

# 금융과인공지능 4차과제

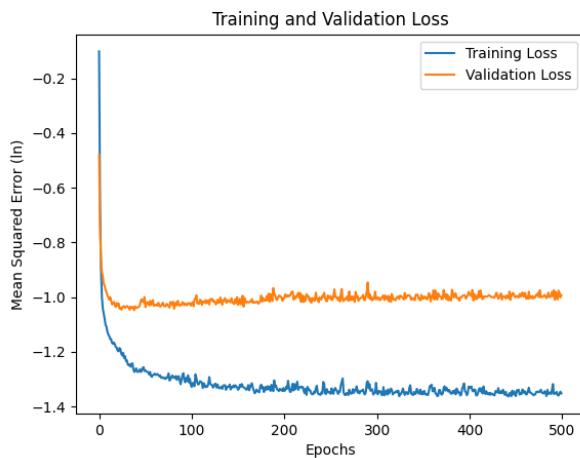
2021510010 이상민

1. (보고서 내용 없음)

2. (보고서 내용 없음)

3. 학습 및 검증

3-1) 단일은닉층 ANN에 대한 학습 및 손실 그래프는 아래의 <그림 1>과 같다.

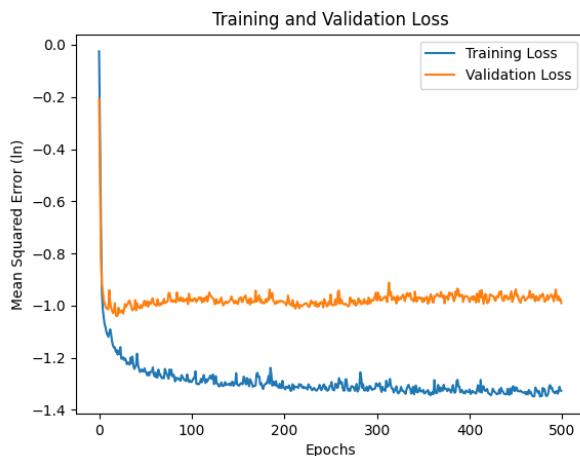


<그림 1. 단일은닉층 ANN Loss 그래프>

3-2) 단일은닉층 ANN의 결과 Epoch 32에서 검증손실 0.0923381820321083으로 가장 낮은 검증손실을 기록했다.

4. 딥러닝 모형

4-1) DNN모형에 대한 학습 및 손실 그래프는 아래의 <그림 2>와 같다.



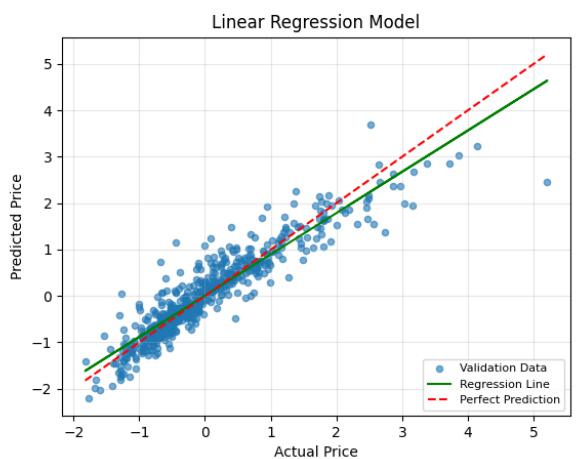
<그림 2. 은닉층 3개의 DNN Loss 그래프>

또한, 3개의 은닉층 DNN의 결과 Epoch 20에서 검증손실 0.09097807854413986으로 가장 낮은 검증손실을 기록했다.

4-2) 본 결과에서 DNN은 ANN 대비 미미한 수준의 검증 성능을 향상 시켰다. 하지만, 복잡도 증가가 검증셋 오차를 의미 있게 감소시키지 않았다. 이는 모델 복잡도를 단순히 올리는 것보다, 데이터 특성에 맞는 적절한 모델 구조와 정규화, 하이퍼파라미터 최적화 전략이 필요함을 시사한다. 즉, 복잡한 모델은 잠재적 표현력을 커질 수 있으나, 그것이 항상 검증 성능 향상으로 이어지는 것은 아니라는 점을 확인할 수 있다.

## 5. 선형회귀모형

5-1) Linear Regression을 사용하여 훈련 데이터를 학습하고, 학습된 모델로 검증 데이터를 이용하여 검증을 진행했다. 이 때 훈련 데이터의  $MSE = 0.10947956921858988$ , 검증 데이터의  $MSE = 0.1338803794854407$ 로 기록됐다. 그리고 Linear Regression을 사용한 시각화 결과는 아래 <그림 3>과 같다.



<그림 3. 선형회귀모형 검증 데이터 시각화 결과>

5-2) 검증셋에 대한 ANN lowest valid loss는 약 0.0927, DNN lowest valid loss는 약 0.0939, Linear Regression MSE는 약 0.1339이다.

먼저, 각 모델 별 성능을 분석해보면 선형 회귀 모델은 0.1339의 검증 MSE를 기록하여 가장 높은 오차를 보였다. 이는 선형 회귀가 데이터의 비선형적 관계를 충분히 포착하지 못했음을 시사한다. 데이터가 선형적 특성만으로 설명되기 어려운 경우, 선형 회귀 모델의 성능은 제한 적일 수 있다.

ANN은 은닉층 1개(5개 노드)로 구성된 신경망으로, Validation Loss가 0.0927로 선형 회귀보다 현저히 낮은 오차를 보였다. 이는 신경망 모델이 선형 회귀보다 더 복잡한 데이터 관계를 학습하고 비선형성을 잘 포착했음을 보여준다. ANN모델은 과적합 없이 적절히 학습되었으며, 간단한 구조만으로도 우수한 성능을 달성했다.

DNN은 은닉층 3개(각 5개 노드)로 구성된 모델로 Validation Loss가 0.0939로 나타났다. 이는 ANN과 비교했을 때, 약간 더 높은 MSE를 기록하여 큰 성능 향상을 보이지 못했다. DNN은 이론적으로 더 복잡한 데이터 패턴을 학습 할 수 있는 잠재력을 가지지만, 실제 성능은 ANN과 유사한 수준에 머물렀다. 이는 주어진 데이터 특성 및 하이퍼파라미터 설정 하에서 추가적인 복잡도가 성능 향상에 크게 기여하지 않음을 나타낸다.

결론적으로, ANN은 가장 낮은 Validation. MSE를 달성하며, 가장 효율적인 모델로 평가된다. DNN은 ANN과 비슷한 성능을 보였지만, 복잡도를 증가시킨 것에 비해 추가적인 성능 향상이 미미했다. 선형 회귀는 데이터의 비선형적 특성을 반영하지 못하여 성능이 가장 낮았다. 이에 IOWA Price데이터는 ANN모델을 사용하여 분석을 진행하는 것이 가장 적절할 것으로 판단된다.