

신경망 모델의 은닉층에 관한 연구*

오성빈 임현일
 경남대학교 컴퓨터공학부
 e-mail : ginbeat21@gmail.com

A Study on Hidden Layer for Neural Network Model

Sung-Bhin Oh Hyun-il Lim
 Dept. of Computer Engineering, Kyungnam University

요 약

딥러닝에서는 기본 층을 겹겹이 쌓아 올려 구성된 신경망이라는 모델을 사용하여 데이터를 학습한다. 신경망 모델에서 층(layer)은 신경망의 핵심 구성 요소로서 입력된 데이터로부터 주어진 문제에 더 의미 있는 표현을 추출하고 표현할 수 있다. 이러한 층의 개수와 층 내에 노드의 개수는 신경망 설계에서 가장 기본적인 문제 중에 하나이다. 본 논문에서는 층의 개수와 노드의 개수가 신경망 학습에 어떠한 영향을 미치는지 실험을 통하여 평가해본다.

1. 서론

딥러닝은 머신 러닝의 특정한 한 분야로서 연속된 층에서 점진적으로 의미 있는 표현을 배우는 데 강점이 있는, 데이터로부터 표현을 학습하는 새로운 방식이다. 딥러닝의 딥이란 단어는 연속된 층으로 표현을 학습한다는 개념을 나타내며, 데이터로부터 모델을 만드는 데 얼마나 많은 층을 사용했는지가 그 모델의 깊이가 된다[1].

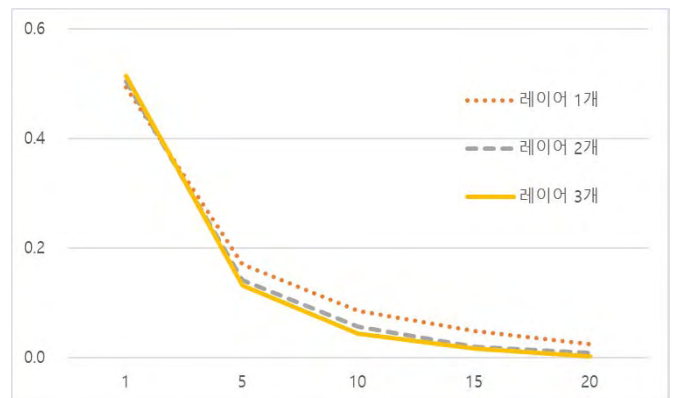
이러한 딥러닝의 층은 여러 개의 인공 뉴런으로 구성되어 있으며 이 각각의 인공 뉴런을 노드라고 부른다. 인공 뉴런인 노드는 입력 값을 받아 가중치를 곱하고 활성화 함수를 사용하여 예측 값을 출력한다. 그리고 신경망의 손실 함수를 통해 예측 값과 실제 값의 차이를 점수로 계산한다. 딥러닝은 이 점수를 피드백 신호로 사용하여 손실 점수가 감소되는 방향으로 가중치 값을 조금씩 수정하는 것으로 이런 일련의 과정들을 ‘학습’이라고 부른다.

본 논문에서는 신경망의 설계에서 중요한 부분을 차지하는 층의 개수와 층 내에 노드의 개수가 학습에 어떠한 영향을 미치는지 실험을 통하여 평가해보았다.

