

시장 내 자금 흐름 변수를 활용한 변동성 지표 예측

경제학과 김겨레
경제학 석사과정 김민호

Index

I. Research Question

II. Variables

III. Model : Multi Layer LSTM

IV. Results

- Boruta Selection Result
- RMSE, MAE
- GW Test
- MCS Test

V. Contribution

VI. Related Literature

0. Outline

1. VKOSPI 지수 예측을 목표로 하며(Y)
2. 주식시장 외부 변수, 주식시장 자금흐름 변수, 주식시장 매매동향 변수를 구분하여(X)
3. Multi Layer LSTM를 통한 예측 실시
* 위 그룹화된 X 변수를 각 Layer에 순차적으로 적용
4. HAR, Boruta-RF, Boruta-XGB 등의 예측방법과 RMSE, MAE 비교, GW Test, MCS Test를 통한 검증을 실시하고,
5. 우리의 변수와 예측방법이 VKOSPI 예측에 우수하게 활용될 수 있음을 보이며,
6. VKOSPI의 정확한 예측을 통해 향후 시장의 움직임에 대해 유연하게 대응할 수 있도록 하는데 기여하고자 한다.

I. Research Question

1. Research Question

- : 어떻게 하면 주식 시장의 변동성을 더욱 잘 예측할 수 있을 것인가?
- : 더 정확한 VKOSPI 예측 → 시장 방향에 대한 예측력 상승
- : 이에 따르는 시장 참여자들의 선제적이거나, 유연한 대응 기대

2. Why, VKOSPI?

- : 변동성에 대한 시장 참여자들의 기대
- : 다른 보조지표가 필요 없는 유용한 지표
- : VKOSPI를 활용한 연구에 비해 그 자체를 예측하는 연구는 상대적으로 드물

