**Abstract**

본 과제에서는 RNN을 활용하여 100개의 주식 데이터에 대해서 각각 미래의 20 business days에 대해서 예측을 수행하였다. 과거의 50 business days 주식 종가 정보를 Input으로 활용하였디. RNN을 통해 미래의 20 business days 주식 종가 정보를 예측하는데, 좋은 성능을 얻을 수 있었다.

1. **Introduction**

본 과제에서는 100개의 주식 데이터에 대해서 각각 미래의 20 business days에 대해서 예측을 하고자 한다. 이 때, 딥러닝 모델인 RNN을 활용하여 예측의 정확성을 최대한 높이고자 한다. 100개의 주식 데이터는 각각 12년 (3,000 business days)의 학습 데이터를 보유하고 있다. 데이터에 저장된 값은 주식의 종가 정보이다.

**그림 1.**

1. **Experiment**
2. **Data Preprocessing**

주식마다 가격에 대한 상대적인 크기가 다르며, 또한 변동에도 큰 차이가 있으므로, Min-max 정규화를 사용하였다. 또한 RNN의 Input으로 활용하기 위해서 특정 시퀀스를 설정하여, 과거의 몇 개의 종가 정보로 미래의 몇 개 종가 정보를 예측할 지를 정해야 한다. 본 과제에서는 과거 50 business days를 Input으로 하여, 미래의 20 business days를 예측하도록 한다. Train 데이터는 window를 1로하여, 각 주식별로 2931개의 학습데이터가 생성된다.

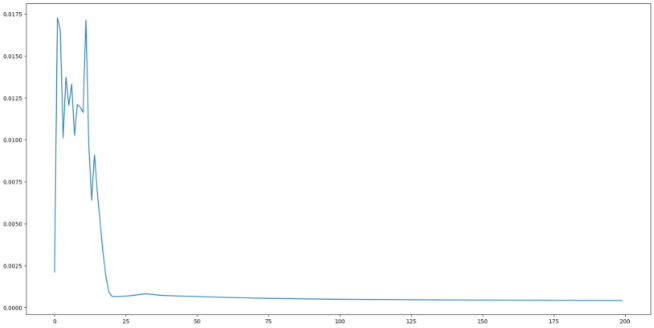
1. **Models**

Python의 Pytorch로 RNN 모델을 구성하였다. RNN Layer는 2개로 구성하였으며, hidden size는 8이다. 또한 2개의 RNN Layer를 통과한 후, FC Layer를 통해 20 business days를 예측할 수 있도록, 20차원의 output을 생성한다.

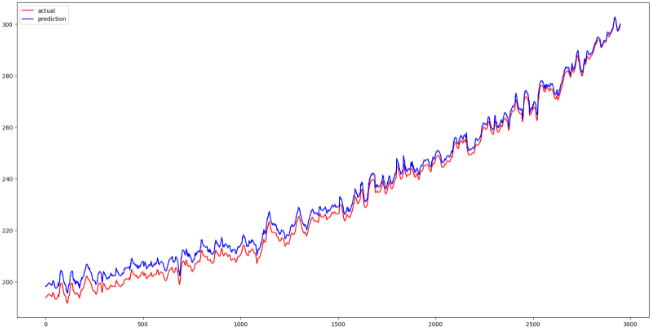
마지막으로 고정된 파라미터는 다음과 같다. Batch size: 16, Epoch: 200, Initial Learning Rate: 0.001Loss: Mean Squared Error, Activation: ReLU

1. **Result**

첫 번째 주식에 대한 학습 과정은 그림 1과 같다. 약 20 Epoch부터 어느정도 낮은 Loss에 수렴한다. 실제 학습 데이터에 대한 정규화한 상태의 MSE는 0.0007이며, 역정규화 후 MSE는 9.1337이다. 그림 2는 역정규화 후 실제 값과 예측 값을 보여준다.



**그림 1.**



**그림 2.**

1. **Conclusion**

본 과제에서는 RNN을 활용하여 100개의 주식 데이터에 대해서 각각 미래의 20 business days에 대해서 예측을 수행하였다. 과거의 50 business days 주식 종가 정보로 미래의 20 business days 주식 종가 정보를 예측하는데, 좋은 성능을 얻을 수 있었다.