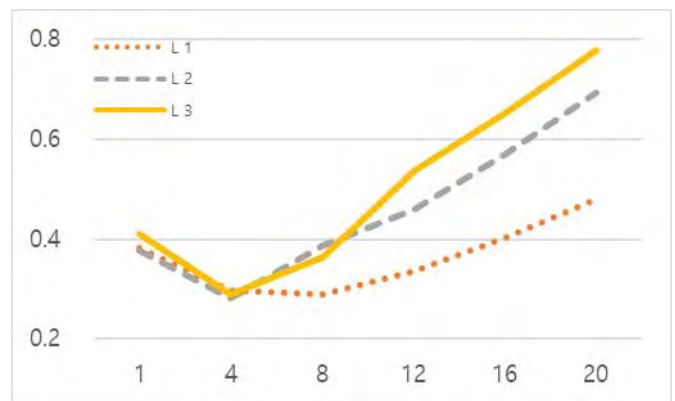
2. 신경망 층의 개수에 따른 영향

본 실험에서 층의 개수에 따라 학습의 결과가 어떻게 달라지는지 실험해 보았다. 실험은 TensorFlow[2]를 이용하여 IMDB 데이터셋에 대해 영화 리뷰를 긍정 또는 부정으로 분류하는 모델을 이용해 진행하였다. 모델의 층 내에 노드의 개수는 16 개로 동일하며, 원본 훈련 데이터에서 10,000 개의 샘플을 떼어 검증

세트를 만들어 검증 샘플 데이터에서 손실과 정확도를 계산하였다.



(그림 1) 층 개수에 따른 훈련 손실 변화

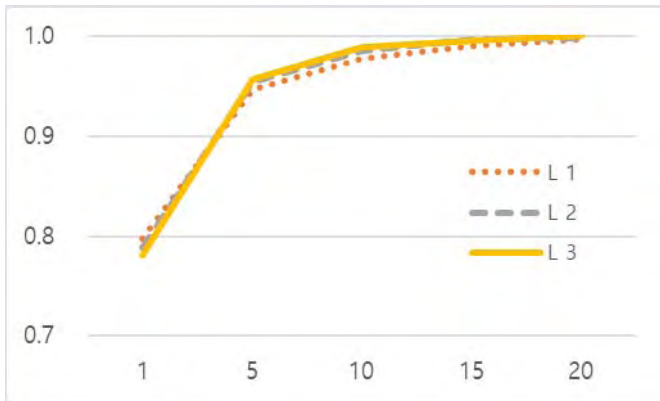


(그림 2) 층 개수에 따른 검증 손실 변화

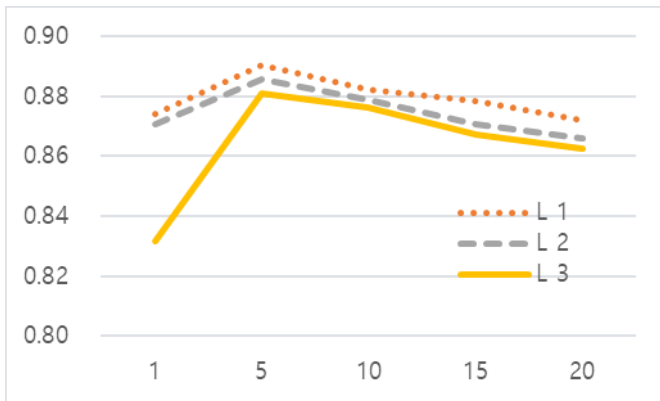
각각 층 개수를 1, 2, 3 개로 하여 학습을 진행한 결

* 이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2017R1D1A1B03034769).

과 그림 1 에서 보이듯이 3 종류의 모델 모두 훈련 데이터에 대해선 손실이 감소하여 0 에 가까워졌지만 그림 2 에서처럼 검증 데이터에 대해선 층의 수가 많을수록 손실이 증가하는 모습을 보였다.



