

# ỨNG DỤNG MỘT SỐ MÔ HÌNH HỌC MÁY TRONG DỰ BÁO CHIỀU BIẾN ĐỘNG CỦA THỊ TRƯỜNG CHỨNG KHOÁN VIỆT NAM

*Lê Văn Tuấn, Nguyễn Thu Thủy, Lê Thị Thu Giang*

Bộ môn Toán, Đại học Thương mại, Hà Nội, Việt Nam

---

**Tóm tắt.** Bài viết sử dụng một số mô hình/thuật toán học máy để dự báo xu hướng biến động (tăng/giảm) của chỉ số thị trường chứng khoán của Việt Nam. Kết quả cho thấy, trong các mô hình hồi quy Logistic, mô hình phân tích phân biệt tuyến tính (LDA), phân tích phân biệt toàn phương (QDA) và mô hình K – lân cận (KNN): mô hình KNN(10) có độ chính xác dự báo tốt nhất. Câu lệnh R được cung cấp đầy đủ tới bạn đọc.

---

## 1. MỞ ĐẦU

Chuỗi chỉ số thị trường chứng khoán thường được đặc trưng bởi một hành vi theo đám đông và phi tuyến tính khiến cho việc dự báo trở thành một nhiệm vụ đầy thách thức. Các yếu tố tạo ra sự biến động trong lĩnh vực này rất phức tạp và bắt nguồn từ nhiều nguyên nhân với bản chất khác nhau. Từ các quyết định kinh tế, chính trị và đầu tư cho đến những nguyên nhân không rõ ràng, ở một khía cạnh nào đó, đều gây ra nhiều khó khăn trong dự đoán chỉ số thị trường chứng khoán. Thị trường chứng khoán luôn thu hút các nhà đầu tư do khả năng sinh lời cao, tuy nhiên, nó cũng chứa đựng rất nhiều rủi ro. Vì vậy, cần có một công cụ thông minh để giảm thiểu rủi ro với hy vọng có thể tối đa hóa lợi nhuận. Ngày nay, các mô hình Học máy (Machine Learning) đã trở thành một công cụ phân tích mạnh mẽ được sử dụng để trợ giúp và quản lý đầu tư hiệu quả. Các mô hình này đã được sử dụng rộng rãi trong lĩnh vực tài chính để cung cấp những phương pháp mới nhằm giúp các nhà đầu tư đưa ra quyết định tốt hơn trong việc đầu tư chứng khoán.

Thị trường chứng khoán mỗi quốc gia ngày càng phát triển sâu rộng và hòa nhập vào thị trường quốc tế vì chúng có lợi nhuận hấp dẫn (He và cộng sự, 2015; Chou và Nguyễn, 2018) và là một tài sản có tính thanh khoản tương đối cao do chúng có thể được mua - bán nhiều lần thông qua các sở giao dịch chứng khoán. Mặc dù vậy, đầu tư cổ phiếu có rủi ro cao do sự không chắc chắn và biến động trên thị trường chứng khoán (Hyndman và Athanasopoulos, 2018). Do đó, các nhà đầu tư phải hiểu bản chất của từng cổ phiếu riêng lẻ và các yếu tố phụ thuộc của chúng ảnh hưởng đến giá cổ phiếu để tăng cơ hội đạt được

lợi nhuận cao hơn. Quan trọng nhất, các nhà đầu tư cần phải đưa ra các quyết định đầu tư hiệu quả vào đúng thời điểm (Ijegwa và cộng sự, 2014) bằng cách sử dụng lượng thông tin chính xác và thích hợp (Nguyen và cộng sự, 2015).

