

Research Article

EXPLAINING A STOCK RETURN VOLATILITY PREDICTION MODEL USING A MACHINE LEARNING APPROACH

Author: ¹Mehdi Rezaei Pour , ²Ali Saeedi, ³Jafar Jola

¹M.Sc. in Financial Management, Islamic Azad University, North Tehran Branch, Tehran, Iran

²Ph.D. in Finance, Islamic Azad University, Science and Research Branch, Tehran, Iran

³Ph.D. in Finance, Islamic Azad University, Science and Research Branch, Tehran, Iran

KEYWORDS

Stock return volatility,
Machine learning, Stochastic
prediction

ARTICLE HISTORY

Received Date: 25/04/2026

Revised Date: 29/04/2026

Accepted Date: 14/05/2026

Published Date: 29/05/2026

DOI: 10.69557/ujrra.v5i2.282

CORRESPONDING AUTHOR

Mehdi Rezaei Pour

<https://orcid.org/0009-0004-8057-836X>

ABSTRACT

The present study aims to develop a machine learning model for predicting stock return volatility. This research is applied in purpose and quantitative-analytical in method. The statistical population includes all companies listed on the Tehran Stock Exchange and Securities Market. The required variables in this study consisted of end-of-day stock price data (opening price, highest price, lowest price, average price, and volume) collected over the required time period. Using fuzzy neural networks and Markov methods, the intended data were analyzed. For this purpose, MATLAB software was used. Libraries such as Pandas, NumPy, and Scikit-learn were used to generate noise-free data and construct the model. To develop an algorithm for predicting stock returns, machine learning models were employed. For the initial model, a linear regression was built in Python using scikit-learn to predict stock returns. The results showed that, after accounting for transaction costs, the machine learning classifiers used in this study indicate that Support Vector Regression cannot generate positive Sharpe ratios in all cases. In contrast, logistic regression and multilayer perceptron can achieve them in some cases. Trading strategies based on a designed four-day stochastic prediction can generate positive returns when total transaction costs are below 0.015–0.040 per transaction.



INTRODUCTION

Predicting stock price movements and consequently forecasting stock returns has long been a persistent puzzle for investors. The Efficient Market Hypothesis and the Random Walk Theory (Cootner, 1964; Fama, 1965) suggest that all past information is fully reflected in stock prices, and therefore stock prices do not change due to past information; in such a situation, it is not possible to earn abnormal excess returns using historical information (weak form informational efficiency). In his article, Fama stated that, under the weak-form of market efficiency, stock prices follow a random walk. However, some researchers argue that the Efficient Market Hypothesis cannot explain certain abnormal behaviors observed in many stocks, such as sharp short-term stock price movements (Weng, Ahmed, & Magahed, 2017).

They claim that existing stock exchanges, especially in developing countries, are not efficient financial markets; therefore, investors may use information not reflected in the market to predict stock prices and obtain abnormal excess returns. In addition, many behavioral finance theories suggest that investor sentiment can predict stock returns (Loewenstein, 2003; Jiang & Zhou, 2018). For example, on August 30, 2018, media outlets accused Qiangdong Liu, CEO of the Chinese e-commerce giant JD.com, of a non-financial allegation. Although the court later announced that there were no charges against Liu, this incident had significant negative effects in its early stages. As a result, JD's stock price fell by 5.97% on the first trading day after the allegation and by 16% on the second day. The company's market value lost more than 7.2 billion USD in U.S. capital markets. This not only challenged the validity of the Efficient Market Hypothesis but also demonstrated that online news and investor sentiment can strongly influence the stock market.

This is a small but critical fact in financial markets: news and information about investor behavior and signals can significantly affect stock prices and potentially be used to predict price movements. Previous research on stock price prediction has mainly focused on traditional indicators (Cao, Li, & Li, 2019; Chung & Shin, 2019; Zhang, Li, & Shang, 2017), such as historical transactions and technical indicators, and has mostly addressed issues from a quantitative perspective. In contrast, new methods, such as data-driven qualitative theories, have also emerged and can effectively predict the impact of news and financial information on stock price changes. With the development of the information and communication space, the academic landscape of stock price prediction has been partly reshaped.

Information and communication platforms provide spaces for investors to exchange ideas and information. News and information create numerous research opportunities for stock price forecasters, including how to transform qualitative information—such as news and posts—into quantitative data. Textual information can, to some extent, predict stock prices; however, when converted into quantitative data, it can become more useful to investors (Chen & Chen, 2017; Feuerriegel & Gordon, 2018).