(그림 3) 층 개수에 따른 훈련 정확도 변화

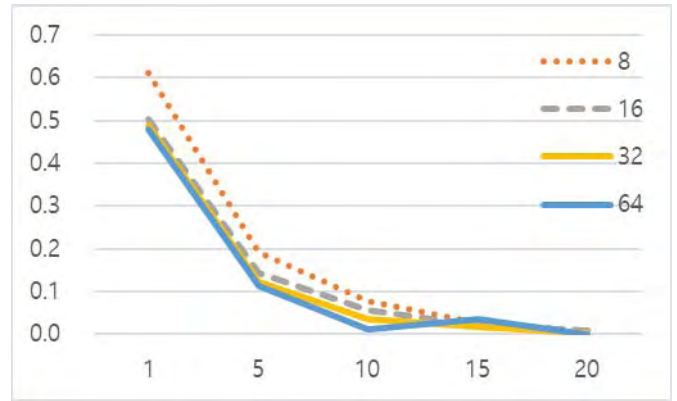


(그림 4) 층 개수에 따른 검증 정확도 변화

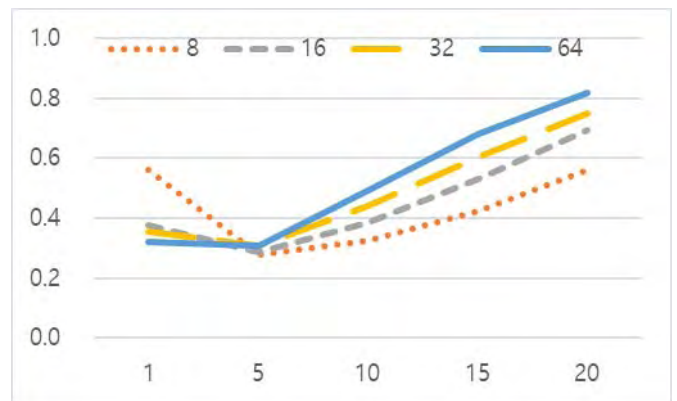
정확도의 경우에는 그림 3 에서와 같이 훈련 데이터에 대해선 모두 1 에 가까워지는, 완벽하게 학습된 모습을 보였지만, 그림 4 에서는 검증 데이터에 대해선 네 번째 에포크 이후에는 과적합이 시작되어 정확도가 정체되고 미미하게 감소되는 모습을 보였다.

3. 층의 노드 개수에 따른 영향

노드 개수에 관한 실험도 동일한 환경에서 수행하였다. 층의 개수는 2 개로 동일하게 유지하고, 노드의 개수를 8, 16, 32, 64 로 조정하여 개수에 따라 어떤 변화가 있는지 실험하였다.

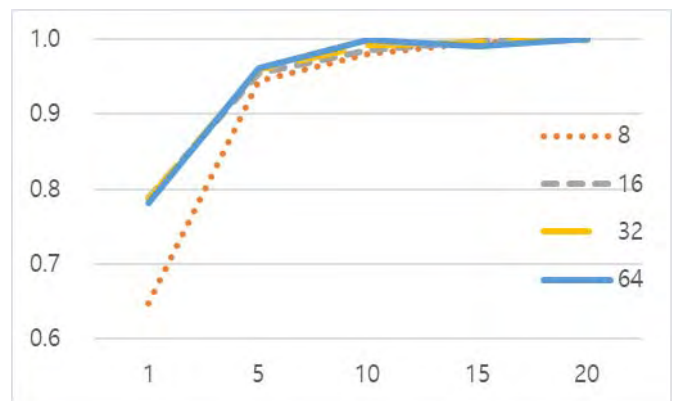


(그림 5) 노드 개수에 따른 훈련 손실 변화

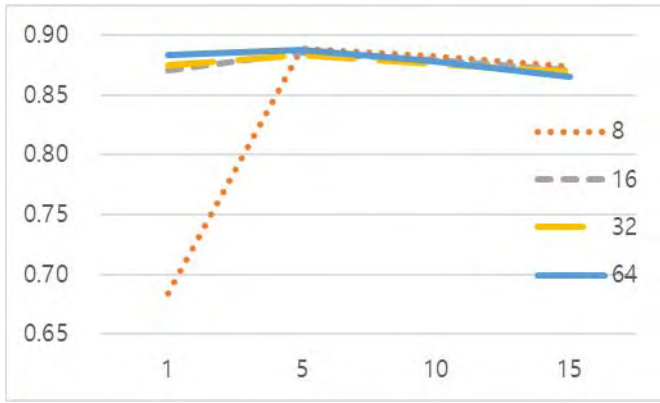


(그림 6) 노드 개수에 따른 검증 손실 변화

학습 결과, 그림 5 에서 보이는 바와 같이 훈련 데이터에 대한 손실은 0 에 가까워지며 손실값을 줄이며 학습이 되고 있음을 확인할 수 있다. 그림 6 에서 노드 개수에 따른 검증 손실 변화에서는 검증 데이터에 대해 학습이 진행 될수록 과적합에 따른 오류로 인해 손실값이 상승하였으며, 노드의 개수가 많을 수록 과적합의 영향이 크다는 것을 알 수 있다.



