

빅데이터를 활용한 인공지능을 통한 주식 예측 분석 사례

최민기, 조광익, 전민기, 최훈*

부산가톨릭대학교

Cases of Stock Analysis through Artificial Intelligence Using Big Data

Min-gi Choi · Kwang-ik Jo · Min-gi Jeon, hun Choi*

Catholic University of Pusan

E-mail : alsrl9606@gmail.com / dlr0359@naver.com / alsrlrkd72@naver.com / chlgn@cup.ac.kr

요 약

21세기는 4차 산업혁명을 맞아 빅데이터를 활용한 여러 분야의 연구가 진행되며 세상에 혁신적이고 유용한 기술이 끊임없이 나오고 있다. 빅데이터 시대에 최근 여러 기술 중 인공지능의 여러 알고리즘을 활용한 다양한 분야 중 금융 분야에서 빛을 발하여 핀테크, 금융사기 탐지 및 위험 요인 관리 등에 쓰이며, 최근 붐이 일어나고 있는 주식시장에도 인공지능 알고리즘 모델을 활용한 투자 예측 및 투자 요인 분석 등에 활용되고 있다. 본 논문에서는 빅데이터를 활용한 인공지능을 통해 주식시장에서는 어떻게 쓰이고 있는지 다양한 연구 사례를 살펴봄에 동향을 살펴볼 예정이다.

ABSTRACT

In the 21st century, as we enter the Fourth Industrial Revolution, research in various fields utilizing big data is being conducted, and innovative and useful technologies are constantly emerging in the world. Among several technologies recently in the big data era, among various fields utilizing some algorithms of artificial intelligence, it shines in the field of finance and is used for pin tech, financial fraud detection and risk management, etc., and recently Even in the booming stock market, it is used for investment prediction and investment factor analysis using artificial intelligence algorithm models. In this paper, we plan to investigate various research cases and investigate trends in how they are used in the stock market through artificial intelligence that utilizes big data.

키워드

Big Data, AI, stock, finance

1. 서 론

최근 저금리 기조가 지속되면서 일반인들의 투자처가 일반 시중 은행의 예금에서 부동산과 주식 시장으로 변화하는 추세다. 전민종(외 1명, 2020)은 2020년 1월, 국내 주식거래계좌의 개수가 약 3천 만개 이상 돌파했다고 말하며, 국내 증시가 호황과 관계없이 해외 증시 호황이 증시 투자 심리를 만들었다고 했다[1]. 또한 핀테크의 등장으로 비대면 계좌 개설이 개인 투자자에게 유용해지면서 계좌 개설이 급증했다. 하지만 일반 투자자들이 기관과 외국인을 상대로 주식 투자를 하기에는 정보 비대칭 및 시기적절한 시장 대응이 실시간으로

대응하기 힘들어 주식 투자 성공을 하기 매우 힘들다. 전민종(외 1명, 2020)은 투자자들의 목적은 이익을 내기 위해서 주식에 투자한다고 말하며, 이러한 투자자들의 요구를 충족시키기 위해 주식시장을 분석하고 예측하는 연구가 이루어지고 있다고 주장했다. 그래서 투자자들이 다양한 기업 정보들을 활용하여 주식 투자에 활용하고 있으나 성공적인 주식 투자를 하기에는 한계가 있다. 따라서 최근 들어 다양한 주식 예측 프로그램들을 활용하여 주식 투자에 활용하고 있다.

이에 따라 본 연구에서는 다양한 주식 예측 프로그램 중에서도 최근 가장 널리 사용하고 있는 기법인 인공지능을 활용한 주식 예측 프로그램에 대해 살펴보고자 한다. 이를 위해, 본 연구에서는 기존에 나와 있는 주식 예측 프로그램에 대해 살

* corresponding author

해보고 각각에 대한 장점 및 문제점을 비교 분석하고자 한다. 이를 통해 인공지능을 활용한 주식 예측 프로그램이 얼마나 활용도가 높은지와 기대를 해도 되는지에 대해 알아보하고자 한다.