Considering the macroeconomic structures of each country and the markets within each economy, it is clear that one of the most important financial sources supplying other sectors of the economy is the capital markets. The stock exchange and securities are components of the capital market and, as part of the economy (the financial sector of the economy), are subject to the country's macroeconomic variables. Accordingly, the present study examines and explains

a machine learning-based model for predicting stock return volatility. Based on the searches conducted, no research with this breadth and approach has been conducted.

Stock Return Volatility

Stock return volatility is an important concept in the field of finance. This concept refers to the amount of share price changes in a specific period of time. Volatility of a stock indicates how much the stock price changes over a given period of time; in other words, it measures the degree of volatility of the stock price. Various methods are used to calculate the volatility of stock returns. One common method for calculating volatility of stock returns is to use the standard deviation. Standard deviation indicates the dispersion of data; here, it refers to the degree of dispersion of stock returns over a given period. The larger the standard deviation, the greater the volatility of stock returns and vice versa.

Chandwani and Saluja (2014) used technical indicators, fundamental indicators, and their combination as inputs to predict stock direction for companies in the Indian stock market. Artificial neural network models, support vector regression optimized through genetic algorithms, and simple (non-optimized) versions of artificial neural networks and support vector regression were used to generate the forecasts. The results showed that the optimized artificial neural network model, trained with a genetic algorithm and a composite indicator set, achieved the highest accuracy (80.51%), followed by the simple support vector regression model with technical indicators, which achieved an accuracy of 79.4%.

Stock price forecasting is a complex and challenging task for companies, investors, and stock traders alike. Stock markets are intrinsically noisy, nonparametric, nonlinear, and deterministic environments (Bettman, 2009). These create challenges for effective and efficient forecasting of future stock prices. Feature selection from financial data is another difficult task in stock forecasting for which many approaches have been proposed (Kumar, 2021). There is a trend in which some researchers use only technical indicators, while others use historical data (Davies et al., 2007; Kumari et al., 2017; Wu and Lee, 2017; Vanstone and Finnie, 2017; Zhong and Enke, 2017; Li Chong, 2017).

Machine Learning Approach

Machine learning is an important field in computer science and artificial intelligence that enables machines and computer systems to learn without being explicitly programmed. In machine learning, machines can recognize patterns and rules in data and, using these patterns and rules, make predictions and intelligent decisions.

With advances in computing power over the past few years, the use of computing in business decision-making has increased proportionally (Cavalcante et al., 2016; Yang et al., 2016). Although machine learning techniques have been “widely accepted in the study and evaluation of stock market behavior” (Atsalakis and Valavanis, 2009), there is no set of indicators or methodology that can be consistently used to forecast stock prices (Cavalcante et al., 2016). Cavalcante et al. (2016) provide a financial trading framework (Figure below) along with a forecast, which is a customized version of the systems approach to forecasting by Zhang et al. (2020). It illustrates the steps in forecasting financial time series with machine learning methods.

Cavalcante et al. (2016) conducted a recent study on the use of machine learning methods in

financial markets. They reviewed research papers from 2009 to 2015 and found that “technical analysis is the most widely used in the study of financial markets in the reviewed papers.” Of the 56 papers reviewed, 47 papers used technical indicators, and 9 papers used fundamental indicators. In addition, other older surveys show a very similar trend. Atsalakis and Valavanis (2009) noted researchers' preference for using technical indicators in surveys, focusing on the use of artificial neural networks and neuro-fuzzy models in stock market forecasting, while the majority of the reviewed papers used index technology as input.

Similarly, Zwart (2009) conducted a review of 46 research papers on machine learning models for stock index forecasting, focusing on the technologies used, forecasting time horizons, input variables, and evaluation methods.

Research Background

Kumar et al. (2021), a study titled Stock Market Forecasting Using Computational Intelligence: A Survey Study... In this paper, the selected papers are organized and discussed based on six main perspectives: (1) stock market analysis and related datasets, (2) The type of input variables investigated, (3) the preprocessing techniques used, (4) the feature selection techniques for selecting the effective variables, (5) the forecasting models to tackle the stock price forecasting problem, and (6) the performance metrics used to evaluate the models. The major contribution of this work is to provide a systematic approach for financial researchers and analysts to develop an intelligent methodology for stock market forecasting. This paper also provides the outlines of the proposed work aimed at enhancing the performance of the existing techniques.

Omer et al. (2020) conducted a study titled Financial Time Series Forecasting with Deep Learning: A Systematic Literature Review: 2005-2019. The motivation for this paper is to provide a comprehensive literature review of deep learning studies on the implementation of financial time series forecasting. The researchers not only categorized the studies by their target forecasting implementation domains, such as index, forex, and commodity forecasting, but also grouped them by deep learning models, including convolutional neural networks, deep belief networks, and short- and long-term memory. The researchers also sought to envision the field's future by highlighting its potential obstacles and opportunities for interested researchers.

