진 각각의 출력 노드가 가진 시그모이드 함수 결과값을 비교하여 가장 큰 값을 가지는 노드의 분류기가 곧 입력된 이미지의 숫자를 의미한다[2].

3.1 과 같은 dataset 을 이용하여 각각 100 회(10*100 = 총 1000 회)씩 학습시킨 후 test data 10000 장의 이미지를 분류하자 98.13%의 정확도를 얻을 수 있었다(학습 시간: 약 5 시간). 그 이상의 학습 횟수에도 정확도의 큰 변화는 없었다.

두 방법은 시간과 정확도에서 차이를 보였다. 같은 횟수의 학습일 경우(1000 회), 분류기가 하나인 방법의 학습 시간이 독립 분류기를 사용하는 방법에 비해 4/5 정도 걸렸으나 정확도가 17% 이상 차이가 났다<그림 1>. 정확도를 높이기 위하여 36 시간 정도 학습을 수행하였을 때에도 독립 분류기에 비해 약 2% 낮은 정확도를 보였다<그림 2>.



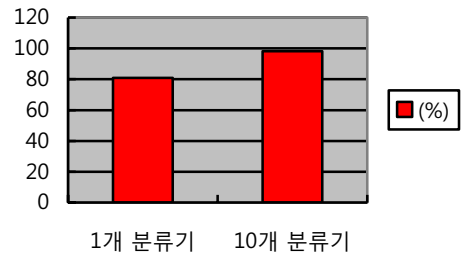
<그림 1> 학습 시간 비교(두 방법 모두 95% 이상의 정확도

를 가지는 분류기를 학습할 때)

4. 결론

본 논문은 다층 퍼셉트론 신경망을 이용한 숫자 분류기에 대해 새로운 접근 통해 더 빠른 학습과 높은 정확도를 가지는 설계를 제안하였다.

하나의 분류기를 이용하는 방식의 경우 학습률이 현저하게 떨어지기 때문에 결과적으로 독립적인 분류기를 사용하는 방법보다 약 6 배 이상의 시간을 학습시켜야 비슷한 정확도의 분류기를 얻을 수 있었다. 이와 같이 복잡한 문제를 단순한 문제로 분산하여 따로 해결하는 방식의 다층 퍼셉트론 신경망 학습은 문자 인식이나 사물 인식 등 다양한 분야에서 이용될 수 있다.향후 연구 계획으로는 은닉 노드의 개수의 최적화 알고리즘 도입과 cascading 을 이용하여 계산량을 최소화할 예정이다.



<그림 2> 정확도 비교(traing dataset 을 각각 1000 회 학습시켰을 때)

참고문헌

- [1] 전소연, 김수진, 박구만 “성능 개선을 위한 픽셀 값 기반 글자 인식 알고리즘”, 한국통신학회 2015년도 추계종합학술발표회
- [2] Cybenko, G. 1989. Approximation by superpositions of a sigmoidal function *Mathematics of Control, Signals, and Systems*, 2(4), 303–314.
- [3] Rumelhart, David E., Geoffrey E. Hinton, and R. J. Williams. "Learning Internal Representations by Error Propagation". David E. Rumelhart, James L. McClelland, and the PDP research group. (editors), *Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1: Foundations*. MIT Press, 1986.
- [4] Haykin, Simon (1998). *Neural Networks: A Comprehensive Foundation* (2 ed.). Prentice Hall. ISBN 0-13-273350-1.
- [5] Neural networks. II. What are they and why is everybody so interested in them now?; Wasserman, P.D.; Schwartz, T.; Page(s): 10-15; IEEE Expert, 1988, Volume 3, Issue 1
- [6] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: Gradient-Based Learning Applied to Document Recognition, *Proceedings of the IEEE*, 86(11):2278-2324, November 1998, \cite{lecun-98}.