II. 선행 연구 사례

요즘 붐을 일으키는 주식 예측은 굉장히 높은 관심 주제이며, 현재 21세기는 전부터 주식 분석과 예측에 관한 자료가 쏟아져 나오고 있고, 정확한 예측을 위해 많은 연구가 진행 중이다. 천세학(2019)은 주식시장 예측에 있어 학문적으로 많은 연구가 있다고 말했다. 하지만 그 연구 결과를 사용해 실제 투자에 이용한 사례는 그리 많이 알려지지 않았고, 실제로 주식시장은 예측하기 상당히 어렵고, 학술 논문이 투자에 이용할 수 있을 정도의 많은 요인을 반영하기 어렵다고 말했다[2]. 안현철(외 1명, 2009)은 주식시장(stock market)을 일컫는 용어로 랜덤 워크(random walk)가 존재한다고 말했다. 이 말은 주식시장의 방향을 예측하는 것이 어렵다는 것을 얘기한다고 주장했다. 주식시장은 고려해야 하는 많은 양의 정보와 데이터가 있고, 불확실성이 꽤 높게 존재하며, 예상치 못한 영향 요인이 다양하게 존재한다고 했다. 정치적인 요소와 개인적인 요소뿐만 아니라, 일반적인 기업, 경제 상황, 기후, 환경 등의 여러 가지 변수들이 복잡하게 주식시장에 영향을 미치고 있다고 말했다[3]. 김규영(외 1명, 2004)은 1984년 이후 2003년까지의 월별 자료에 대하여 OLS Lewellen(2004)의 방법을 적용해, 한국의 주식시장에서 재무적 변수들이 주식수익률의 예측 가능성이 있는지를 검증하였다. 이들 연구 결과에 따르면, 전체 표본 기간에서 배당수익률이 주식수익률에 관한 예측력을 가지고 있지 못한 것으로 나타났다[4]. 이처럼 기존의 문서자료로는 주식 예측에 대한 불확실성이 높게 드러나는 것으로 나타났다.

III. AI를 활용한 주식 예측 사례

정욱(2018)은 주식시세를 예측하는 딥 러닝 최적화 방법을 연구를 LSTM(Long Shot Term Memory)를 활용해 진행했다. 하이퍼 파라미터의 최적화를 통한 Model 튜닝과 관련 거래시간 feature 제공을 추가해 주가 예측치를 높은 예측 정확도를 제공하여 하이퍼 파라미터의 튜닝 여부에 따라서 학습 정확도가 9.68%, 예측 정확도가 6.43% 차이를 보이는 것을 확인했다. 이러한 LSTM의 장점으로는 기존의 주식 예측 시스템들과는 높은 차이를 보이고 LSTM에 추가하는 하이퍼 파라미터의 종류나 다양한 데이터를 기반으로 한다면, 정확도가 90%까지도 무궁무진하게 발전할

수 있다. 다만, 이러한 결과를 내기 위해서는 종목별로 어떤 하이퍼 파라미터의 튜닝을 해야 하는지가 전부 다르므로, 이러한 하이퍼 파라미터의 종류를 잘 선택해야 예측률을 높일 수 있으며, 실제 자동화를 실행해도 많은 시간이 소요되기 때문에 최적화된 소프트웨어와 하드웨어가 뒷받침되어야 한다. 또한, 매번 예측할 때마다 학습과 하이퍼 파라미터 탐색을 진행하는데 시간과 컴퓨팅 리소스 비용이 들기 때문에 한 번 생성된 모델에 대해 품질 관리와 재학습을 실행하는 로직이 필요하다[5].

IV. 결 론

주식시장에 예측률 및 정확률에 대한 변수는 무궁무진하게 많으면서 불확실성이 매우 크다. 이러한 주식 예측 시스템에 관한 개발과 연구는 꽤 오래전부터 끊임없이 이루어지고 있으며, 4차 산업혁명을 통한 빅데이터 시대의 인공지능을 통한 주식 예측 프로그램은 기존의 거시경제지표와 같이 문서자료에 의한 예측보다는 현저히 높은 예측률을 나타내고 있다. 이러한 주식 예측 프로그램에 더 연구에 정진한다면, 인공지능을 통한 매수, 매도를 결정하는 날이 올 것이다. 하지만 여전히 기계에 의한 예측이므로 많은 문제를 일삼고 있으며, 어떤 변수가 나타날지가 문제이다. 이러한 문제를 잡고 비용에 관한 문제에 해결 방법을 제시한다면 좋은 주식 예측 프로그램이 될 것으로 기대된다.

Acknowledgement

이 보고서(과제, 결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 사회 맞춤형 산학협력 선도대학(LINC+) 육성사업의 연구 결과입니다.

References

- [1] 전민종, 이욱(2020), NARX 신경망 최적화를 통한 주가 예측 및 영향 요인에 관한 연구
- [2] 천세학(2019), 데이터 크기에 따른 k-NN의 예측력 연구: 삼성전자 주가를 사례로
- [3] 안현철, 이형용(2009), 투자 의사결정 지원을 위한 유전자 알고리즘 기반의 다중 인공지능 기법 결합 모형
- [4] 김규영·김영빈 (2004), “한국 주식시장에서 수익률의 예측에 관한 연구:재무 변수를 이용한 예측을 중심으로,”산업경제 연구, 제17권 제 4호, 1255-1271. 한국증권거래소, 증권통계 연보, 1983 - 2004 각년도.
- [5] 정 욱 (2018). 주식시세 예측을 위한 딥러닝 최적화 방법 연구