Wafi et al. (2016) conducted a study titled Forecasting Stock Price Movement with Multiple Data Sources: Evidence from the Stock Market in China. In this study, 14 data source combinations were used across 5 heterogeneous data sources and 3 forecast horizons (1 day, 2 days, and 3 days) to investigate the predictive effects of stock price movements in China. The study found that the optimal data source combination for active and passive stocks differs. Active stocks achieve the highest accuracy when combining multiple non-traditional data sources, while passive stocks achieve the highest accuracy when combining traditional and unconventional data sources. Also, across most data source combinations, the more active the stock, the more accurate the results, indicating that the present research approach is more powerful for predicting stock price movements during active and highly active periods.

Fan et al. (2015) used the ARMA-GARCH-NN method and daily stock market patterns to predict stock market shocks. The results and initial evidence showed that the ARMA-GARCH-NN models are generalizable and that daily stock market patterns can be an efficient method for detecting shocks in the stock market.

Chen et al. (2016) predicted stock price volatility based on volume information in the stock

and foreign exchange markets. The results of this study indicate that information related to the volume of the stock market affects stock returns and the volume-price relationship, and this effect is indirectly created through net capital on the stock market. Also, currency depreciation and current volatility do not lead to a recovery in the stock market.

Zhang et al. (2014) examined the periodicity of important dates of the Chinese market price bubble. In this study, three periods of price bubbles are identified for the time periods July 2006 to October 2007, December 2007 to October 2008, and October 2014 to June 2015, based on the CSI 300 index. In this study, the periodic law of price technique was used to examine price fluctuations. The results showed that this technique identified the crash period in 2007 and 2015 well. Price gaps can serve as a warning of a price bubble.

Kumari et al. (2017) studied the factors affecting unsystematic volatility in the Indian stock market. In this study, they examined five factors: company size, book-to-market ratio, momentum, liquidity, cash flow-to-price, and return on assets. The results of the study showed that smaller companies have higher liquidity, lower momentum, higher book-to-market ratios, lower cash flow-to-price ratios, and higher unsystematic volatility.

Bao et al. (2017) examined volatility risk and the predictability of stock returns during the Global Financial Crisis. This study was conducted over the period 2001–2010 in the United States stock exchange and the Chicago Stock Exchange. The results of the study indicated that idiosyncratic volatility enables the prediction of stock returns in both bear and bull markets. Considering cross-sectional effects, these volatilities have both positive and negative impacts on stock return prediction.

Method

The present research is applied in terms of objectives and quantitative-analytical in its analysis method. The statistical population of the present research comprises all companies listed on the Tehran Stock Exchange; however, for ease of data access and to avoid high costs, the statistical sample consists of several categories of listed companies. The conditions for selecting the sample include that the company under study does not experience any changes in the financial period during the mentioned period. The company's data is available daily from 2013 to 2023. The company does not have a trading break. Its financial information is available.

Research Variables

Required variables. In the present study, end-of-day stock price data (opening price, maximum price, minimum price, average price, and volume) were collected for the required period.

Independent variables

Independent variables, also known as predictors or characteristics, are factors that are manipulated or controlled in a study. In predicting stock return fluctuations, independent variables include historical stock prices, trading volume, market indices, economic indicators, and company-specific financial metrics. These variables serve as inputs to the machine learning model and are used to predict fluctuations in stock returns.

Dependent variable

The dependent variable is stock return fluctuations. This is the variable that the machine learning algorithm is trying to predict based on the independent variables.

Mediating variables

In predicting stock return volatility, mediating variables include market sentiment indicators, volatility clustering indicators, and sector-specific performance measures. These variables may not directly affect stock return volatility, but they can affect the relationship between the independent and dependent variables.

Control variables

In predicting stock return fluctuations, control variables include interest rates, inflation rates, political events, and seasonal effects. Using fuzzy and Markov neural network methods, the desired data will be analyzed. For this, Matlab software was used. Libraries such as Pandas, NumPy, and Scikit-learn were used to create noise-free data and build the model. Machine learning models were used to develop an algorithm for predicting stock returns. To create the initial model, a linear regression model is created using Python and scikit-learn to predict stock returns:

Table 1: Linear regression model for predicting stock returns

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn import metrics

#Load historical stock data
stock_data = pd.read_csv('stock_data.csv')

#Define features and target variable
X = stock_data[['Feature1', 'Feature2', 'Feature3']]
y = stock_data['Return']

#Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)

#Create a Linear Regression model
model = LinearRegression()
model.fit(X_train, y_train)

