



텐서플로우를 이용한 주가 변동 예측 딥러닝 모델 설계 및 개발

A Design and Implementation of Deep Learning Model for Stock Prediction using TensorFlow

저자 (Authors)	송유정, 이종우 Yoojeong Song, Jongwoo Lee
출처 (Source)	한국정보과학회 학술발표논문집 , 2017.06, 799-801 (3 pages)
발행처 (Publisher)	한국정보과학회 KOREA INFORMATION SCIENCE SOCIETY
URL	http://www.dbpia.co.kr/Article/NODE07207386
APA Style	송유정, 이종우 (2017). 텐서플로우를 이용한 주가 변동 예측 딥러닝 모델 설계 및 개발. 한국정보과학회 학술발표논문집, 799-801.
이용정보 (Accessed)	한밭대학교 203.230.104.*** 2018/03/21 16:00 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독 계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

텐서플로우를 이용한 주가 변동 예측

딥러닝 모델 설계 및 개발

송유정[○] 이종우

숙명여자대학교 IT공학과

{yjsong, bigrain}@sookmyung.ac.kr

A Design and Implementation of Deep Learning Model for Stock Prediction using TensorFlow

Yoojeong Song[○] Jongwoo Lee

Dept. of IT Engineering, Sookmyung Women's University

요 약

최근 딥러닝에 대한 연구가 활발히 진행되고 있으며, 국내외 할 것 없이 다양한 기술 개발 및 서비스를 제공하고 있다. 딥러닝은 다양한 분야에서 이용될 수 있는데, 그 중 과거부터 현재까지 풀리지 않는 난제로 남아있는 것이 바로 주식 시장 예측이다. 다양한 시도들이 있어왔지만 아직까지 정확한 미래 주가를 예측하는 것은 불가능했다. 하지만, 현재 많은 발전을 거듭한 인공지능 기술인 딥러닝을 이용한다면 주식 데이터를 통해 특정 패턴을 찾아낼 수 있고, 이를 통해 미래의 주식 시장 변동을 예측하는 것이 가능할 것이다. 본 논문에서는 주식데이터를 이용해 주가 변동을 예측하기 위해 실제 주식 데이터의 변동 패턴을 학습하는 딥러닝 모델을 설계하고 구현한다.

1. 서 론

최근 구글에서 개발한 인공지능 로봇인 알파고(AlphaGo)의 바둑 대결 승리로 인해 인공지능이 매우 주목을 받고 있으며, 이에 따라 ‘딥러닝’에 대한 연구가 활발히 진행되고 있다. 구글, 바이두, 페이스북 등 해외에서도 매우 활발한 연구개발이 이루어지고 있으며 국내 또한 대기업/중소기업, 스타트업 할 것 없이 딥러닝을 이용한 연구를 진행하고 있다.

딥러닝은 다양한 분야에서 이용될 수 있는데, 그 중 과거부터 현재까지 풀리지 않는 난제로 남아있는 것이 바로 주식 시장 예측이다. 주가를 과학적으로 예측하기 위한 다양한 시도들이 있어 왔지만, 아직까지 정확한 미래 주가를 예측하는 것은 불가능했다. 하지만 주가에 대한 예측은 경제학, 수학, 물리학, 그리고 전산학 등 여러 관련 분야에서 오랜 관심의 대상이 되어 왔으며[1], 과거 미국 주식시장에 대한 실증분석 결과를 살펴보면 다양한 데이터 분석을 통한 주식수익률 예측이 통계적으로 유의미한 것으로 나타났고, 이는 주가예측이 가능하다는 증거로 제시되어 왔다. 딥러닝을 이용해 높은 확률로 주가 변동을 예측할 수 있다면, 예측 결과를 기반으로 투자를 진행하여 최종적으로 수익을 얻을 수 있을 것이다. 따라서 본 논문에서는 오픈소스 딥러닝 프레임워크인 텐서플로우(TensorFlow)[2]를 이용하여 주가 변동을 예측을 위한 학습모델을 설계 및 개발하고자 한다.

