

# 재무 정보를 바탕으로 선별한 종목의 주가 예측

장유진<sup>o</sup> 전석희

경희대학교 컴퓨터공학과

2015104214@khu.ac.kr jeon@khu.ac.kr

## Stock price prediction

### selected based on financial statement

Yujin Jang<sup>o</sup>, Seokhee Jeon

Department of Computer Science and Engineering, Kyung Hee University

#### 요약

전문 주식 투자자들은 투자할 종목을 선택할 때 재무 정보를 참고한다. 하지만 재무 정보를 바탕으로 선별한 종목이 항상 수익을 내지는 못한다. 현재 머신러닝은 각종 분석을 위해 사용되고, 이 분석한 정보들을 바탕으로 여러 서비스를 제공한다. 머신러닝을 적용하면 좋은 상황 중 하나가 빅데이터인데 사람이 미처 파악하지 못할 데이터간의 관계를 파악할 수 있기 때문이다. 특히나 주식은 수십 년의 주가 지수와 여러 종목별 주가가 있는 빅데이터다. 이 데이터들을 이용하여 주가를 예측할 수 있다면 투자자가 더 안정적이고 효율적인 투자를 할 수 있게 될 것이다. 원하는 재무 정보를 사용하여 종목을 선별한다. 선별한 종목들을 인공지능을 통해 주가를 예측한 후, 예측 결과가 주가 상승인 종목을 2차적으로 선별하여 수익률을 높인다.

#### 1. 서론

돈을 많이 벌고자 하는 욕망은 대부분의 인간이 가지고 있는 욕구이다. 하지만 노동을 통해 얻을 수 있는 소득은 정해져 있다. 따라서 사람들은 투자를 하기 시작했고, 주식은 현재 가장 많은 사람들이 이용하는 투자 수단이 되었다. 하지만 단순히 기업 이미지, 기업 대표, 인기 등을 보고 주식을 구매하는 것은 투자가 아니라 투기에 가깝다. 그래서 합리적이고 효율적인 투자를 위한 여러 분석 기법들이 탄생했다. 실제로 이러한 분석 기법을 이용한 투자는 수익률을 봤을 때 통계적으로 유의미한 결과를 보여줬다.

최근 머신러닝 분야가 빠르게 성장함에 따라 여러 분야에 머신러닝들이 적용되고 있다. 사람이 해결하기 어려웠던 문제나 정확도가 떨어졌던 문제들이 머신러닝들로 해결되고, 분석하기 어려운 데이터들에 대한 분석도 머신러닝으로 이루어진다. 나아가 머신러닝은 단순히 문제 해결에만 사용되는 것에 그치지 않고 이를 적용한 서비스를 제공하거나, 최근에는 사업 아이템 자체가 머신러닝이 메인인 스타트업도 생겨나고 있다. 특히 머신러닝이 사람보다 더 좋은 성능을 발휘하는 분야 중 하나가 빅데이터다.

주식시장은 공식적인 데이터들이 아주 많이 쌓여있는 빅데이터라고 볼 수 있다. 위에서 언급한 분석 기법들이

아직도 유용하게 사용되고 있지만, 주식시장이 빅데이터라면 머신러닝을 적용해보지 않을 이유가 없다. 합리적인 분석 기법 중 하나인 재무 정보를 통한 종목 분석을 하여 투자할 종목을 선별한다. 그 후 이 종목들이 실제로 수익을 낼 지 예측하는 딥러닝 모델을 만들어 투자할 종목을 2차적으로 선별하면 수익률을 더욱 높일 수 있을 것이다.

## 2. 관련연구

### 2.1. PER(주가수익률)과 주가의 관계

PER이란, 기업의 현재 주가를 EPS(주당순이익)으로 나눈 값을 나타낸다. EPS란, 기업의 당기 순이익을 총 발행 주식수로 나눈 것으로 기업이 1년 동안 1주당 얼마의 순이익을 벌었는지 계산한 것이다. 즉, PER의 의미는 기업의 현재 주가가 순이익에 비해 어떠한지 나타내는 것이다. PER이 높다는 것은 1주당 순이익에 비해 1주의 가격이 높게 형성되어 주가가 고평가되었다는 것이고, PER이 낮다는 것은 1주당 순이익에 비해 1주의 가격이 낮게 형성되어 주가가 저평가되었다는 것이다. PER이 낮은 기업은 대체로 저평가되어있기 때문에 주식을 매수하기 좋은 기업으로 평가할 수 있다. 다만 당기 순이익이 일회성 수익일 가능성이 있으므로 주의해야 한다.

### 2.2 ROE(자기자본이익률)과 주가의 관계

ROE란 당기 순이익을 자본총액으로 나눈 값을 나타낸다. 즉, 기업이 자본을 이용하여 얼마만큼의 이익을 냈는지 나타내는 지표이다. 예를 들어, 10억의 자본으로 1억의 순이익을 냈다면 ROE는 10%가 된다. ROE가 높다는 것은 투자금 대비 더 큰 수익을 낸다는 것이므로 ROE가 높은 기업을 선택하는 것이 좋다. ROE도 PER과 마찬가지로 일회성 수익일 가능성이 있으므로 보통 3년 평균 ROE 값을 보는 것이 좋고, 최소한 은행 이자보다 높은 ROE를 기록하는 기업을 선택하는 것이 좋다.

