

기술적 분석지표와 딥러닝을 활용한 주가 예측 기법

원종민(*), 황현수(**), 정유현(***), 박희동(****)

(*)-(****) 나사렛대학교 IT융합학부, (*)-(****){kceeck, bbss67117, braker051, knudolby}@gmail.com

Stock Price Prediction Technique Using Technical Analysis Index and Deep Running

Jong-Min Won(*), Hyun Su Hwang (**), You-Hyun Jung(***), Hee-Dong Park(****)

(*)-(****)Korea Nazarene University, School of IT Convergence,

(*)-(****) { kceeck, bbss67117, braker051, knudolby}@gmail.com

요 약

주가 예측에 사용되는 분석방법으로 기본적 분석과 기술적 분석의 문제점이 제기되었다. 바둑 인공지능 알파고가 많은 관심을 받은 이후 금융업계에서는 인공지능을 이용한 주가 예측에 대한 기술 개발에 많은 관심을 가진다. 장점으로는 데이터 기반의 객관적인 추천, 낮은 최소투자금액과 저렴한 비용이 있다. 본 프로젝트에서는 기술적 분석 지표와 딥러닝의 LSTM모형을 결합하여 대회 측에서 제공한 자료를 이용해 예측 데이터를 생성한다. 지도 학습 방법인 분류분석을 이용해 그래프를 생성하는 모델을 만들어 금융시장의 미래에 대한 상황을 예측할 수 있는 방안들을 제안하였다.

1. 서론

주가 예측이란 미래의 주식시장 가격을 예측하기 위한 기술로, 2016년 3월 구글 (Google)의 인공지능 시스템인 알파고 (AlphaGo)가 이세돌 9단과의 바둑대결에서 승리한 이후 인공지능 기술을 활용한 주가 예측 또한 많은 사람들의 관심을 받고 기술 개발이 이루어지고 있다 [1, 2].

주가 예측에 사용되는 분석방법으로는 크게 기본적 분석(fundamental analysis)과 기술적 분석(technical analysis)이 있다. 기본적 분석은 기업의 내재가치 분석을 통해 종목을 선정하는 방법으로 경제흐름의 대세를 판단하지만 내재가치로 수렴하는데 어느 정도의 시간이 걸릴지 알 수 없다. 기술적 분석이란 전통적인 증권분석 방법으로 과거 주식의 가격이나 거래량 같은 자료를 이용하여 주가 변화의 추세를 발견해내어 미래의 주가를 예측하는 방법이다 [3].

본 논문에서는 기술적 분석 지표와 딥러닝의 LSTM

(Long-Short Term Memory) 모델 [4]을 결합하여 일부 종목 주식자료에 대한 예측하는 모델을 개발하고자 두 가지 연구를 진행한다. 첫 번째로, 딥러닝에서 RNN의 변형인 LSTM 모델을 이용해 예측 데이터를 생성한다. 두 번째로, 딥러닝에서 LSTM 모델을 이용해 예측한 데이터를 대상으로 지도 학습 방법인 분류(Classification) 분석을 이용해 그래프를 생성하여 해당 종목의 다음 한달 주식 상황이 상승 또는 하강하는지 실제 값과 비교해 모델의 효율성을 확인하고자 한다.

2. 제안하는 모델

2.1 Deep Learning(LSTM)

실제 일 년 주가 데이터를 불러와 각각의 가격 데이터를 정규화 처리하여 학습과 테스트 데이터로 나누어 기존 하이퍼파라미터에 지정한 값을 이용해 모델을 생성한다. 모델 생성 과정 중 손실함수와 최적화함수를 이용하여 입

력 값에 넣어주고 Cell들의 연결과 마지막 hidden값을 출력하는 학습을 진행하여 many-to-one 형태로 시가, 고가, 저가, 종가, 거래량 데이터를 각각 반복하여 다음 날의 주가 데이터를 예측하고 이를 한 달 반복하여 예측 데이터를 생성한다.

2.2 SVM(Support Vector Machine)

딥러닝에서 LSTM 모델을 이용하여 예측한 한 달 데이터들 중 주가에 영향을 줄 수 있는 종가를 토대로 초평면에 그리게 되면 어떤 직선을 그어도 두 범주를 완벽하게 분류하기 어려운 경우인 비선형 분리 면을 갖게 된다. 이때 커널함수를 도입하여 입력벡터, 즉 입력 변수를 더 높은 차원의 특징 공간으로 사상하여 그 공간에서 각각의 데이터 간 거리를 측정하여 여러 개의 데이터 사이의 중심을 구한 후 그 가운데에서 최적의 초평면을 얻음으로써 비선형으로 분포되어 있는 패턴도 분리하여 마진이 가장 큰 결정 경계를 찾아내어 예측한 한 달 종가에 대한 상승과 하락을 구분한다.

3. 실험교육 및 평가

(그림 1)은 실험에 사용된 미래에셋대우의 원본 및 예측 주가 데이터를 나타낸다.