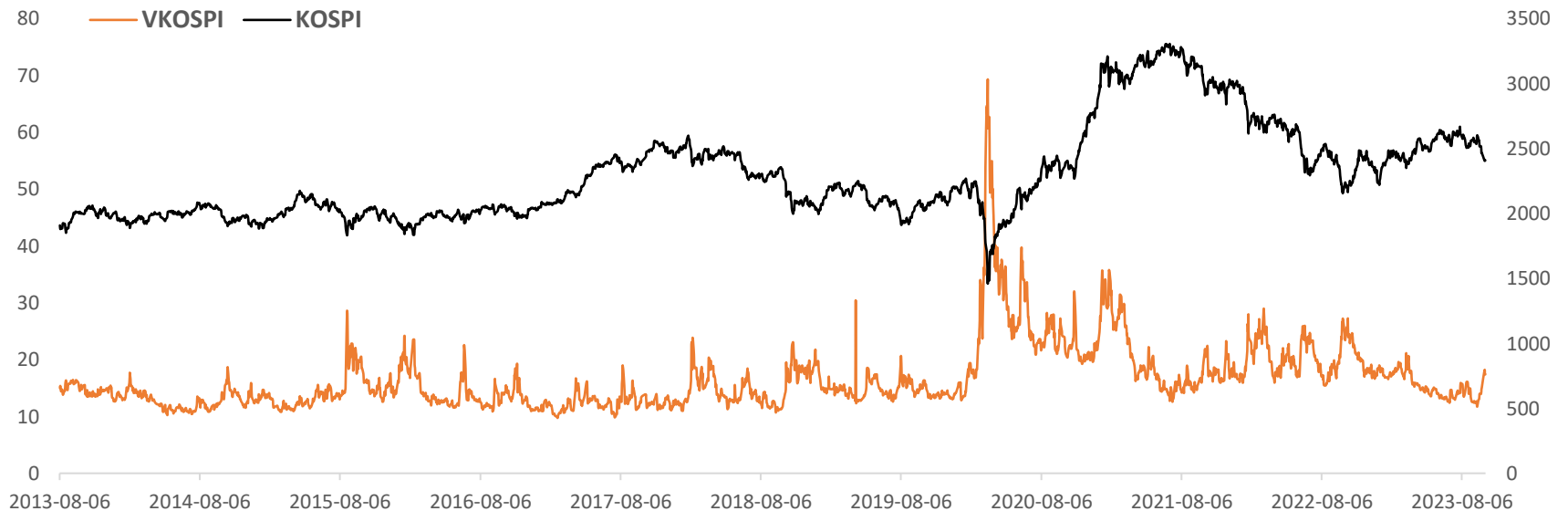
3. 주식시장 자금 흐름 변수, Multi Layer LSTM 활용

- : 경제 내 자금 흐름은 시장 심리를 나타내는 지표
- : 시장 심리 → 변동성에 대한 기대

II. Variables – VKOSPI

- **KOSPI 200 옵션가격 이용**
: 옵션 투자자들이 예상하는 KOSPI 200지수의 미래 변동성 측정
- **Derivation : 옵션가격 결정모형에서 역산**
: Option price = $f(\text{행사가격, 현재 상품가격, 남은 만기, 이자율, 변동성})$

VKOSPI and KOSPI, 2013.08.16 ~ 2023.10.10, Daily



II. Variables – Independent Var.

주가 수익률 및 변동성 관련 선행 논문의 변수 사용 목록

1) Han et al. (2015)

: VKOSPI, KOSPI 200 Spot index, VIX, S&P 500, Exchange Returns, interest rates

2) Chun et al. (2020)

: Exchange Rates, Commodity spot – Gold, Oil Price

3) Kim and Han (2022)

: 채권 금리 Spread 등

4) 이승희와 한희준 (2016)

: 경기동행, 선행지수, VKOSPI, VIX, 주택매매가격지수, 콜금리, CD금리, 국고채금리, 장단기금리차

5) 유한수(2006), 류형선 외(2013)

: 개인, 외국인, 기관투자자 순매수

II. Variables – Independent Var.

기존 연구들에서 활용된 변수 및 자금 흐름 변수를 추가 활용, 3개 그룹 총 51개 변수 사용
* VKOSPI 포함시 52개, 신규변수 15개(파란색), 기존 변수 36개

주식시장 매매동향(11)

- KOSPI Index
- KOSDAQ Index
- 개인
- 외국인
- 기관
- 차익거래(매수)
- 차익거래(매도)
- 차익거래(순매수)
- 비차익거래(매수)
- 비차익거래(매도)
- 비차익거래(순매수)

주식시장 자금흐름 (9)

- 신용잔고
- 주식형펀드
- 신용비율
- 고객예탁금
- 장내파생상품 예수금
- 대고객 RP 매도잔고
- 위탁매매 미수금
- 반대매매 금액
- 미수금 대비 반대매매 비중

주식시장 외부변수 (31)

- 콜금리_1일
- KORIBOR_3M, 6M, 12M
- CD91
- 국고채 1Y, 3Y, 5Y, 10Y, 20Y, 30Y
- 통화안정증권 91D, 1Y, 2Y
- 회사채 AA, AA-, BBB
- MMF 7D
- KOSPI 200 Futures
- US T.bill 3M, 5Y, 10Y, 30Y
- Gold Price
- Dollar Index, KRW/USD
- IXIC, DJI, S&P 500 Index
- VIX
- WTI 현물가격(선물X)

분석기간 : 2013.08.06 ~ 2023.10.10

분석단위 : Weekly data, 531 Weeks

실제활용 : (52개+Target Factor 4) * lag 4

II. Variables - Transformation

한희준(2023)등을 참고하였으며, FRED-MD Working Paper 기준 T-Code 부여

주식시장 매매동향

Variables	T
• KOSPI Index	5
• KOSDAQ Index	5
• 개인	1
• 외국인	1
• 기관	1
• 차익거래	1
• 비차익거래	1

주식시장 자금흐름

Variables	T
• 신용잔고	5
• 주식형펀드	5
• 신용비율	2
• 고객예탁금	5
• 장내파생상품 예수금	5
• 대고객 RP 매도잔고	5
• 위탁매매 미수금	1
• 반대매매 금액	5
• 미수금대비반대매매 비중	1