2. 관련연구

본 장에서는 주식 예측과 관련된 기계학습 및 딥러닝 연구에 대해 설명한다. 국내에는 대표적으로 주식 데이터를 다루는 기업인 코스콤이 SNS 데이터인 소셜 데이터를 분석하여 주식 시장을 예측하는 시스템을 개발 하였다. 코스콤의 주가예측 시스템에서는 데이터 분석 운영 환경을 구축하기 위해 오픈소스 분산처리 기술 하둡과 분석용 프로그래밍 언어 R을 적용하였으며 70% 내외의 적중률을 보이는데 성공하였다[3]. 벤처회사인 티그레이프(Tgrape)사는 핑거스톡(FingerStock) 서비스를 제공하는데 이는 그림 1과 같다. 기계학습 알고리즘인 SVM, NN 등을 이용해 주가 예측 결과를 모바일로 제공하도록 설계되어있다[4]. 또한 그림 2는 스마트포어캐스트(Smartforecast)사의 어플리케이션인 빅터(Bigta)의 화면이다. 빅터는 SNS 데이터 등의 빅데이터 분석을 통해 주가 예측, 매수/매도 타이밍을 알려준다[5]. 스타트업인 딥넘버스(DeepNumbers)는 주가 변동 상황 예측을 위해 딥러닝을 연구하고 있으며 알고리즘 기반 자문서비스를 연구하고 있고 아직 서비스는 제공되고 있지 않다[6]. 딥넘버스를 제외한 코스콤, 티그레이프, 스마트포어캐스트의 서비스들은 모두 딥러닝이 아닌 기계학습 알고리즘인 SVN, NN 등을 이용해 개발되었다. 마지막으로 금융시장을 컴퓨터 과학으로 풀기 위한 서울대학교 문병로 교수는 학내 벤처 회사인 오픈투자산운용[7]을 설립해 실제 운용을 하고 있으며 2009년 2월 이후 현재까지 261.8%의 수익을 내는 등 매우 높은 연구 성과를 거두고 있다.



그림 1. Tgrape의 Finger Stock



그림 2. Smartforecast의 Bigta

해외에서는 증권사를 중심으로 금융데이터를 이용해 주식예측, 주식 트레이딩 시스템 개발 등을 진행하고 있으며 국내보다 활발하게 딥러닝을 이용하는 실정이다. 미국의 운용사인 골드만삭스자산운용은 AI를 이용해 100만 개에 달하는 애널리스트 리포트를 분석, 주가 자료 탐색 시스템 개발하였다[8]. 뉴욕의 대형 헤지펀드 회사인 브릿지워터 어소시에츠(Bridgewater Associates)는 2012년 말 IBM의 인공지능 컴퓨터 왓슨의 개발자인 데이비드 페루치(David Ferrucci) 박사 등의 전문인력을 영입해 인공지능 주식거래 시스템 개발에 착수하였다[9].

3. 딥러닝 모델 설계

본 연구에서는 텐서플로를 이용해 삼성전자의 일별 주가 증감 패턴을 예측하기 위한 증감패턴 학습모델을 설계 및 개발하고자 한다. 딥러닝은 기계학습과 달리 인간의 개입을 최소화 하고 데이터를 있는 그대로 학습하여 데이터에 대한 특징 또한 기계 스스로 학습한다. 딥러닝을 구현하는 방법으로는 다양하게 제공되는 프레임워크로 구현하는 방법이 있다. 딥러닝 프레임워크는 테아노(Theano)[10], 토치(Torch)[11], 텐서플로우, 카페(Caffe)[12] 등 여러 가지가 있으며 이 중 텐서플로우는 다양한 하드웨어 환경에서 인공신경망 모델을 쉽게 생성하고 학습할 수 있는 인터페이스를 제공하여 해결하고자 하는 문제에 최적화 된 모델 생성에 집중할 수 있다는 장점이 있다[8]. 따라서 본 연구에서는 텐서플로우 이용하여 주식 데이터를 위한 딥러닝 학습모델을 구현한다.

예측모델은 심층신경망(Deep Neural Network, DNN)의 가장 기본적 구현인 다층퍼셉트론(Multilayer Perceptron)을 이용해 구현을 하며, 구조는 그림 3과 같다. 비록 현재는 합성곱 신경망 등의 구조가 아닌 얇은 수준의 신경망이지만, 현재 간단한 주식 데이터와 단순한 신경망의 구성으로도 유의미한 결과를 낼 수 있었으며 텐서플로를 이용해 주식 예측 연구를 진행할 때에 성능 비교 대상이 될 수 있는 연구적 가치가 있다. 네트워크 전체적인 구조는 입력층에는 5개의 뉴런으로 구성되어 있으며

10개의 뉴런으로 구성 된 7개의 은닉층과 2개의 뉴런으로 구성 된 출력층으로 이루어진다.

