

LFD Project#4: Stock trading using RL

Team 9: 김 현, 김 석, 이 은, 한창진

1. Objective and Dataset

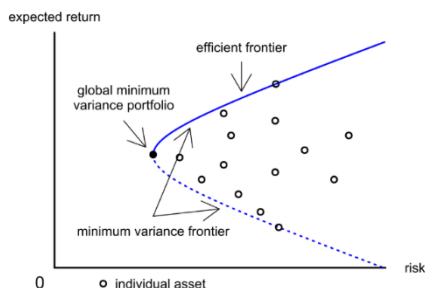
본 프로젝트의 목표는 주어진 과거 주식 데이터로 RL을 이용한 모델을 학습시켜 특정 기간 동안 거래를 진행하면서 수익을 최대화하는 것이다. 주어진 데이터셋은 2010.01.04~2020.05.19까지의 KOSPI 시가총액 상위 11개 기업의 주가와 KOSPI 지수를 포함하고 있으며 각 종목은 Date, Open, High, Low, Close, Volume의 정보를 가지고 있다. Training set은 2018.01.01~2020.04.01 기간, Test set은 2020.05.06 ~ 2020.05.19 (10일) 기간의 데이터를 활용하였으며 두 기간 사이를 Validation set으로 두어 최종 모델을 선정하였다.

2. Feature preprocessing

우선 주어진 데이터를 살펴봤을 때 결측값은 없었고 volume 값에 있는 K를 포함한 문자열을 제거하고 정수형으로 변환해주었다. 그리고 네이버 같은 경우는 액면분할을 하면서 가격 데이터가 0으로 채워져 있는 구간이 존재하는데 이 부분은 pandas 라이브러리의 dropna를 통해 제거해주고 프로젝트를 진행했다. 주어진 가격을 그대로 사용하기보다는 가공하여 사용하였을 때 더 좋은 성능을 낼 수 있다는 선행연구에 기반하여^[1] 각 종목에 대한 price change, open change/price, high change/price, low change/price, volume change 를 생성해 이들만을 feature로 활용했으며, hyperparameter인 input_days로 windowing하여 사용했다. 최종 feature에 대한 Scaling 방법으로는 standard scaling을 사용하였다.

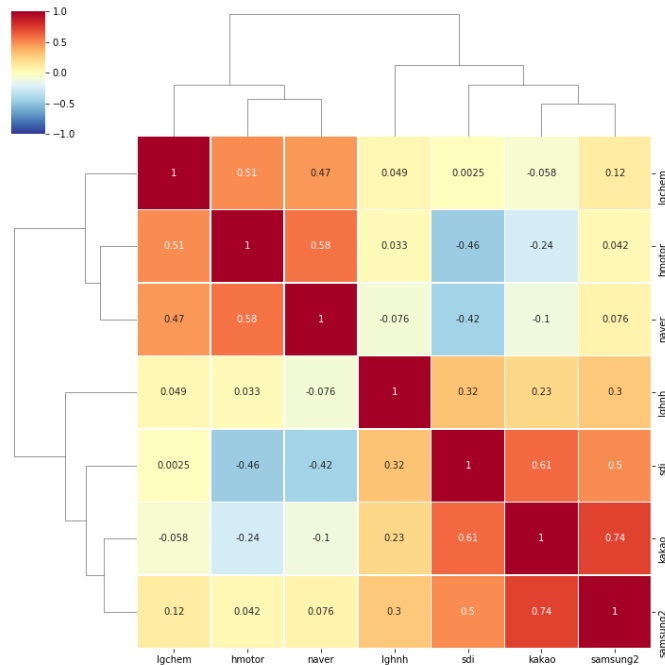
3. Trading strategy

본 프로젝트는 모델의 성능 뿐만 아니라 투자 전략 또한 수익률에 중요한 영향을 끼칠 것이라 판단이 되어 우리는 해리 마코위츠(Harry Markowitz)의 Modern portfolio theory 중 minimum variance portfolio 방식을 투자 전략으로 채택하였다. 이는 포트폴리오를 구성할 때, 상관도가 낮은 주식을 같이 매입하여 risk를 최소화하는 기법이다.



위 전략을 채택한 이유는 risk를 최소화하여 여유자금을 모두 투자하는 데에 대한 위험을 줄이기 위함이다. 상관도는 Pearson correlation 결과는 다음과 같다. (본 모델의 최종 training 기간 2018.01.01~2020.05.19). 상관도가 서로 낮은 ['hmotor', 'naver', 'lgchem', 'kakao', 'lgsh', 'samsung2', 'sdi'] 를 trading company로 채택하였다.

Action 은 "not_buying"와 각 주식을 선택할 수 있으며, Q-learning model 이 주식을 선택할 경우 총 자금의 50%를 해당 주식에, 이외의 종목에는 나머지 금액을 투자한다. "Not_buying" 일 경우 1000 만원을 골고루 투자한다.



4. RL components

4-1. Policy

Action 을 취할 때 exploration 의 비율을 결정하는 epsilon 을 simulation 이 진행될수록 감소하게끔 decay 를 적용하였다. Q matrix 를 학습하는 모델로는 Feedforward Neural Network 를 사용했으며 size 64 의 hidden layer 하나를 가지고 있는 3 layer 구조이다. 그리고 training 시에만 오버피팅을 방지하기 위해서 layer 간에 dropout 이 적용되도록 하였다. 또한 agent 의 기록을 랜덤하게 섞어주는 replay buffer 를 사용하여 학습이 더 빠르게 수렴하도록 유도하였으며, 하루 뒤에 무조건 갖고 있는 주식을 모두 팔아야 하는 우리 조의 전략에 맞게 gamma 값을 0 으로 설정하였다.

4-2. State

State는 기존 뼈대코드에 있던 budget과 num_stocks_list를 제외하고 input으로 들어가는 company list의 feature들로만 나타내었다.

4-3. Action

Action은 3.Trading Strategy에서 언급한대로 실행하였다.

4-4. Reward

다음날 포트폴리오와 현재 포트폴리오의 차이 즉, 하루 간 포트폴리오의 변화량으로 정의하였다.

5. Result

1000 번의 simulation (epoch)을 진행하며 valid portfolio 값을 기준으로 개선될 때마다 새로운 모델을 저장하였고 75epoch 에서 1 억 1951 만 9650 원이 최고 값으로 기록되었으며, 이 모델을 사용했을 때 Test portfolio 는 1 억 1063 만 8500 원이 나왔다. 따라서 이번 프로젝트에서 학습한 RL 모델로 10 일간 자본금 1 억을 가지고 매일 1000 만원 이상의 주식을 매입해야 한다는 조건을 지키면서 투자하였을 때 약 10.64%의 수익률을 기록하였다.

6. Reference

[1] Park, H., Sim, M. K., & Choi, D. G. (2020). An intelligent financial portfolio trading strategy using deep Q-learning. *Expert Systems with Applications*, 113573.