

R 신경망을 이용한 주가예측 모의실험

Stock Price Estimation Experiments using R Neural Nets

저자 (Authors)	우미영, 박소영, 한영우, 박우창 Miyoung Woo, Soyoung Park, Youngwoo Han, Uchang Park
출처 (Source)	Proceedings of KIIT Conference , 2013.5, 498-500(3 pages)
발행처 (Publisher)	한국정보기술학회 Korean Institute of Information Technology
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02172742
APA Style	우미영, 박소영, 한영우, 박우창 (2013). R 신경망을 이용한 주가예측 모의실험. Proceedings of KIIT Conference, 498-500
이용정보 (Accessed)	송실대학교 203.253.***.153 2020/09/29 18:05 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 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.

R 신경망을 이용한 주가예측 모의실험

우미영***, 박소영***, 한영우**, 박우창*

Stock Price Estimation Experiments using R Neural Nets

Miyoung Woo***, Soyoung Park***, Youngwoo Han**, and Uchang Park*

요 약

주식시장에서 주식의 과거 데이터를 이용하여 주가를 예측하여 자동으로 거래를 시도하는 분야는 비즈니스에서 높은 관심사항이다. 본 연구는 개별 기업의 주가 예측을 분기별로 주식의 내재변수인 PER, EPS, ROE 등 4개 지표와 거시지표인 다우지수, KOSPI 지수 2개의 지표를 이용하여 R 언어의 신경망 예측 모델 nnet를 이용하여 예측을 시도하였다. 실험에서 5개의 대표적인 기업의 과거 주가를 분기별로 조사하고 예측 결과를 실제 값들과 비교한 결과를 제시하였다.

Abstract

In stock market, stock price estimating for automatic trading using past trading data is a difficult and long term interesting topic. In this research, we apply R neural net model to a company stock price estimation using input variables, PER, EPS, ROE, with macro input variables DOW JONES INDEX, KOSPI. In the experiment, 5 major company stock price is estimated and compared.

Key words

R, stock price, neural net

1. 서 론

주식은 많은 사람으로부터 관심을 받는 분야이며, 이를 예측하는 방법에는 많은 연구가 이루어져 왔다. 연구에서 사용되어온 주 예측 방법들은 ANN(artificial neural network), SVM(support vector

machine)이 주로 이용되고 있다[1][2].

본 연구에서는 R에서 제공되는 신경망 패키지를 이용하여 주가예측을 시도하였다. 신경망 알고리즘은 연상기억의 특성을 이용한 알고리즘으로서 test 집합의 반복적인 학습을 통하여 규칙을 발견하고 예측할 수 있다. 그러므로 수학적으로 표현하기 어

* 덕성여자대학교 컴퓨터학과 교수

** 한국예탁결제원

*** 덕성여자대학교 컴퓨터학과 학부생

렵거나 관계를 파악하기 힘들 때, 입출력 패턴의 관계를 충분히 생성할 test 집단의 자료가 있다면 컴퓨터에 의해 일반적인 관계추출이 가능하다.

본 연구에서 예측하고자 하는 주가지수는 기업내부의 주당 순이익(EPS), 주가 수익배수(PER), 자기자본 이익률(ROE) 등의 영향 뿐 아니라 환율, 금리, 타 국가의 국가지수 등 외부영향 등과 상호 연관성을 갖는다. 하지만 그들의 관계나 각각의 가중치를 파악하고, 생성된 모델을 통하여 수학적으로 주가를 정확히 예측해내는 것 또한 한계점이 있다[3]. 따라서 본 연구에서는 실험적으로 각 기업들의 과거 데이터들을 test 집단으로 설정해 신경망을 통해 학습한 이후 주가 예측을 시도해보았다.

II. 주가 예측 모델

예측을 위한 기업은 한국 증시에 상장되어 있는 1000여 개의 종목 중, 대표적인 5개 종목을 대상으로 하였다. SK, LG, CJ, KT, 삼성전자를 선정하였고 위의 코드는 SK(주)의 주가를 예시로 하였다. 예측을 위한 데이터는 2002년 3월부터 2012년 9월까지의 분기별 주가의 평균이다.