#Make predictions
predictions = model.predict(X_test)
#Evaluate the model
print('Mean Absolute Error:', metrics.mean_absolute_error(y_test, predictions))
```

Research Model

Figure 1 shows the different layers that make up the proposed framework:

Input
Market Status
Model and output.

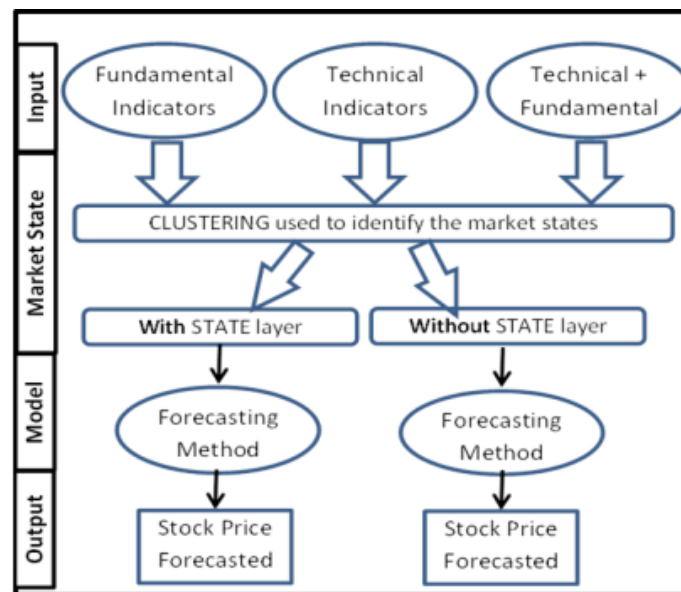


Figure 1: Different layers that make up the proposed framework

The input layer determines the type of financial indicators for the model layer. These indicators are classified into technical indicators (based on technical analysis), fundamental indicators (financial variables based on fundamental analysis), and their coherent combination. This layer helped determine the input set related to the target stock.

The market-state layer was used to identify different stock market conditions and to determine the target stock's sensitivity to market fluctuations. While incorporating market conditions into the forecasting process, a clustering algorithm is used to identify distinct regions of the general stock market, and corresponding forecasting models are built for the target company for each market condition. In addition, the framework also includes stocks that may not be affected by various market fluctuations. The model layer includes the machine learning method(s) (e.g., ANN, SVM, etc.) used for output forecasting, and the output layer represents the generated forecast.

The framework is trained on 147 companies to predict the percentage change in the stock price of a selected company over the next year (252 trading days), using various machine learning methods that are exposed to technical indicators, fundamental indicators, and their combinations across the member companies of the Tehran Stock Exchange. The performance of the framework's predictor, as measured by Root Mean Square Error (RMSE), is compared with that of the random walk and ANN models using technical indicators. The following sections provide details on the implementation of the framework (Figure 2).

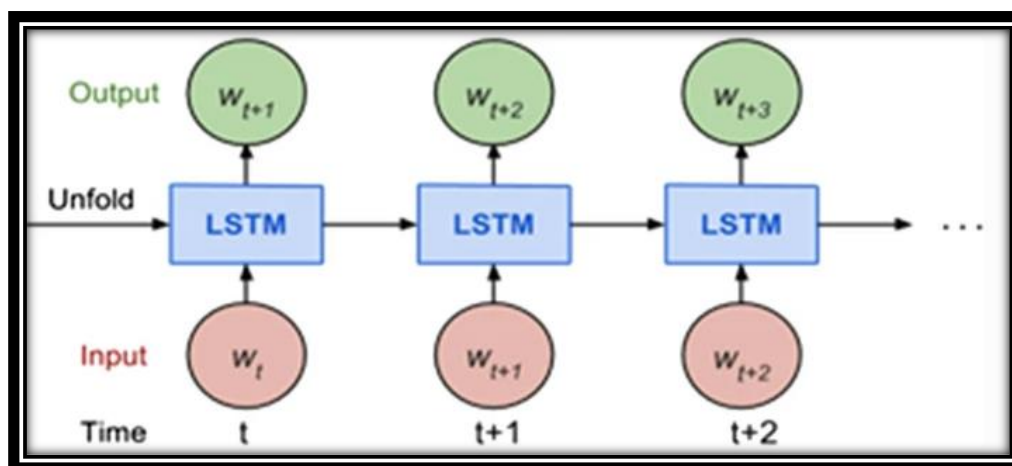


Figure 2: Framework implementation details

For each company in the present study, end-of-day stock price data (opening price, maximum price, minimum price, average price, and volume) were retrieved for the required period. As described by Romero and Balch (2014), a 2-price-adjusted set was preferred over the raw data from the retrieved financial time series (Jiang and Zhou, 2018). The functions in the TTR software package (Ulrich, 2015) were used to calculate the required technical indicators.

For each company, the daily values of the return ratio and the short- and long-term EPS growth rates were calculated based on IBES (2016) estimates. As an overview, the steps taken are as follows, which are described in more detail below:

- Estimate the average EPS (1%, 2%, and long-term growth)
- Convert the monthly frequency to daily
- Estimate the 1-year EPS to the dividend yield ratio
- Calculate a proxy for the short-term EPS growth rate (2-year EPS / 1-year EPS)