주식시장 외부변수

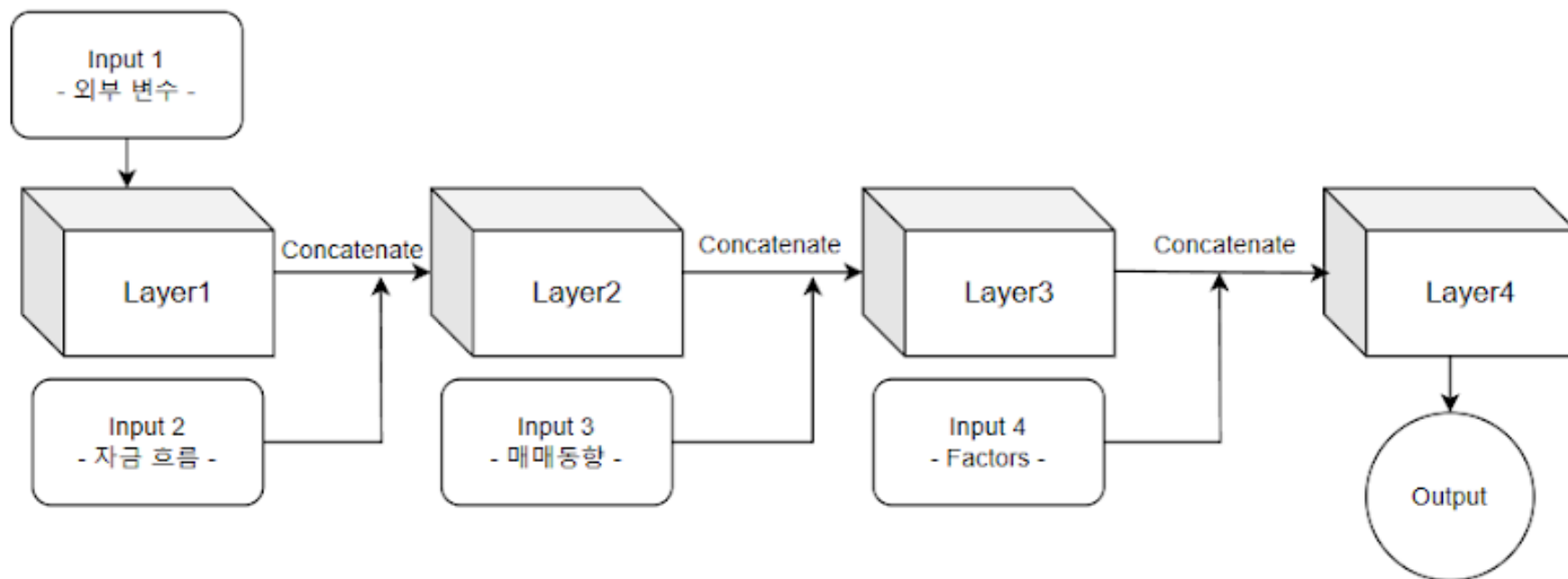
Variables	T
• 콜금리_1일	2
• KORIBOR	2
• CD91	2
• 국고채	2
• 통화안정증권	2
• 회사채	2
• MMF	2
• KOSPI 200 Futures	5
• US T.bill	2
• Gold Price	5
• D.Index, KRW/USD	5
• IXIC, DJI, S&P	5
• VIX	1
• WTI 현물	5

T-Code :

(1) No Transformation (2) ΔX_t (5) $\Delta \log X_t$

III. Model : Multi Layer LSTM

< Multi Layer LSTM의 구조 >



III. Model : Multi Layer LSTM

< Multi-Layer LSTM Code >

```
input_layer_1 <- layer_input(shape = c(4, 31))
input_layer_2 <- layer_input(shape = c(4, 9))
input_layer_3 <- layer_input(shape = c(4, 12))
input_layer_4 <- layer_input(shape = c(4, 4))

# Level-1 LSTM layers
lstm_layer_1_1 <- layer_lstm(units = unit_n, return_sequences = TRUE)(input_layer_1)
lstm_layer_1_2 <- layer_lstm(units = unit_n, return_sequences = TRUE)(input_layer_2)

# Concatenation after Level-1
concatenated_layer_1 <- layer_concatenate(c(lstm_layer_1_1, lstm_layer_1_2))

# Level-2 LSTM layer
lstm_layer_2_1 <- layer_lstm(units = unit_n, return_sequences = TRUE)(input_layer_3)
lstm_layer_2 <- layer_lstm(units = unit_n, return_sequences = TRUE)(concatenated_layer_1)

# Concatenation for Level-2 and Level-3
concatenated_layer_2 <- layer_concatenate(c(lstm_layer_2, lstm_layer_2_1))

# Level-3 LSTM layer
lstm_layer_3_2 <- layer_lstm(units = unit_n, return_sequences = TRUE)(input_layer_4)

# Concatenation for Level-3
concatenated_layer_3 <- layer_concatenate(c(concatenated_layer_2, lstm_layer_3_2))

# Level-4 LSTM layer
lstm_layer_4 <- layer_lstm(units = unit_n)(concatenated_layer_3)

# Output layer
output_layer <- layer_dense(units = 1)(lstm_layer_4)

# Model
model <- keras_model(inputs = list(input_layer_1, input_layer_2, input_layer_3, input_layer_4),
                      outputs = output_layer)

# Compile the model
model %>% compile(loss = 'mse', optimizer = 'adam')

# Fit the model to the training data
model %>% summary()

history = model %>% fit(x = list(X.arr3, X.arr2, X.arr1, X.arr4), y, epochs = epochs,
                       batch_size = batch_size, shuffle = FALSE, verbose = 2)
```