(그림 1) 미래에셋대우의 원본 및 예측 주가 데이터

미래에셋대우										
Date	Open	High	Low	Close	Volume	Open	High	Low	Close	Volume
2011-01-03	23435.52	23435.52	23255.44	23255.75	1916799	23287.4	23473.34	22801.84	23128.77	17604.68
2011-01-04	23479.84	23524.06	23037.66	23214.53	1634302	22956.95	23274.96	22821.94	22690.22	17426.86
2011-01-05	23126.09	23170.31	22462.82	22551.26	3039914	22965.19	23297.27	22170.9	22648.21	17650.85
2011-01-06	22728.13	22905	22241.73	22418.6	2461253	22875.04	23340.47	22799.23	22783.34	17645.95
2011-01-07	22418.6	22639.69	22197.51	22418.6	2194664	22734.4	23222.91	23078.58	23030.67	17380.12
2011-01-10	22330.17	22860.78	22285.95	22551.26	1858189	22952.76	23269.75	22377.85	22606.58	17385.37
2011-01-11	22639.69	23214.53	22462.82	23214.53	2721510	22559.76	22381.97	21709.65	22055.88	17345.33
2011-01-12	23347.15	23615.49	23136.09	23347.15	2067243	20871.39	20646.97	21114.51	21063.49	17009.94
2011-01-13	23568				176	22121				57
2011-01-14	23833				906	19702				41
2011-01-17	24187				341	20281				5
2011-01-18	24451				523	22281				78
2011-01-19	24562				894	2114				29
2011-01-20	25077.59	25205.95	25061.47	24955.21	3297971	21845.95	23483.77	22674.54	23107.89	17225.77
2011-01-21	25558.09	25602.31	24275.76	24452.64	3190504	24312.66	23957.74	22688.77	24983.41	17892.93
2011-01-24	24452.64	24939.04	24275.76	24762.16	1703403	25978.19	25213.27	25830.14	26661.26	16793.82
2011-01-25	25027.47	25248.56	23745.15	24054.67	2625903	26114.02	30203.05	23541.96	27293.63	17007.09
2011-01-26	24231.55	24717.95	24010.45	24585.29	1521856	27476.74	30643.32	28829.14	28697.74	20815.93
2011-01-27	24585.29	24873.73	24010.45	24319.98	1524166	24559.41	24523.36	24520.2	26433.45	20684.89
2011-01-28	24187.33	24595.29	24010.45	24452.64	1708737	23576.59	25583.74	21789.95	20602.98	20785.69
2011-01-31	23877.8	23966.24	23126.09	23170.31	4036267	22075.92	18969.17	20511.12	18929.72	19369.74

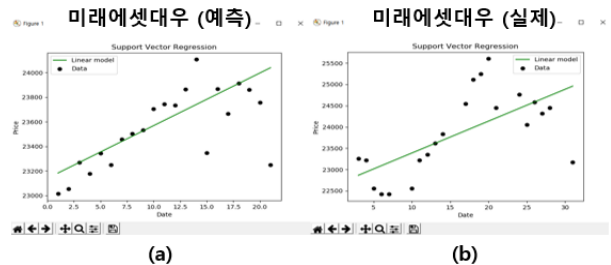
(Figure 1) Source and prediction stock price data of Mirae Asset Daewoo

(그림 2) 그래프들의 점들은 각각의 날짜에 대한 종가 데이터를 나타내며, 초록색 그래프는 최적의 선형 결정 경계를 찾는 선형 SVM(Linear SVM)이다. (a) 그래프는 딥러닝의 LSTM 모델을 이용하여 2010년 데이터들을 학습하여 다음 2011년 1월 한 달 종가에 대한 등락을 예측하는 그래프이고, (b) 그래프는 2011년 1월 실제 종가에 대

한 등락을 보여준다.

(그림 2) 2011년 1월의 미래에셋대우 주가 움직임.

(a) 예측 그래프 (b) 실제 그래프



(Figure 2) Mirae Asset Daewoo's stock price movement during Jan. 2011. (a) Forecast graph. (b) Actual graph.

4. 결론

본 논문에서는 많은 금융시장의 변수와 불규칙적으로 변하는 상황 때문에 기술적 분석과 기계학습을 이용하더라도 정확한 데이터를 예측하는 것은 한계가 있어 금융시장의 상황을 예측하고자 종가에 대한 미래의 등락을 예측하는 모델을 만들어 금융시장의 미래에 대한 상황을 예측할 수 있는 방안들을 제안하였다.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2017R1D1A3B06035024).

참고 문헌

- [1] L. Arnat, "Stock Price Prediction by Deep Learning," Dec. 2016.
- [2] 송유정, 이재원, 이종우, "텐서플로우를 이용한 주가 예측에서 가격-기반 입력 피쳐의 예측 성능 평가," 정보과학회 컴퓨팅의 실제 논문지, Vol. 23-11, pp. 625-631, 2017년 11월.
- [3] 기본적분석과 기술적분석 비교 정리 : <http://md2biz.tistory.com/362>
- [4] LSTM, [Online]. Available: https://en.wikipedia.org/wiki/Long_short-term_memory