Dự báo sự biến động trên thị trường chứng khoán là một chủ đề quan trọng trong lĩnh vực tài chính. Việc dự báo hiệu quả sẽ giúp nhà đầu tư xây dựng được chiến lược đầu tư tối ưu cũng như phòng ngừa rủi ro. Dự báo một số chỉ số tài chính dựa trên một số yếu tố tác động sẽ dễ dàng nhưng kết quả có thể không chính xác vì các yếu tố chưa được đưa vào mô hình cũng có thể quan trọng trong việc giải thích sự biến động của chỉ số tài chính. Chẳng hạn, giá của cổ phiếu hoặc chỉ số của thị trường có thể bị ảnh hưởng bởi nhiều yếu tố khác nhau, ví dụ: tăng trưởng kinh tế (Perwej và Perwej, 2012; Selvin và cộng sự, 2017). Rất khó để phân tích tất cả các yếu tố theo cách thủ công (Nguyen và cộng sự, 2015; Sharma và cộng sự, 2017), vì vậy, sẽ tốt hơn nếu có các công cụ hỗ trợ phân tích dữ liệu này. Trong đầu tư chứng khoán, việc đưa ra quyết định đúng đắn trong khoảng thời gian kịp thời là một thách thức lớn vì cần một lượng thông tin đồ sộ để dự đoán sự biến động của giá thị trường chứng khoán. Những thông tin này rất quan trọng đối với các nhà đầu tư vì sự biến động của thị trường chứng khoán có thể dẫn đến tổn thất đầu tư đáng kể. Do đó, việc phân tích thông tin lớn này rất hữu ích cho các nhà đầu tư và cũng hữu ích cho việc phân tích xu hướng biến động của các chỉ số thị trường chứng khoán (Kim và Kang, 2019).

Với sự thành công rực rỡ của các mô hình Học máy (Machine Learning) trong nhiều lĩnh vực, ứng dụng của Học máy trong tài chính ngày càng được chú ý và phát triển không ngừng (Nguyen et al., 2015; Attigeri et al., 2015; Kim and Kang, 2019). Việc áp dụng các mô hình Học máy để dự báo chứng khoán, cổ phiếu cũng rất phổ biến trên thế giới, không chỉ trong học thuật mà còn trong thực tiễn. Ở các bài toán dự báo ở những nghiên cứu trước đây, các tác giả thường sử dụng dữ liệu lịch sử của chuỗi thời gian để cung cấp các ước lượng cho các giá trị trong tương lai. Gần đây, các nhà nghiên cứu đã bắt đầu phát triển các kỹ thuật Học máy giống với quá trình sinh học và tiến hóa để giải quyết các vấn đề phức tạp và phi tuyến. Kỹ thuật này trái ngược với cách tiếp cận truyền thống, tức là các phương pháp thống kê cổ điển. Ví dụ về các kỹ thuật học máy có thể nhắc đến như Mạng thần kinh nhân tạo (Artificial Neural Networks - ANN), Máy vectơ hỗ trợ (Support Vector Machines - SVM) và Lập trình di truyền (Genetic Programming - GP), thuật toán K-lân cận (K-nearest neighbors - KNN), Mô hình hồi quy Logistic, phân tích phân biệt tuyến tính (Linear Discriminant Analysis - LDA), phân tích phân biệt toàn phương (Quadratic Discriminant Analysis - QDA).

Bài báo này sẽ khám phá một ứng dụng của Học máy trong tài chính. Cụ thể, chúng tôi sẽ sử dụng các mô hình Học máy quen thuộc để áp dụng dự báo cho thị trường chứng khoán Việt Nam. Kết quả cho thấy, trong các mô hình hồi quy Logistic, LDA, QDA, và KNN: mô hình KNN(10) có độ chính xác dự báo tốt nhất.

## 2. MỘT SỐ MÔ HÌNH HỌC MÁY

Vấn đề dự báo tăng/giảm của thị trường chứng khoán thực ra là bài toán phân lớp (2 lớp) trong đó:

Biến đầu ra:  $Y \in \{Up, Down\}$

Biến đầu vào:  $X = (X_1, X_2, \dots, X_p)$ .

Bài toán sẽ dự báo  $Y$  thuộc lớp Up, với  $Y$  tương ứng với đầu vào  $x_0$ , nếu:

$$\Pr(Y = Up|X = x_0) > 0.5; \text{ và ngược lại với lớp Down.}$$

### 2.1. Phương pháp K- láng giềng gần nhất (KNN)

Về mặt lý thuyết, chúng ta thường muốn dự báo các hành vi bằng cách sử dụng công thức Bayes. Nhưng đối với dữ liệu thực tế, chúng ta không biết phân phối xác suất có điều kiện của  $Y$  khi đã biết  $X$ , và vì vậy việc tính toán sử dụng công thức Bayes là không thể thực hiện được. Do đó, công thức xác suất Bayes được xem như một tiêu chuẩn vàng khó có thể đạt được. Chúng ta cần tìm kiếm các phương pháp khác khả thi hơn. Có nhiều cách tiếp cận hướng tới mục tiêu ước lượng phân phối xác suất có điều kiện của  $Y$  khi đã biết  $X$ , và sau đó phân loại các quan sát với xác suất ước tính cao nhất. Một trong những phương pháp đó là phương pháp K-láng giềng gần nhất. Cho trước một số nguyên dương  $K$  và một quan sát thử nghiệm  $x_0$ , phân lớp KNN đầu tiên xác định  $K$  điểm trong dữ liệu định hướng gần nhất với  $x_0$ , được ký hiệu là bằng  $N_0$ . Sau đó, xác suất có điều kiện cho lớp Up là tỷ lệ các điểm trong  $N_0$  có giá trị bằng Up:

$$\Pr(Y = Up|X = x_0) = \frac{1}{K} \sum_{i \in N_0} I(y_i = Up)$$

Với  $K$  là hằng số cho trước,  $N_0$  là  $K$  điểm gần  $x_0$  nhất trong tập dữ liệu,  $I$  là biến chỉ báo. Trong bước cuối cùng, KNN sử dụng công thức Bayes và phân lớp các quan sát thử nghiệm với các lớp có xác suất lớn nhất.

Một số ưu điểm của mô hình KNN là dễ sử dụng và cài đặt, việc dự báo kết quả của dữ liệu mới dễ dàng, và độ phức tạp tính toán vừa phải. Tuy nhiên, phương pháp KNN còn tồn tại một số nhược điểm như KNN nhiều dễ đưa ra kết quả không chính xác khi  $K$  nhỏ, ngoài ra cần thời gian lưu tập huấn luyện, khi tập huấn luyện và tập kiểm tra tăng lên nhiều sẽ mất thời gian tính toán.

## 2.2. Mô hình hồi quy Logistic

Phân tích hồi qui logistic là một kỹ thuật thống kê để xem xét mối liên hệ giữa biến độc lập (biến số hoặc biến phân loại) với biến phụ thuộc là biến nhị phân. Trong hồi qui logistic, biến phụ thuộc  $y$  chỉ có 2 trạng thái 1 (ví dụ tăng) và 0 (ví dụ giảm). Muốn đổi ra biến số liên tục người ta tính xác suất của 2 trạng thái này. Nếu gọi  $p$  là xác suất để một biến cố xảy ra (ví dụ: tăng), thì  $1-p$  là xác suất để biến cố không xảy ra (ví dụ: giảm). Ký hiệu:  $p(X) = \Pr(Y = Up|X)$ . Mô hình mô hình hồi quy Logistic có dạng:

$$\log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$$

Hay

$$p(X) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}$$

Trong đó:  $\beta_0, \beta_1, \dots, \beta_p$  là các hệ số cần ước lượng.

Mô hình và ý tưởng này khá đơn giản nhưng cho kết quả rất khả quan và được áp dụng rất nhiều trong cuộc sống.

## 2.3. Phân tích phân biệt tuyến tính (LDA)

LDA (Linear Discriminant Analysis) là một phương pháp giảm chiều dữ liệu cho bài toán phân lớp. LDA có thể được coi là một phương pháp giảm chiều dữ liệu (dimensionality reduction), và cũng có thể được coi là một phương pháp phân lớp (classification), và cũng có thể được áp dụng đồng thời cho cả hai, tức giảm chiều dữ liệu sao cho việc phân lớp hiệu quả nhất. Số chiều của dữ liệu mới là nhỏ hơn hoặc bằng  $C-1$  trong đó  $C$  là số các lớp. Từ 'Discriminant' được hiểu là những thông tin đặc trưng cho mỗi lớp, khiến nó không bị lẫn với các lớp khác. Từ 'Linear' được dùng vì cách giảm chiều dữ liệu được thực hiện bởi một ma trận chiếu (projection matrix), là một phép biến đổi tuyến tính (linear transform).

Ký hiệu:

$\pi_{Up}$  là xác suất (tiên nghiệm) một quan sát ngẫu nhiên thuộc lớp  $Up$  (tương tự cho  $\pi_{Down}$ )

$f_{Up}(X) = \Pr(X = x|Y = Up)$  là hàm mật độ của  $X$  cho một quan sát thuộc lớp  $Up$  (tương tự cho  $f_{Down}(X)$ ).

Trong mô hình LDA:

$$\Pr(Y = Up|X = \mathbf{x}) = \frac{\pi_{Up} f_{Up}(\mathbf{x})}{\pi_{Up} f_{Up}(\mathbf{x}) + \pi_{Down} f_{Down}(\mathbf{x})}$$

Mô hình LDA giả thiết quan sát  $X$  thuộc lớp  $Up$  có phân phối chuẩn nhiều chiều:  $N(\mu_{Up}, \Sigma)$ , với  $\mu_{Up}$  là vectơ trung bình của lớp  $Up$ ,  $\Sigma$  là ma trận hiệp phương sai chung của cả 2 lớp (tương tự cho quan sát  $X$  thuộc lớp  $Down$ ).