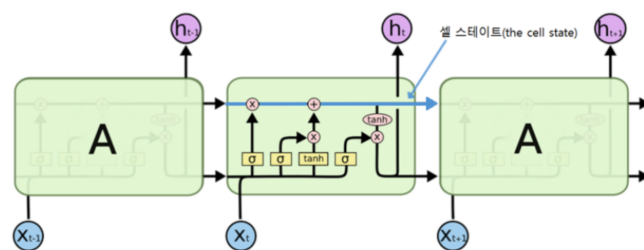
### 2.3 PBR(주가순자산비율)과 주가의 관계

PBR이란, 현재 주가를 현재 BPS로 나눈 값을 나타낸다. BPS란, 기업의 현재 순자산을 총 발행 주식수로 나눈

것으로 만약 기업이 현재 청산한다면 1주당 얼마씩 돌아가는지 나타내는 지표다. 일반적으로 PBR이 낮은 기업이 높은 주가수익률을 기록한다. 하지만, PBR은 부채를 반영하지 않으므로 주의해야 한다. 또한, 동종 업계 현황이 안좋거나, 회사 내부에 문제가 있어도 PBR은 낮아지므로 주의해야 한다.

## 2.4 LSTM

RNN의 경우, 데이터간의 거리가 멀어질수록 데이터의 상관관계가 불명확해지는데 이를 개선한 모델이 바로 LSTM(Long Short Term Memory Network)이다. 가장 기본적인 RNN 모델의 경우, 학습이 거듭될수록 초기의 weight값이 유지되지 않아 학습률이 저하되는 Long-Term Dependency 현상이 생긴다.



[그림 1] LSTM 구조와 셀 스테이트

[그림 1]과 같이 LSTM에서는 Cell State라는 layer를 추가해서 weight를 계속 기억할 것인지 결정하여 이 문제를 개선했다.[2] 이를 통해 직전의 데이터만 참고하는 Vanilla RNN과 달리 LSTM은 cell state에서 다른 weight를 참고할 수 있게 되었다.

## 3. 프로젝트 내용

### 3.1. 재무 정보를 바탕으로 설정한 조건에 맞는 종목을 선별해주는 프로그램

여러 기업의 재무 정보를 수집하여 데이터베이스에 저장한다. PER, ROE, PBR 조건을 설정한다. 데이터베이스에 접근하여 저장된 재무 정보들을 계산하여 설정한 조건에 맞는 기업을 선별한다.

선별한 기업의 매수 날짜, 매도 날짜를 설정한다. 설정한 매수 날짜에 일정 금액만큼 주식을 매수한 후,

매도 날짜에 전부 매도할 때 얼마큼의 손익이 발생했을지 계산해주는 프로그램을 개발한다.

### 3.2. 재무 정보를 바탕으로 선별한 종목들의 주가를 예측하는 모델

위의 프로그램에서 선별한 종목들을 전부 매수, 매도할 경우 주가가 하락하는 종목에 대한 만큼 손실이 발생한다. 따라서 선별한 종목의 주가를 딥러닝 모델을 통해 예측한다. 그 후, 주가가 상승할 것이라고 예측한 종목에 대해서만 매수, 매도할 경우 전체 종목을 매수, 매도한 것에 비해 더 좋은 수익을 기록할 것이다.

## 4. 결론 및 향후연구

본 논문에서는 재무 정보를 바탕으로 설정한 지표에 해당하는 종목을 투자했을 경우 얻었을 손익과 딥러닝을 사용하여 주가를 예측하는 모델에 대해 연구하였다.

수백개에 해당하는 기업들 중 투자할 종목을 선택하기 쉽지 않다. 게다가 전문가도 한 기업의 재무 정보를

분석하는 데에 최소 30분은 필요하다고 하는데, 비전문가는 시간이 더 오래걸릴 것이다. 따라서 종목을 선택할 때 사용하는 지표의 조건을 설정하여 이에 해당하는 종목들을 알려주는 프로그램은 투자 종목 선택시 편리함을 제공할 것이다. 또한, 선택한 종목들 중 좋은 수익을 기록할 것 같은 종목을 예측해주는 딥러닝 모델은 수익률을 더 높일 수 있을 것이다.

다만 재무 정보에서 설정하는 조건은 종목 선택에 완벽하지 않다. 당기 재무 정보엔 좋은 지표가 사실은 일회성 지표고, 몇년간의 재무 정보를 봐야 알 수 있는 내용도 있기 때문이다. 이러한 몇년간의 재무 정보도 같이 계산하여 사용하면 수익률을 더욱 높일 수 있을 것이다.

### 참고문헌

- [1] 이지훈, 딥러닝을 이용한 주가 예측 모델, 국내석사학위논문 숭실대학교 대학원, 2017, 서울
- [2] Christopher Olah, Understanding LSTM Networks, 2015.  
<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- [3] Andreas C.Muller and Sarah Guido, Introduction to machine learning with python, 2017.