(그림 7) 노드 개수에 따른 훈련 정확도 변화



(그림 8) 노드 개수에 따른 검증 정확도 변화

정확도 역시 그림 7, 8 에서 확인할 수 있듯이 훈련 데이터에 대해선 정확도가 1 에 근접해 가지만 검증 데이터에 대해선 네 번째 에포크 이후에 과적합이 시작되어 정확도가 정체되며 떨어지는 모습을 보여주고 있다. 특히 노드의 개수가 많을수록 과적합의 영향을 많이 받게 돼서 정확도의 감소폭이 더 컸다.

4. 실험 결과

실험 결과, 층의 개수가 많을수록, 층 내에 노드의 개수가 많을수록 모델은 훈련 데이터에 대해선 손실이 적고 정확도가 높게 나오며 더욱 타겟 값에 가깝게 예측 값을 출력하는 모습을 보인다. 하지만 검증 데이터에 대해선 그러한 모습을 보이지 않고 있다.

손실은 타겟 값과 모델의 예측 값 간의 차이인데 층의 개수와 노드의 개수가 많을수록 검증 데이터에 대한 손실이 더 크게 증가하는 것을 그림 1, 2, 5, 6 에서 확인 가능하다. 이는 신경망 모델들이 과적합되었음을 뜻한다. 훈련 데이터에 과도하게 최적화되어 훈련 데이터에 특화된 표현을 학습하므로 훈련 세트 이외의 데이터에는 잘 작동하지 않는 것이다. 따라서 층의 개수와 노드의 개수가 많을수록 훈련 데이터를 더 잘 학습하게 되어 과적합되는 정도가 더 세지면서 손실이 늘어나는 것이다.

정확도의 경우는 그림 3, 4, 7, 8 에서 확인할 수 있듯이 훈련 데이터에 대해선 학습을 반복할수록 1 에 가까워지지만 검증 데이터에 대해선 네 번째 에포크 이후부터 정확도가 상승하지 않고 정체되는 모습을 보인다. 이 또한 과적합이 발생한 결과로 훈련 데이터에만 최적화되어 검증 데이터에 대해선 정확한 예측 값을 출력하지 못하는 것이다.

또한, 이진 분류 모델을 사용하였기 때문에 정확히 예측했던 값은 계속 정확하게 예측하고 그러지 못한 경우는 예측 값과 타겟 값 간의 차이가 계속 늘어나서 검증 데이터에 대해 손실은 계속 증가하지만 정확도는 감소하지 않고 정체되어있다.

본 실험을 통해 노드의 개수 및 층의 개수는 문제를 해결하기 위한 적정 크기를 유지하는 것이 좋다는 것을 알 수 있다. 필요 이상으로 노드의 수 또는 층의 수를 늘리면 과적합으로 인해 오히려 성능을 떨어

뜨릴 수 있다는 것을 확인할 수 있었다.

5. 결론

본 논문의 실험을 통해 층의 개수와 층 내에 노드의 개수가 증가할수록 훈련 데이터에 과도하게 최적화되어 다른 데이터에는 제대로 작동하지 않을 수 있다는 것을 확인할 수 있었다. 이러한 과적합 문제는 머신 러닝의 성능에서 매우 중요한 역할을 하며, 과적합 문제를 어떻게 다루는지에 따라 좋은 모델과 그렇지 않은 모델을 나누게 될 것이다.

따라서 좋은 신경망 모델을 만들기 위해선 모델에 맞는 층의 개수와 층의 노드 개수를 찾는 것이 중요한 요소라고 할 수 있다.

감사의 글

이 논문은 2017 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2017R1D1A1B03034769).

참고문헌

- [1] François Chollet. “Deep Learning with Python”, Manning Publications, December 22, 2017.
- [2] TensorFlow. <https://www.tensorflow.org/>