In the present study, analytical methods will be used, including library methods and quantitative analyses of time series data. Also, seven stocks of interest in the Securities and Exchange Commission will be used. Detailed information about these stocks, including codes, names, and industry sectors, will be provided. Investors usually choose their desired stocks based on their investment criteria, so in the present study, the stocks that have attracted the most investor attention are identified through interviews. Seven stocks belonging to different industries are randomly selected as research targets, which not only attract the current high attention of investors but also diversify the investor's portfolio, so our sample is a good representation of the Iranian stock market. In the present study, 12 months from April 1, 2020, to March 29, 2021, are studied. The forecast targets will be set as $(t + 2)$ -Open (t) , Open $(t + 3)$ -Open (t) , and Open $(t + 4)$ -Open (t) , indicating forecast horizons of 1 day, 2 days, and 3 days, respectively.

To smooth the data, the TTR section (Ulrich, 2015) was used to transform the raw price data into moving-average and moving-average convergence/divergence (MACD) indices for the short term (26 and 12 days) and the medium term (126 and 12 days).

FINDINGS

Market Trend Detection

Chi (2014) described market indicators for tracking market movement, and from these indicators, the VIX index, the Relative Strength Index, and the Put-to-Call ratio were selected as variables. When variables were selected to represent states, the potential market states were further defined by the number of states the market could represent, as indicated by the selected cluster sizes (3, 5, and 7). To investigate which of these market-feature indicators is more effective at capturing the overall stock market state, the forecasting models were tested on test data using different state-layer definitions (3 indicators with 3 cluster sizes). Each company, machine learning method (artificial neural network, support vector regression, decision tree, learning rate), and input set (technical, fundamental, hybrid) was combined based on the table showing root mean square error values for each market feature indicator across different business sectors. The averages were calculated from the model's performance for each company, using a specific market characteristic index.

The results showed that the overall VIX and Put-Call ratios were similar (although the VIX was slightly better at 0.0695 vs. 0.0711), and the Relative Strength Index did not perform as well as the other two indices in market trend identification. An examination at the industry sector level shows a similar trend: except for the IT sector, the average root mean square error for models using the market characteristic index or the VIX index was slightly lower than for models using the Put-Call ratio.

The root mean square error values indicate the best performance of the market characteristic index across industry sectors. Both the Put-Call ratio and the VIX index outperformed the Relative Strength Index across all industry sectors. In addition, the VIX performed best in 61% of cases, followed by the Put-Call Ratio in 39% of cases, and the Relative Strength Index was not included in any category. Some sectors, such as Energy, Financials, Industrials, and Materials, seem to favor the Relative Strength Index more than others. One explanation for the Relative Strength Index's poor performance relative to others is that the VIX and Put-Call Ratio are forward-looking indicators, whereas the Relative Strength Index is backward-looking.

Model Validation