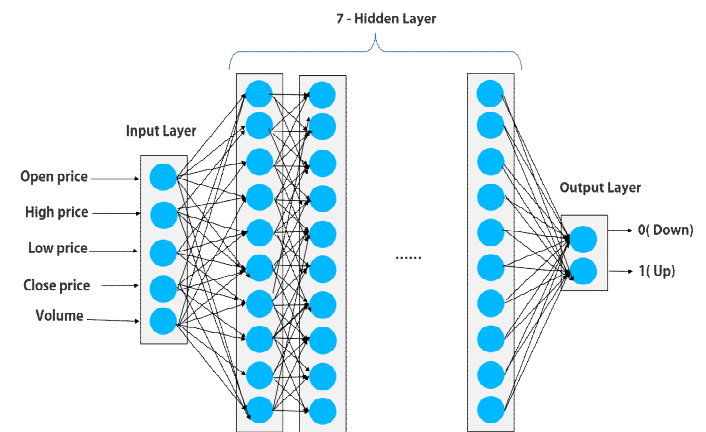


그림 3. 다층 퍼셉트론, (Multilayer Perceptron, MLP)의 구조

4. 실험 및 결과

본 장에서는 주가 변동 예측 학습 모델을 위한 실험 데이터와 환경에 대해 설명한다. 실험 데이터의 경우 최근 매우 주가가 상승한 삼성전자 주식 한 종목의 데이터를 이용해, 주가의 경우 각 종목마다 가격대가 서로 많이 다르기 때문에 학습 시 고려할 사항이 많다. 따라서 본 논문에서는 높은 성능을 내는 데에 초점을 맞추기 위해 삼성전자 한 종목만의 주가 데이터를 이용해 실험을 진행한다. 이 모델은 향후 다양한 종목을 학습할 수 있는 모델 또는 다른 종목에 최적화 된 모델로 발전할 수 있다. 학습을 위한 입력데이터는 일별 시가, 고가, 저가, 종가, 거래량 이렇게 5개의 지표로 구성된다. 이러한 입력 데이터를 주면 단순 주가 패턴을 인식하게 된다. 각 일별 주가 데이터에 대한 결과 값으로는 5일 후의 종가의 상승 또는 하락을 1 또는 0으로 나타낸 값을 설정한다. 상승했을 경우 1이 되며, 하락했을 경우 0이 된다. 약 2000개의 데이터로 학습을 진행하며 약 500개의 데이터로 학습모델의 성능을 검증한다. 심층신경망의 과적합(Overfitting) 문제를 방지하기 위해 Dropout 기법을 사용하였고, 활성화 함수로는 sigmoid, 최적화기는 AdamOptimizer를 사용하였다. 학습률은 0.001로 설정하였으며 500번의 반복학습을 수행하였다. 구현 한 학습모델은 그림 4와 같이 학습을 진행 할수록 점점 오차가 낮아지는 결과가 나왔으나, 크게 줄지는 않았다. 최종적으로 약 56%의 성능을 갖는 학습 모델이 도출되었다. 가공되지 않은 데이터를 사용했음에도 불구하고 50%이상의 성능이 도출되었으며, 비록 딥러닝을 이용한 주가예측 모델 연구[13]에서 구현 된 다층 퍼셉트론 모델을 이용해 측정 된 76.92%의 학습 데이터에 대한 성능에 비해 낮은 성능을 보였으나, 이는 입력데이터를 단순히 제시하여 학습한 결과이기 때문에 입력 데이터의 고도화를 통한 성능 향상을 기대해 볼 수 있다. 또한 추후 연구에 대한 비교 대상으로 사용될 수 있다.