IV. Result - Boruta Selection

Boruta Algorithm을 통한 변수선택 결과, 신규 도입 변수들이 상위권에 위치

< Boruta Selection Result >

	1 Week Ahead	2 Weeks Ahead	4 Weeks Ahead	8 Weeks Ahead
1 st	VKOSPI	VKOSPI	VKOSPI	차익거래(매수)
2 nd	VIX	VIX	위탁매매 미수금	위탁매매 미수금
3 rd	위탁매매 미수금	위탁매매 미수금	VIX	VIX
4 th	차익거래(매도)	차익거래(매도)	차익거래(매수)	VKOSPI
5 th	차익거래(매수)	비차익거래(매도)	차익거래(매도)	차익거래(매도)
6 th	비차익거래(매도)	차익거래(매수)	비차익거래(매도)	비차익거래(매도)
7 th	S&P 500	미수금 대비 반대매매 비중	비차익거래(매수)	미수금 대비 반대매매 비중
8 th	미수금 대비 반대매매 비중	비차익거래(매수)	미수금 대비 반대매매 비중	비차익거래(매수)
Number of Var.	47	43	51	62

* Kim and Han(2022), 보루타 알고리즘과 교차검증 그리고 랜덤포레스트를 통한 변수선택

IV. Result - RMSE, MAE

* MAE 기준 Increasing order

Model	1-week ahead	
	MAE	RATIO
HAR	0.073	1.00
Adaptive LASSO	0.073	1.00
LASSO	0.075	1.02
Random Walk	0.076	1.03
Adaptive E.NET	0.077	1.05
Target Factor	0.078	1.07
Boruta - RF	0.079	1.08
E.NET	0.079	1.08
Boruta - XGB	0.079	1.08
Random Forest	0.082	1.12
XGBoost	0.083	1.14
Multi Layer LSTM	0.092	1.26
HARX	0.101	1.38
LSTM	0.111	1.52

Model	4-week ahead	
	MAE	RATIO
Boruta - XGB	0.109	1.00
Boruta -RF	0.115	1.05
XGBoost	0.120	1.10
Random Forest	0.121	1.11
HAR	0.127	1.15
Adaptive LASSO	0.128	1.18
LASSO	0.131	1.21
E.NET	0.132	1.21
Adaptive E.NET	0.133	1.22
Multi Layer LSTM	0.139	1.28
Random Walk	0.141	1.29
Target Factor	0.143	1.31
LSTM	0.156	1.43
HARX	0.168	1.55

Model	2-week ahead	
	MAE	RATIO
HAR	0.097	1.00
Boruta - RF	0.097	1.00
Boruta - XGB	0.101	1.05
Random Forest	0.103	1.06
XGBoost	0.103	1.07
Random Walk	0.105	1.08
LASSO	0.111	1.14
E.NET	0.112	1.15
Target Factor	0.112	1.15
Adaptive E.NET	0.114	1.18
Adaptive LASSO	0.115	1.18
Multi Layer LSTM	0.117	1.21
LSTM	0.122	1.26
HARX	0.148	1.53

Model	8-week ahead	
	MAE	RATIO
Boruta -RF	0.121	1.00
Multi Layer LSTM	0.124	1.03
Boruta -XGB	0.127	1.05
Random Forest	0.131	1.09
LSTM	0.132	1.09
XGBoost	0.134	1.11
Adaptive LASSO	0.134	1.11
Adaptive E.NET	0.135	1.12
HAR	0.140	1.13
LASSO	0.141	1.17
E.NET	0.142	1.18
Target Factor	0.158	1.31
Random Walk	0.162	1.35
HARX	0.187	1.55