Even with a relatively simple neural network, an MLP, we can see the efficiency and potential of deep learning technology when predicting the price changes of HT and OPTX stocks. As shown by the blue line in Figure 3, the multilayer perceptron network successfully predicted stock prices at several time points. For example, during the 110-120 time period, the MLP predicted a relative price stagnation followed by a sharp increase in the 100 time period. The prediction success is visible throughout the entire period. However, during testing, at several time points, the network failed to predict price movements accurately and exhibited a time lag (e.g., at time points 60, 160, and 200). These deviations can be partially corrected by further tuning the M model. LP (weight correction) or with additional training and optimization of the multilayer perceptron neural network.

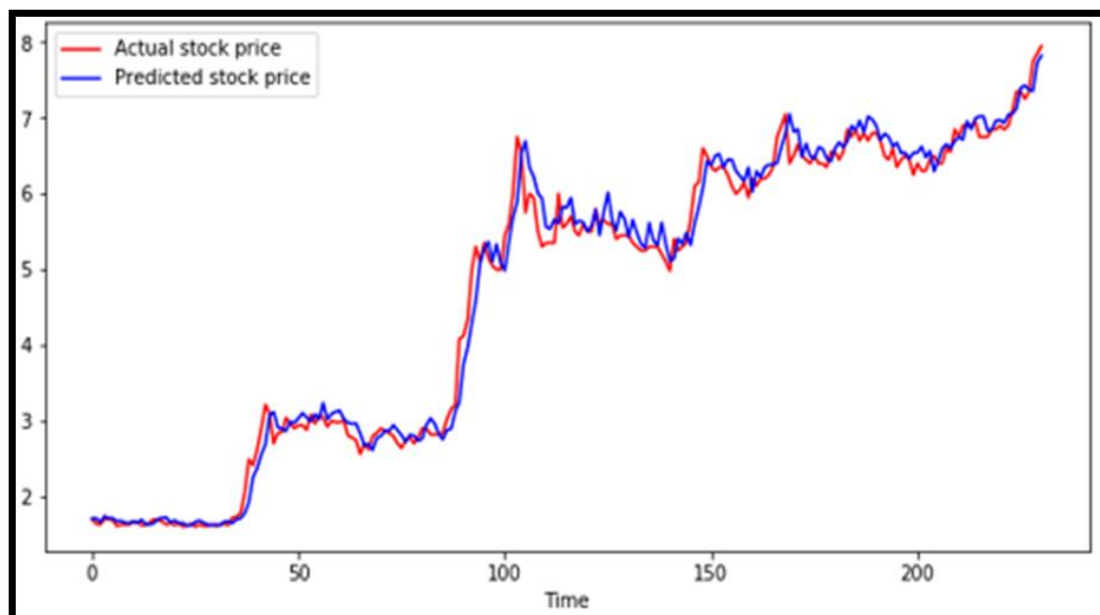


Figure 3: Simulation of the total index over one year for the time horizon of August 2018 to August 2019

Stock Price Prediction with Long Short-Term Memory Network

The Long Short-Term Memory Network has proven very effective with a simple model built with the Keras package in Python. Figure 4 shows the actual price movement from 1392 to 1402 (red line) and the prices predicted by the Long Short-Term Memory Network on the test data (blue line). From the chart, it can be seen that the network has noticed the movement of the OPTE stock price and predicted it. Of course, this model needs to be further tested on future stock price data to determine its real efficiency and usefulness in stock market analysis and investment.

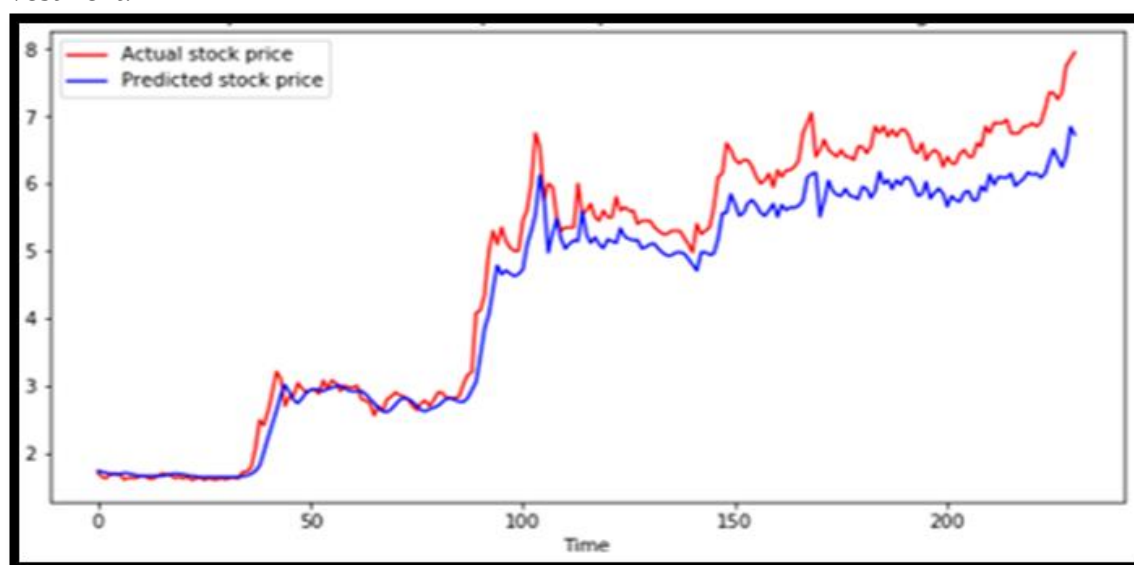


Figure 4: Stock price prediction with long short-term memory network

Based on the analysis performed, the models designed in the field of stock price time series prediction are summarized as follows (Table 2).

Table 2: Models designed for stock price time series forecasting

Method	RMSE ¹	MAE ²	MAPE ³ (%)	U ₁	U ₂
ABC-ANFIS-SVM	0.0092	0.0088	0.7283	0.0012	0.0009
ARIMA-SVM	1.0021	0.0142	2.3769	0.0489	0.0324
ABC-SVM	0.8864	0.0098	1.8237	0.0272	0.0256
ABC-ANFIS	0.9235	0.0095	1.2920	0.0195	0.0144
ANFIS-SVM	1.8934	0.0117	0.9436	0.0121	0.0132
ARIMA-ANFIS	1.1473	0.0182	2.9124	0.1025	0.0763
WANFIS	0.0378	0.0136	1.4515	0.0196	0.0056
SVM	0.0276	0.0129	0.9307	0.0021	0.0049
QGGABC-FFNN	0.0635	0.0155	1.1250	0.1036	0.0042
ANN-SVM	0.0414	0.0372	0.8403	0.1128	0.0053
HWT-ANN-DABC	0.0526	0.0436	1.4937	0.0255	0.0071
FA-MSVR	0.0387	0.0130	1.0533	0.1072	0.0178
PSO-SVM	0.0443	0.0227	1.1714	0.1369	0.0144
GA-ANFIS	0.0725	0.0316	0.9325	0.0141	0.0058
ANFIS	0.0847	0.0407	1.2746	0.0095	0.0082
ANN-SVM	0.0481	0.0521	1.0255	0.0122	0.0039
NN-SVR	0.0532	0.0368	1.0367	0.0086	0.0052
SOFM-SVR	0.0654	0.0445	0.9366	0.0074	0.0095
KPCA-SVR	0.0793	0.0504	0.8941	0.0098	0.0077
ARIMA-ANN-SVM	0.0449	0.0172	0.8452	0.0059	0.0086
SVM-RF	0.0172	0.0245	1.3417	0.0132	0.0047

CONCLUSION

This study aims to develop a machine learning-based model for predicting stock return volatility. The performance of the machine learning classifiers used in this study shows that, after transaction costs, support vector regression cannot predict ratios that produce positive Sharpe ratios in all cases; logistic regression and multilayer perceptron can obtain positive Sharpe ratios in some cases. These results may be due to the shortcomings of each machine learning model. Different classifiers may be suitable for different datasets. Support vector regression does not perform well when the dataset is noisier, meaning that if the data is not properly separated and the target classes overlap, it performs poorly. In addition, decision trees are also affected by noise.

¹ - Root Mean Square Error

² - Mean Absolute Error

³ - Mean Absolute Percentage Error