Có thể nói, LDA là một phương pháp giảm chiều dữ liệu có sử dụng thông tin về nhãn của của dữ liệu. Ý tưởng cơ bản của LDA là tìm một không gian mới với số chiều nhỏ hơn không gian ban đầu sao cho hình chiếu của các điểm trong cùng 1 lớp lên không gian mới này là gần nhau trong khi hình chiếu của các điểm của các lớp khác nhau là khác nhau. LDA có giả sử ngầm rằng dữ liệu của các lớp đều tuân theo phân phối chuẩn và các ma trận hiệp phương sai của các lớp là gần nhau. LDA hoạt động rất tốt nếu các lớp là tuyến tính riêng biệt (linear seperable), tuy nhiên, chất lượng mô hình giảm đi rõ rệt nếu các lớp là không tuyến tính riêng biệt. Điều này dễ hiểu vì khi đó, chiếu dữ liệu lên phương nào thì cũng bị chồng lấn, và việc tách biệt không thể thực hiện được như ở không gian ban đầu.

#### 2.4. Phân tích phân biệt toàn phương (QDA)

Mô hình QDA được xây dựng tương tự như LDA, sự khác biệt là cho phép các quan sát ở mỗi lớp có hiệp phương sai khác nhau. Nghĩa là giả thiết  $X$  thuộc lớp  $Up$  có phân phối chuẩn nhiều chiều:  $N(\mu_{Up}, \Sigma_{Up})$  (tương tự cho quan sát  $X$  thuộc lớp  $Down$ ).

### 3. THỰC NGHIỆM TRÊN TTCK VIỆT NAM

Phần này sẽ thực hiện so sánh độ chính xác trong dự báo của các mô hình/thuật toán học máy được giới thiệu trong Mục 2. Độ chính xác được tính bằng: số lần dự báo đúng/tổng số quan sát (trong bộ dữ liệu test). (Các kết quả được các tác giả thực hiện trên R, câu lệnh tham khảo trong (James và cộng sự, 2013)).

#### 3.1. Dữ liệu

Bài viết sử dụng dữ liệu giao dịch theo ngày của chỉ số VNIndex (chỉ số đại diện cho TTCK Việt Nam), download từ trang web của Công ty Cổ phần Chứng khoán VNDIRECT. Dữ liệu được lấy từ ngày 26 tháng 8 năm 2016 đến ngày 26 tháng 8 năm 2020, gồm 1000 quan sát. Bộ dữ liệu trên được chia thành 2 tập, tập huấn luyện (training set) từ ngày 26 tháng 8 năm 2016 đến ngày 26 tháng 8 năm 2019 (gồm 748 quan sát), tập kiểm tra (test set) từ ngày 27 tháng 8 năm 2019 đến ngày 26 tháng 8 năm 2020 (gồm 252 quan sát).

Cấu trúc dữ liệu có 9 biến (tương tự như các biến được sử dụng trong (James và cộng sự, 2013)):

DATE: ngày giao dịch

Volume: khối lượng cổ phiếu giao dịch

Today: lợi suất % trong ngày của VNIndex

Direction: chiều tăng/giảm (Up/Down) của chỉ số VNIndex

Lag1, ..., Lag5 là lợi suất % của các ngày thứ 1, ..., thứ 5 liền trước ngày hiện tại

### 3.2. Độ chính xác trong dự báo của mô hình hồi quy Logistic

Trước hết, ta sẽ sử dụng mô hình hồi quy Logistic với biến đầu ra là Direction, các biến đầu vào là Volume, Lag1, ..., Lag5.