IV. Result - GW Test

GW Test - Forecast Horizon 우수모형 대비

: 8 Weeks Ahead Forecast에서, Random Forest with Boruta Selection과 예측력 차이 없음

Forecast Horizon	비교모형	Statistics	P-Value
1 Week ahead	HAR	-3.961	0.000
2 Weeks ahead	HAR	-3.320	0.001
4 Weeks ahead	XGB-Boruta	-2.100	0.035
8 Weeks ahead	RF-Boruta	-0.362	0.717

* 각 Week ahead forecast별 최소 MAE모형과 GW Test 실시

GW Test - LSTM 대비

: 4 Weeks ahead 이상의 Forecast horizon에서 기존 LSTM대비 유의한 예측력 차이

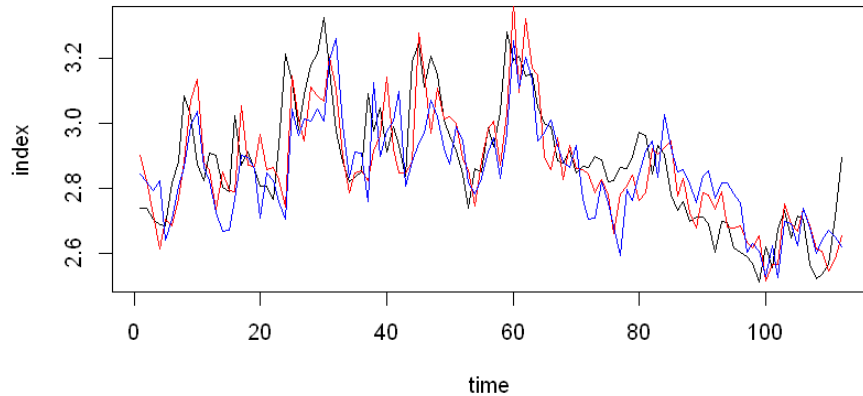
Forecast Horizon	Statistics	P-Value
1 Week ahead	1.254	0.210
2 Weeks ahead	0.066	0.947
4 Weeks ahead	2.361	0.018
8 Weeks ahead	2.857	0.004