Although this paper makes some observations about multilayer perceptron and long-short-term memory networks, I did not discuss them in detail because they cannot predict intraday price movements with high accuracy. Both of the aforementioned models involve a large number of parameters, and it is difficult to achieve robust accuracy across all possible parameter combinations. Therefore, slippage costs are considered in strategies to approximate contingency costs.

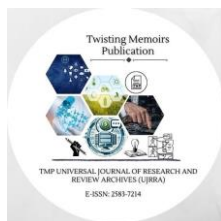
In general, trading strategies based on 4-day-designed stochastic forecasting can generate positive returns when total transaction costs are less than 0.015-0.040 per trade. The performance of machine learning-based trading strategies was investigated. After examining the robustness and comparing the models, it is shown that the most successful trading strategy in this study is based on a 4-day-designed stochastic forecasting model.

According to the results, it is proposed to present a new model based on the learning rate of the grasshopper optimization algorithm to accelerate convergence, and to present a new model based on the number of iterations, error, and sensitivity.

REFERENCES

1. Atsalakis, G. S., and Valavanis, K. P. (2009). Surveying stock market forecasting techniques- Part II: Machine learning methods. *Expert Systems with Applications*. 36, 5932-5941.
2. Bao W, Yue J, Rao Y (2017) A deep learning framework for financial time series using stacked autoencoders and long-short-term memory. *PLoS ONE* 12(7): e0180944.
3. Bettman, J. L., Sault, S. S., and Schultz, E. L. (2009). Fundamental and technical analysis: substitutes or complements? *Accounting and Finance*, 21-36.
4. Cao, J., Li, Z., & Li, J. (2019). Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 519, 127-139.
5. Cavalcante, C. R., Brasileiro, C. R., Souza, L. F. V., Nobrega, P. J., Oliveira A. L. (2016). Computational Intelligence and Financial Markets: A Survey and Future Directions. *Expert Systems with Applications*. 55, 194-211.
6. Chandwani D., Saluja S. M. (2014). Stock Direction Forecasting Techniques: An Empirical Study Combining Machine Learning System with Market Indicators in the Indian Context. *International Journal of Computer Applications*. 92(11), 8-17.
7. Chen, H., Cheng-Few, L., Wei-Kang, S. (2016). Technical, Fundamental, and Combined Information for Separating Winners from Losers. *Pacific-Basin Finance Journal*. 39, 224- 242.
8. Chen, M.-Y., & Chen, T.-H. (2017). Modeling public mood and emotion: Blog and news sentiment and socio-economic phenomena. *Future Generation Computing Systems*. <https://doi.org/10.1016/j.future>. 2017.10.028.
9. Chi Wan , , Zhijie Xiao , (2014), Idiosyncratic Volatility, Expected Windfall, and the Cross-Section of Stock Returns, in Yoosoon Chang , Thomas B. Fomby , Joon Y. Park (ed.) *Essays in Honor of Peter C. B. Phillips (Advances in Econometrics, Volume 33)* Emerald Group Publishing Limited, pp. 713–749
10. Chung, H., & Shin, K.-s. (2018). Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction. *Sustainability*, 10, 3765.

11. Cootner, P. H. (1964). The random character of stock market prices. *Operations Research*, 14, 962-965.
12. Davies, Gail & Dwyer, Claire. (2007), Qualitative Method: Are you enchanted or are you alienated, *Progress in Human Geography*, 31(2).
13. Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38, 34–105.
14. Fan, Y.H., Xi, D. and Zheng, J.M. (2015), “Stock market development, real exchange rate and China’s