Kết quả ước lượng các tham số là:

```
Coefficients:
      Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.541e-01  2.553e-01 -0.603  0.54618
Lag1         -2.574e+01  8.895e+00 -2.894  0.00381 **
Lag2          1.111e+01  8.734e+00  1.272  0.20323
Lag3         -8.581e-01  8.871e+00 -0.097  0.92294
Lag4          1.262e+01  8.804e+00  1.434  0.15158
Lag5          1.452e+01  8.700e+00  1.669  0.09515 .
Volume        1.852e-09  1.598e-09  1.159  0.24630
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Giá trị của P-value cho thấy chỉ có biến Lag1 là có ý nghĩa thống kê.

Kết quả dự báo:

```
glm.pred Down Up
Down     26  33
Up       92 101
```

Độ chính xác dự báo: 0,5039683

Như vậy độ chính xác dự báo của mô hình khá thấp, xấp xỉ 50%.

Ta sẽ lặp lại các bước trên tuy nhiên đầu vào chỉ giữ lại Lag1 (là biến duy nhất có ý nghĩa thống kê).

Kết quả dự báo:

```
glm.pred Down Up
Down     27  29
Up       91 105
```

Độ chính xác dự báo: 0,5238095

Kết quả đã tăng đáng kể với độ chính xác hơn 52%.

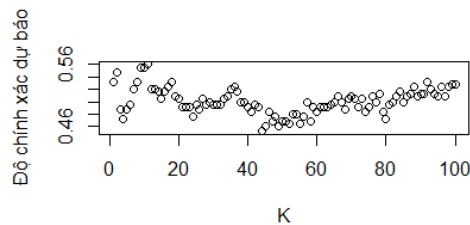
### 3.3. So sánh độ chính xác trong dự báo của các mô hình

Phần này trình bày độ chính xác dự báo của các mô hình/thuật toán trong được giới thiệu trong Mục 2 với đầu ra là Direction, đầu vào là Lag1.

Độ chính xác dự báo:

| Mô hình/thuật toán | Độ chính xác dự báo |
|--------------------|---------------------|
| Hồi quy Logistic   | 0,5238095           |
| LDA                | 0,5198413           |
| QDA                | 0,5277778           |
| KNN(10)            | 0,5555556           |

Chúng tôi chọn KNN(10) vì nó cho kết quả dự báo tốt nhất khi K từ 1 đến 100 (hình dưới).



Như vậy, giải thuật KNN(10) cho kết quả dự báo tốt nhất trên TTCK Việt Nam, xấp xỉ 0,556 (kết quả này tốt hơn kết quả dự báo của các mô hình khi đầu vào là tất cả các biến, xem Phụ lục A). Kết luận này khác với kết quả dự báo cho TTCK Mỹ trong (James và cộng sự, 2013): mô hình/giải thuật tốt nhất là QDA, xấp xỉ 0,599. Có thể kết luận là TTCK Việt Nam khó dự báo hơn TTCK Mỹ.

Lưu ý rằng bài viết chưa sử dụng các mô hình quen thuộc khác như ANN, SVM để so sánh độ chính xác dự báo. Tuy nhiên, hiệu quả dự báo của các mô hình/thuật toán trong mục 2 tốt hơn so với mô hình ANN khi dự báo các cổ phiếu trên TTCK Mỹ (Rivera, 2018).

#### 4. KẾT LUẬN

Như vậy, sự biến động tăng/giảm của TTCK Việt Nam chỉ phụ thuộc vào sự tăng/giảm của ngày ngay trước mà không phụ thuộc vào các ngày xa hơn, và cũng không phụ thuộc vào khối lượng giao dịch.

Kết quả so sánh các mô hình/thuật toán hồi quy Logistic, LDA, QDA và KNN cho thấy thuật toán KNN(10) có độ chính xác dự báo tốt nhất. Độ chính xác của KNN(10) xấp xỉ 55,6%, kết quả này là một bằng chứng phủ định giả thuyết Thị trường hiệu quả của Eugene

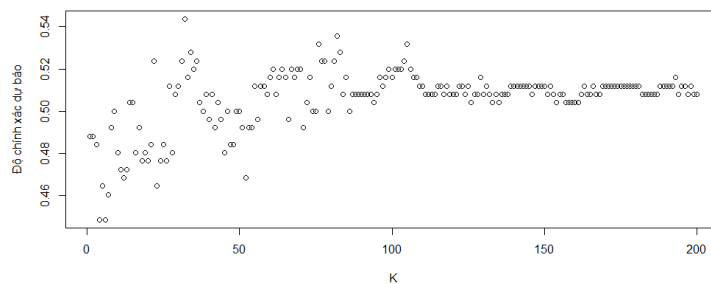
Fama (1960) trên TTCK Việt Nam – giả thuyết này cho rằng “không ai dự đoán được” và “chuyển động của giá cổ phiếu là hoàn toàn ngẫu nhiên”.

#### PHỤ LỤC A.

Độ chính xác dự báo với đầu vào là tất cả các biến:

| Mô hình/thuật toán | Độ chính xác dự báo |
|--------------------|---------------------|
| Hồi quy Logistic   | 0.5039683           |
| LDA                | 0.5                 |
| QDA                | 0.4960317           |
| KNN                | < 0.54              |

Độ chính xác dự báo của KNN với đầu vào là tất cả các biến:



#### PHỤ LỤC B. CÂU LỆNH R:

```
library(MASS)
library(class)
vnindex<-read.csv("D:\\data\\ML\\VNIndex.csv", header=TRUE)
test<-vnindex[1:252,]
train<-vnindex[253:1000,]
####
glm.fit =glm ( Direction~Lag1+ Lag2+ Lag3+Lag4+ Lag5+Volume , data=train , family = binomial )
summary(glm.fit)
glm.probs = predict (glm.fit , test , type ="response")
glm.pred=rep ("Down" ,252)
glm.pred[glm.probs >.5]="Up"
mean(glm.pred == test$Direction)
####
glm.fit =glm ( Direction~Lag1 , data=train , family = binomial )
glm.probs = predict (glm.fit , test , type ="response")
glm.pred=rep ("Down" ,252)
glm.pred[glm.probs >.5]="Up"
mean(glm.pred == test$Direction)
#####
```



```

lda.fit=lda ( Direction~Lag1 , data= train )
lda.pred= predict (lda.fit , test)
lda.class =lda.pred$class
mean(lda.class == test$Direction)
#####
qda.fit =qda ( Direction~Lag1 , data = train )
qda.class = predict (qda.fit , test)$class
mean(qda.class == test$Direction)
#####
knn.pred=knn(cbind(train$Lag1), cbind(test$Lag1),train$Direction ,k=10)
mean(knn.pred == test$Direction)
#####
x <- c(1:100)
results <- 0
for (val in x) {
  knn.pred=knn(cbind(train$Lag1), cbind(test$Lag1),train$Direction ,k=val)
  results <-c(results,mean(knn.pred == test$Direction))
}
results<-results[-1]
plot(x,results, xlab="K", ylab="Độ chính xác dự báo")

```

## TÀI LIỆU THAM KHẢO

Attigeri, G. V., MM, M. P., Pai, R. M., and Nayak, A. (2015), “Stock market prediction: A big data approach”, In *TENCON 2015-2015 IEEE Region 10 Conference*, pages 1-5. IEEE.

- [1] Chou, J.-S. and Nguyen, T.-K. (2018),” Forward Forecast of Stock Price Using Sliding-Window Metaheuristic - Optimized Machine-Learning Regression”, *IEEE Transactions on Industrial Informatics*, 14(7):3132- 3142.
- [2] He, J., Cai, L., Cheng, P., and Fan, J. (2015), “Optimal investment for retail company in electricity market”, *IEEE Transactions on Industrial Informatics*, 11(5):1210-1219.
- [3] Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- [4] Ijegwa, A. D., Rebecca, V. O., Olusegun, F., and Isaac, O. O. (2014). “A predictive stock market technical analysis using fuzzy logic”, *Computer and information science*, 7(3):1.
- [5] James, G., Witten, D., Hastie, T. & Tibshirani, R., *An Introduction to Statistical Learning: with Applications in R* (Springer Texts in Statistics) 1st ed. 2013, Corr. 7th printing 2017 Edition.

- [6] Kim, S. and Kang, M. (2019). “Financial series prediction using Attention LSTM”, *arXiv preprint arXiv:1902.10877*.
- [7] Nguyen, T. H., Shirai, K., and Velcin, J. (2015). “Sentiment analysis on social media for stock movement prediction”, *Expert Systems with Applications*, 42(24):9603- 9611.
- [8] Perwej, Y. and Perwej, A. (2012), “Prediction of the Bombay Stock Exchange (BSE) market returns using artificial neural network and genetic algorithm”, *Journal of Intelligent Learning Systems and Applications*, 4(02):108.
- [9] Rivera, P. (2018). “Forecasting stock movements with Artificial Neural Networks in R”: <https://medium.com/@paul90.hn/forecasting-stock-movements-with-artificial-neural-networks-in-r-f60f97ca7940>
- [10] Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V. K., and Soman, K. (2017). “Stock price prediction using LSTM, RNN and CNN-sliding window model”, In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 1643-1647. IEEE.
- [11] Sharma, A., Bhuriya, D., and Singh, U. (2017). “Survey of stock market prediction using machine learning approach”, In *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)*, volume 2, pages 506-509. IEEE.