* 일반 LSTM과 GW Test 실시

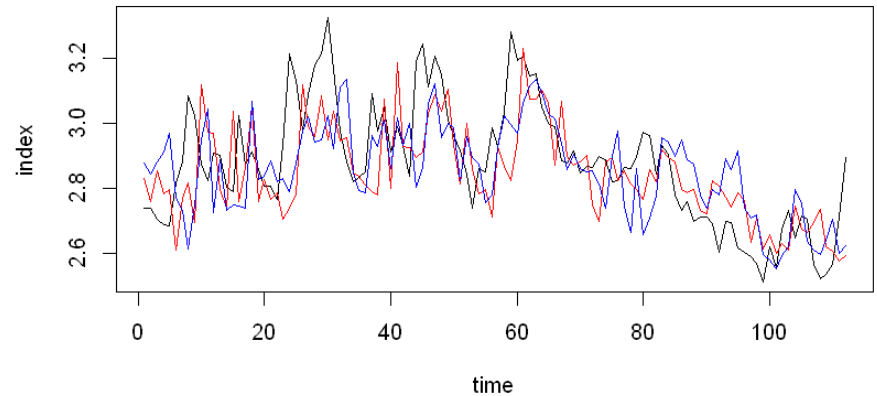
IV. Result – Plot

LSTM과 Multi-Layer LSTM의 비교

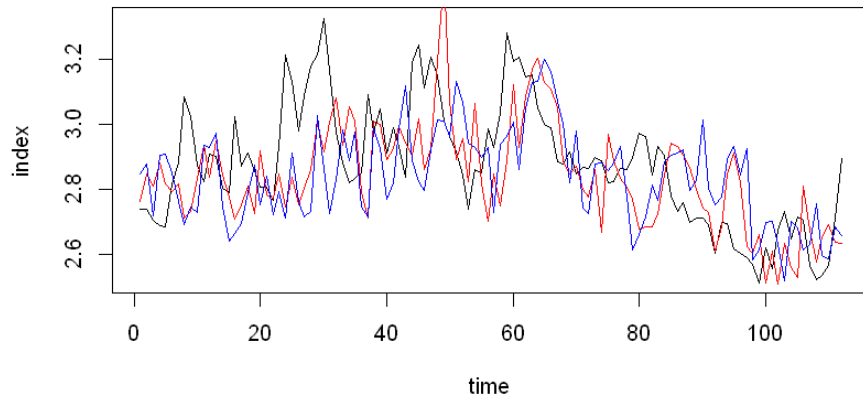
1 Week Ahead Forecast



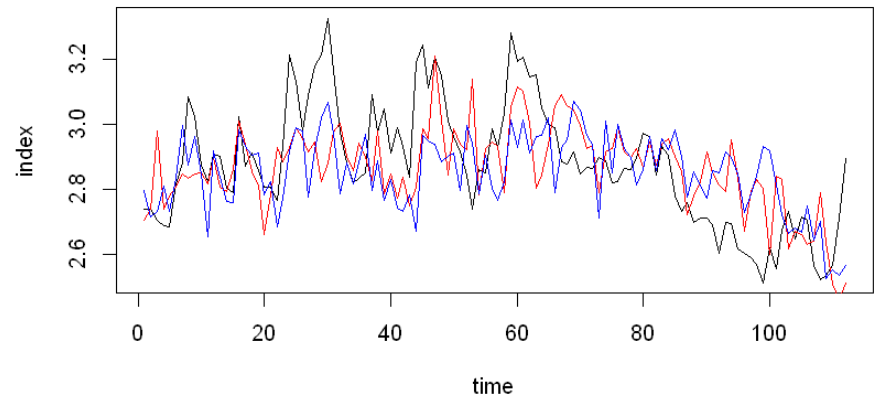
2 Weeks Ahead Forecast



4 Weeks Ahead Forecast



8 Weeks Ahead Forecast



— Real Data — LSTM — Multi Layer LSTM

IV. Result - MCS Test

< Superior Set of Models >

모형 / 알고리즘	1 Week Ahead	2 Weeks Ahead	4 Weeks Ahead	8 Weeks Ahead
Multi Layer LSTM	-	-	-	○
LSTM	-	-	-	○
Boruta - XGBoost	○	○	○	○
Boruta - RF	○	○	○	○
HAR	○	○	-	○
HARX	-	-	-	-

V. Contributions

I. 주식시장 자금흐름 변수의 유용성

: Boruta Selection 결과, 변동성 설명변수로서의 유용성 발견

II. Multi-Layer LSTM의 중장기 예측성능

: 8 Weeks ahead forecast에서의 우수한 예측성능 → 중장기 우수한 예측성능 기대

III. 기존 LSTM 알고리즘 대비 우수한 예측력

: 4 Weeks ahead forecast 이상에서 유의하게 예측력 차이 발생 확인

VI. Related Research

1. Han, H., Kutan, A. M., & Ryu, D. (2015). Effects of the US stock market return and volatility on the VKOSPI. *Economics*, 9(1), 20150035.
2. Chun, D., Cho, H., & Ryu, D. (2020). Economic indicators and stock market volatility in an emerging economy. *Economic Systems*, 44(2), 100788.
3. Kim, B. Y., & Han, H. (2022). Multi-Step-Ahead Forecasting of the CBOE Volatility Index in a Data-Rich Environment: Application of Random Forest with Boruta Algorithm. *Korean Economic Review*, 38(3), 541-569.
4. 이승희, & 한희준. (2016). 경제/금융 변수를 이용한 한국 주식시장의 변동성분석 및 예측. *경제학연구*, 64(2), 67-95.
5. Liu, M., Ye, J., & Yu, L. (2022). Volatility Prediction via Hybrid LSTM Models with GARCH Type Parameters. *Proceedings of Business and Economic Studies*, 5(6), 37-46.
6. Kim, S. W. (2010). A study on developing a VKOSPI forecasting model via GARCH class models for intelligent volatility trading systems. *Journal of Intelligence and Information Systems*, 16(2), 19-32.
7. 류형선, 양성국, & 김봉현. (2013). 한국 주식시장의 투자주체별 투자행태가 주가지수 수익률에 미치는 영향. *글로벌경영학회지*, 10(3), 355-370.
8. 유한수. (2006). 투자주체별 거래와 주가변동성. *기업경영연구*, 13(2), 67-77.