15. Feuerriegel, S., & Gordon, J. (2018). Long-term stock index forecasting based on text mining of regulatory disclosures. *Decision Support Systems*, 112, 88-97.
16. Jiang, Y., & Zhou, Z. (2018). Does the time horizon of the return predictive effect of investor sentiment vary with stock characteristics? A Granger causality analysis in the frequency domain. *arXiv preprint arXiv:1803.02962*.
17. Jiang, Y., & Zhou, Z. (2018). Does the time horizon of the return predictive effect of investor sentiment vary with stock characteristics? A Granger causality analysis in the frequency domain. *arXiv preprint arXiv:1803.02962*.
18. Kumar, G., Jain, S. & Singh, U.P. (2021). Stock Market Forecasting Using Computational Intelligence: A Survey. *Arch Computat Methods Eng* 28, 1069–1101. <https://doi.org/10.1007/s11831-020-09413-5>
19. Kumari, Jyoti, Jitendra Mahakud, Gourishankar S. Hiremath . (2017).Determinants of idiosyncratic volatility: Evidence from the Indian stock market, *Research in International Business and Finance*, Volume 41, October 2017, Pages 172–184
20. Li.Chong. (2017). Log-periodic view on critical dates of the Chinese stock market bubbles, *Physica A: Statistical Mechanics and its Applications*, Volume 465, 1 January 2017, Pages 305–311
21. Loewenstein, G. J. S. L. (2003). The role of affect in decision making. *Handbook of affective science*, 619, 3.
22. Omer Berat Sezer, Mehmet Ugur Gudelek, Ahmet Murat Ozbayoglu (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019, *Applied Soft Computing*, Volume 90, 106181, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2020.106181>.
23. Ulrich, J. (2015). TTR: Technical Trading Rules. R package version 0.23-0. <http://CRAN.Rproject.org/package=TTR>
24. Vanstone, B., Finnie, G. (2009). An empirical methodology for developing stock market trading systems using artificial neural networks. *Expert Systems with Applications*. 36(3), 6668-6680.
25. Wafi, S. A., Hassan, H., & Mabrouk, A. (2015a). Fundamental Analysis Models in Financial Markets – Review Study. *Procedia Economics and Finance*. 30, 939-947.
26. Weng, B., Ahmed, A. M., Megahed, M. F. (2017). Stock market one-day-ahead movement prediction using disparate data sources. *Expert Systems with Applications*. 79, 153-163.
27. Wu, S., Lee, S. (2015). Employing local modeling in machine learning-based methods for time-series prediction, *Expert Systems with Applications*. 42, 341-354
28. Yang A., Li W., Yang X. (2019) Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines. *Knowledge-Based Systems*, 163, 159–173.



29. Zhang X., Zhang T., Young A. A., Li X. (2014) Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance Data. PLoS ONE 9(2): e88075.
30. Zhang, N., Lin, A., & Shang, P. (2017). Multidimensional k-nearest neighbor model based on EEMD for financial time series forecasting. Physica A: Statistical Mechanics and its Applications, 477, 161-173.
31. Zhong, bao. Zhou, Meng. Gao, Qing. Liu Helu Xiao. 2020. Forecasting stock price movements with multiple data sources: Evidence from the stock market in China. Physica A: Statistical Mechanics and its Applications. Volume 542, 123389. <https://doi.org/10.1016/j.physa.2019.123389>
32. Zhong, X., and Enke, D. (2017). Forecasting daily stock market return using dimensionality reduction. Expert Systems with Applications. 67, 126-139
33. Zwart G., M. T. (2009). The economic value of fundamental and technical information in emerging currency markets. Journal of International Money and Finance, 28, 581-604.