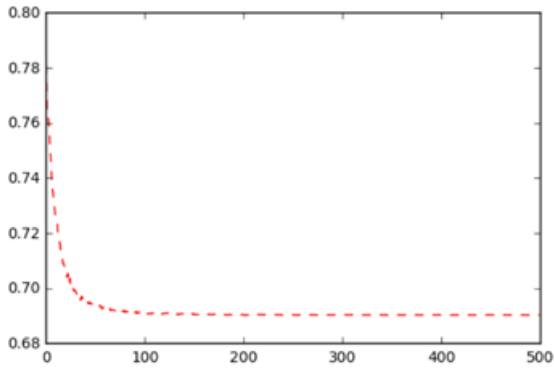


그림 4. 학습 모델 오차 감소율

6. 결론 및 향후 연구

본 논문에서는 5가지 간단한 주가 데이터를 이용해 삼성전자의 증감패턴을 고려하여 설계된 딥러닝 기반 학습 모델을 설계 및 개발하였다. 본 연구의 결과는 약 56% 정도의 성능을 갖는 학습 모델로 도출되었으며, 현재 매우 간단한 구조의 네트워크로 구성되어 있고, 단순한 주가 데이터 특징만을 이용해 실험을 진행하였기 때문에 더욱 성능이 높은 학습모델로 발전할 가능성이 무궁무진하다. 본 연구는 기존의 의미가 없어 보이는 시가, 종가, 고가, 저가 등의 단순 주식 데이터를 통해 우리의 삶에 직접적으로 영향을 주는 가치를 만들어낼 수 있다. 또한 주가 데이터를 이용해 증감패턴을 예측하는 프로그램의 학습 모델을 설계함으로써 프로그램 개발의 기초가 되는 연구가 된다.

추후 연구로는 본 연구에서 개발한 심층 신경망을 합성곱 신경망(Convolutional Neural Network, CNN)이나 순환신경망(Recurrent Neural Network, RNN) 등으로 확장 구현할 수 있다. 합성곱 신경망의 경우 이미지 처리에 매우 높은 성능을 보이는 신경망으로 최근 이미지 처리뿐만 아니라 자연어 처리 등의 여러 유형의 데이터 처리에도 사용되고 있다. 합성곱 신경망은 다른 신경망 구조와 달리 Convolution 및 Pooling 과정을 추가로 수행하게 되는데 Convolution 과정은 데이터의 주위 값들을 반영해 목적하는 작업의 성공률이 높도록 원래 데이터를 변형하는 작업을 뜻한다. Pooling은 잡음 제거와 비슷한 효과를 볼 수 있도록 하는 과정으로 변형된 데이터들의 대표 값들만 취해 성능을 더욱 높이는 과정이다. 따라서 합성곱 신경망을 주식 예측 학습 모델에 적용한다면 보다 높은 성능의 학습 모델을 얻을 수 있을 것이다.

마지막으로 신경망의 입력데이터 측면의 경우 입력 데이터를 설정할 때 시가, 고가, 저가, 종가, 거래량 등의 단순 데이터가 아닌 이동 평균선 등의 주식의 기술적 접근 방식에 따른 다양한 지표를 이용해 입력데이터를 고도화 하여 학습 데이터로 사용할 예정이다. 이를 통해 더욱 성능 높은 주가 예측 모델을 개발할 수 있을 것으로 예상된다.

참고문헌

- [1] 이재원, 고변동 주가 패턴의 감독 학습에 기반한 주식 거래 시스템, 정보과학회, 제19권, 제1호, p167-172, 2013
- [2] 텐서플로우, <https://www.tensorflow.org/>
- [3] 코스콤 주가예측 서비스, http://biz.chosun.com/site/data/html_dir/2014/03/27/2014032704225.html
- [4] FingerStock, <https://ko-kr.facebook.com/fingerstock/>
- [5] Bigta, <https://ko-kr.facebook.com/bigtasnu/>
- [6] Deepnumbers, <http://www.deepnumbers.com/>
- [7] 옵투스자산운용, <http://www.optus.co.kr/>
- [8] Goldman Sachs robo-advisor, <http://www.afr.com/personal-finance/managed-funds/goldman-sachs-buys-robo-investment-adviser-20160314-gnj069>
- [9] 브리지워터 헤지펀드, <http://www.yonhapnews.co.kr/bulletin/2016/12/23/0200000000AKR20161223070600009.HTML>
- [10] 테아노, <http://deeplearning.net/software/theano/>
- [11] 토치, <http://torch.ch/>
- [12] 카페, <http://caffe.berkeleyvision.org/>
- [13] 이지훈, 딥러닝을 이용한 주가 예측 모델, 학위논문, 숭실대학교, 2016