신경망 모델의 입력층에 쓰이는 데이터는 2002년 3월부터 2012년 6월까지 3개월 간격으로 분기마다 평균값을 사용하였다. 입력노드는 크게 두 부분으로 기업 내의 상황을 나타내는 미시적 요인 4개와, 기업의 외적인 거시적 요인 2개로 분류 하였다. 미시적 요인에는 주가 수익 배수(PER), 자기자본 이익률(ROE), 매출액증가율(GRS), 주당 순이익(EPS)이, 거시적 요인에는 다우지수(DOW JONES INDEX)와 코스피(KOSPI) 지수가 포함되었다.

예측을 위한 R 프로그램의 일부분은 다음과 같다. 입력 데이터를 구성하고 nnet를 호출하여 모델을 트레이닝한 후, predict를 호출하여 예측을 시도한다[4][5].

```
>sk1=nnet(PREDICT~+PER+ROE+GRS+EPS+DOW+
KOSPI,data=sk,size=4,decay=5e-3)
> p=predict(sk1,predict_sk,type="class")
```

그리고 알맞은 신경망 모델을 구축하기 위해 약 800개의 신경망 학습 데이터를 수집하였고 그 데이터 중 400개는 학습 집합으로 나머지 400개는 테스트 집합으로 분류하였다. 그리고 적합한 size와 decay 값을 찾기 위하여 다양한 신경망 모델을 만든 후 오분류를 분석을 통하여 가장 성능이 좋았던 모델의 size와 decay 값을 선택하였다.

구축한 신경망모델의 구조에서 은닉층에는 4개의 은닉노드가 있다. 입력층에는 주가를 결정하는 여러 요인들 중에 선정한 6개의 입력노드가 있고, 출력층에는 2012년 9월의 주가지수를 예측하는 1개의 출력노드가 있다. 2012년 9월 주가지수를 출력노드가 buy인 경우는 상승을, hold인 경우 유지를, sell인 경우 하락을 예측한다는 의미이다.

P1: 예측하고자 하는 분기의 주가

P2: 예측하고자 하는 분기의 전 분기 주가

$$\text{predict} \begin{cases} \text{sell} & (P1/P2 < 0.9) \\ \text{hold} & (0.9 \leq P1/P2 \leq 1.1) \\ \text{buy} & (P1/P2 > 1.1) \end{cases}$$

표 1은 은닉층의 size를 4로, decay를 5e-3으로 두었을 때의 예측 결과를 나타낸 것이다.

표 1 : nnet 모델의 예측 빈도표

Table1 : Frequency ratio of nnet model prediction

	buy	hold	sell
buy	35	25	15
hold	30	133	21
sell	39	33	69

III. 예측 결과 및 결론

실험 결과 각각의 기업에 대한 buy, hold, sell의 확률을 표 2와 같이 예측하고 있으며 반올림하여 소수 다섯째자리로 표현하였다.

표 2. 5개 기업 주가 예측 확률

Table 2. 5 Companies stock price estimation probability

기업	buy	hold	sell
SK	0.51749	0.257564	0.19350
LG	0.40238	0.42335	0.13976
CJ	0.63449	0.27608	0.04859
KT	0.72856	0.16137	0.00011
SAMSUNG	0.34885	0.51161	0.11629

실험을 통하여 hold를 제외한 buy와 sell의 예측 정확도(Precision)은 약 0.5 정도로 나타났다. 주가를 예측하기 작업은 어려우며 주가 예측을 위한 변수를 결정하는 작업은 본 연구의 범위를 넘기 때문에 알려진 최소한의 변수를 이용하여 실험을 하였다. 또 본 연구의 예측 실험의 기업의 수와 예측 실험의 횟수가 작아서 정확한 예측이 어렵지만 주어진 한계 내에서 신경망을 이용하여 실험을 시도하였다. 예측 기간과 예측 횟수를 늘여서 모델에 대한 신뢰도를 높이기 위한 검증이 필요하다.

참 고 문 헌

- [1] Luis Torgo, "Data Mining with R", CRC Press, 2011.
- [2] Yangchang Zhao, "R and Data Mining", Elsevier, 2013.
- [3] 임도형, 이일병, "신경망을 사용한 매도/매수 주식 종목 선정", 대한산업공학회/한국경영과학회 2000 춘계공동학술대회 논문집. 2000년 4월 21-22일
- [4] 박창이, 김진석, "R을 이용한 데이터 마이닝", 2008년 12월 11일
- [5] Robert Kabacoff, "R in Action", Manning 2011.