

# 최적의 주식 투자전략

이헌중

강남대학교 데이터사이언스 전공

E-mail: lhj990502@kangnam.ac.kr

## Introduction 서론

### 최적의 주식 투자전략

- 기술의 발전으로 인공지능(강화학습, 딥러닝)을 활용하여 실시간으로 대량의 데이터를 분석, 해석이 가능해짐. 기본적인 분석, 기술적 분석, 정서 분석과 함께 보다 정확하고 트렌드에 적합한 투자 결정이 가능

### 개인 투자자의 현실

- 개인 투자자는 기관 및 외국인 투자자들에 비해 수익률이 현저히 낮은 것이 현실



- 금융 정보에 소외된 저소득층, 초연령층도 쉽게 투자 포트폴리오를 구성할 수 있도록 종목 추천과 추천별 매도/매수/홀딩 타이밍 추천

## Method 방법론

### 데이터 수집 및 전처리

- 한국거래소, DART OPEN API, 증권사(크레온)API, FinanceDataReader에서 API, 크롤링을 사용하여 (종합주가(시가, 고가, 저가, 종가), 기업주가(시가, 고가, 저가, 종가), 상장법인 목록, 유동자산, 부채총계, 자본총계, 매출액, 영업이익, 당기순이익, 영업활동현금, 잉여현금흐름, 시가총액, 한국 국채 3년물, 투자자별 매매 실적) 데이터 수집

- 수집한 데이터를 사용하여 [등락률, 매출액 증가율, 영업이익 증가율, 부채비율, PER, PBR, ROE, PCR, PSR, POR, PFCR, NCAV/MK, 전일 종가대비 당일 시가비율, 당일 종가대비 당일 고가 비율, 당일 종가 대비 당일 저가 비율, 전일 거래량 대비 당일 거래량 비율, 이동평균(5,10,20,60,120)별 종가 대비 당일 종가, 거래량 대비 당일 거래량, 기관/외국인 투자자 순매수 거래량의 전일 대비 당일 비율] 파생변수 생성

- 2018년 ~ 2022년 학습 데이터 2023년 테스트 데이터로 설정

### EDA 탐색적 분석

- 변수별 상관관계확인, 투자자별 매매내역 분석, 업종별 시가총액 확인 등

### 종목 추천

- 기본적, 기술적 분석과 컨트투자를 통한 필터링으로 저평가된 종목 하나 선정 [004710]

- 로테츠 기반 필터링을 통해 앞서 선택한 종목과 유사한 주식 4개 추가 선정 [003670, 004770, 009140, 020150]

### 강화학습

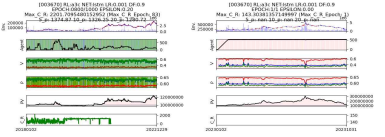
- 강화학습은 Markov Decision Process(MDP)의 개념을 기반으로하여 상태-행동 가치 함수를 신경망을 사용하여 모델링
- 강화학습 모델로는 DDQN, PG, A3C, 신경망 모델로는 DNN, CNN, LSTM을 사용하여 가장 좋은 선정
- 이용과 탐험을 조절하여 투자를 진행하며 이를 통해 경험을 쌓아 잘못된 행동에는 긍정적으로 보상하고, 잘못된 경우에는 부정적으로 보상
- 포트폴리오 가치(PV) = 주식 잔고 X 현재 주가 + 현금 잔고
- 매도, 매수, 홀딩에 따라 PV가 변하며, PV를 높이는 것이 목표



## Result 결과

### 강화학습 과정

- 누적 보상이 높은 모델이 좋은 모델이라는 것을 기준으로 매개변수 선정



- 강화학습 모델: A3C, 신경망: LSTM, 최적화 함수: RMSprop, 학습률: 0.001, 에포크수: 1000, 탐험률: 100%에서 10%씩 감소, 할인 비율: 0.9
- 과적합을 방지하고자 누적 보상 비율이 상위 5% 평균보다 5번 이상 높을 시 조기 종료

### 매수 매도 홀딩 타이밍 추천

- 23년 11월 예측 진행 결과: 매수, 매도, 홀딩에 대한 확률 확인

```
[{"20231107", [0.587322947170667, 0.45868414640426636, 0.427944302588999],
{"20231108", [0.5909274220466614, 0.398074209609094, 0.3460130989551544],
{"20231109", [0.592709481716156, 0.39848726654052734, 0.3457196354866028],
{"20231110", [0.593349814414978, 0.38397786021232605, 0.3318164213180542],
{"20231113", [0.5918446643829346, 0.38031429852352905, 0.33287304639816284],
{"20231114", [0.5884157419284712, 0.387952059173584, 0.345743864774704],
```

## Discussion 토의

### 성과 및 한계

- 2018년부터 2023년까지의 데이터를 활용하여 강화학습을 진행하였고, 그 결과 자트 데이터에 따른 주식의 매도, 매수, 홀딩 전략을 추천받을 수 있음
- A3C 모델은 특정 종목에서 양호한 성과를 보이지만, 다른 종목에서는 과적합 현상이 발생한 케이스도 있었으며, 이를 극복하기 위해 하이퍼파라미터 조정과 모델의 일반화 능력을 향상시키는 노력을 하였으며 추가적인 테스트가 필요함
- 주식 시장은 이슈에 민감하게 반응하는 특성이 있어, 이를 고려하기 위해 기사를 수집하여 감정분석을 진행함, 이 과정에서 아직 모델 학습에 녹여내지 못한 부분이 아쉬움

### 향후 연구 방향

- 거래 수수료와 거래세를 고려하여 모델을 학습
- 데이터의 다양성을 확보하기 위해 기사 데이터와 더불어 다양한 변수들을 추가할 계획
- 향후 최적의 주식 투자전략을 위해 만들어진 이 모델을 활용하여 자동 주식 매매 시스템을 구축 예정

## References 참고

- 강화학습을 이용한 주가 예측 Stock price prediction using reinforcement learning
- 딥러닝 기반 강화학습 모델 성능 비교 - 국내 주식시장 사례연구
- 딥러닝을 이용한 주가 예측 모델 = Stock price prediction model using deep learning
- 비대칭성을 활용한 인공지능 주가 예측 분석
- 파이썬을 이용한 딥러닝/강화학습 주식투자
- 인공지능 모델을 활용한 주가 예측성과 분석: 국제시장 비교 = Prediction Performance of Stock Price using Artificial Intelligence Models: Global Market Approach