

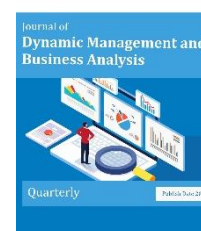


Journal Website

Article history:
Received 27 September 2024
Revised 11 November 2024
Accepted 28 November 2024
Published online 31 Dec. 2024

Dynamic Management and Business Analysis

Volume 3, Issue 5, pp 178-199



E-ISSN: 3041-8933

Analyzing Market shocks with a Hybrid ARMA-GARCH and Deep Learning Model for Enhanced Forecasting

Ebrahim Rahimi ¹ , Ahmad Mohammadi ² *, Ali Asghar Mottaghi ³ , Seyed Ali Payetakhti Oskoueï ³

¹ Ph.D. Student, Department of Accounting and Financial Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.

² Assistant Professor, Department of Accounting and Financial Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran (Corresponding author).

³ Assistant Professor, Department of Accounting and Financial Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran.

* Corresponding author email address: iacp_a-mohammadi@yahoo.com

Article Info

Article type:

Original Research

How to cite this article:

Rahimi E, Mohammadi A, Mottaghi A, Payetakhti Oskoueï A. (2024). Analyzing Market shocks with a Hybrid ARMA-GARCH and Deep Learning Model for Enhanced Forecasting. *Dynamic Management and Business Analysis*, 3(5), 178-199. <https://doi.org/10.22034/dmbaj.2024.2043125.1146>



© 2024 the author(s). Published by Knowledge Management Scientific Association. This is an open access article under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

ABSTRACT

Objective: The primary aim of the present study is to provide a model for forecasting stock market shocks using a deep learning approach in the Tehran Stock Exchange.

Methodology: The current research is applied in terms of its objective, meaning it seeks to solve a specific real-world problem, namely, the prediction of stock market shocks. In terms of nature, this research is descriptive-analytical, as its goal is to examine and analyze market fluctuations over time and it utilizes statistical methods for data analysis. For this purpose, intraday 15-minute data of the overall index in the time interval from March 20, 2018, to December 27, 2018, which includes the opening, closing, highest, and lowest values of the aforementioned index, were used. In the proposed model, statistical information from the time series is first extracted using the ARMA-GARCH model and then, using the deep learning method and one-dimensional convolutional neural networks (CNN), the nonlinear relationships of the time series observations are modeled. Finally, the appropriate models are selected, and based on the criteria of mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE), the prediction accuracies are evaluated.

Findings: The results of this study indicated that the combined ARMA-GARCH-CNN model, leveraging the strengths of both traditional and modern methods, was capable of accurately forecasting market shocks.

Conclusion: This research, focusing on the Iranian stock market and utilizing intraday data from the Tehran Stock Exchange, has specifically and accurately addressed the prediction of Iran's market shocks. This localization of the model for the Iranian market significantly increased the accuracy of the results and enabled the simulation of the specific conditions of this market. Overall, this study presented an innovative hybrid model for analyzing the Iranian stock market, which can be beneficial for future analyses in similar markets.

Keywords: Stock Market Shock Prediction, Deep Learning, Time Series Analysis, ARMA-GARCH-CNN Model

EXTENDED ABSTRACT

Introduction

The rapid advancement of machine learning technologies and the growing need for improved market prediction methods underline the importance of this research. By employing these techniques, this study aims to develop a comprehensive model for predicting stock market shocks in Iran. Such a model can serve as an effective tool for investors and financial analysts, aiding in more accurate market analyses. Additionally, it can assist economic policymakers in preventing financial crises and making informed decisions to guide the national economy (He, 2024; Yazdan Panah & Ahmadi Mousavi, 2023).

Combining the strengths of statistical models and machine learning algorithms represents a significant step toward more accurate predictions of stock market fluctuations and shocks, especially in emerging markets like Iran (Hasanzadeh et al., 2023; Mahmoudi et al., 2023). The primary objective of this research is to propose a hybrid model for forecasting stock market shocks in the Tehran Stock Exchange using intraday data and machine learning algorithms.

This study seeks to answer the critical question: What is the optimal model for predicting stock market shocks in Iran based on a machine learning approach?

Methodology

This research is applied in nature, aiming to solve a specific real-world problem: predicting stock market shocks. The study is descriptive-analytical, focusing on examining and analyzing market changes over time using statistical methods.

The statistical population for this study consists of the Tehran Stock Exchange (TSE) index data for the period from March 20, 2018, to December 27, 2018. The dataset includes the opening, closing, highest, and lowest values for 15-minute intra-day intervals between 9:00 AM and 12:30 PM. These data were used for intra-day analysis and market shock prediction.

The sampling method employed was random sampling, ensuring a good representation of the statistical population. A total of 2,521 observations were used, extracted from the TSE's website and Rahavard Novin software.

The primary tools for data collection included:

Excel: Used for organizing and managing data and performing basic calculations.

Python 3.9: Used for advanced statistical analyses and implementation of ARMA-GARCH models and neural networks.

Rahavard Novin software and the TSE website: Used to obtain the intra-day stock data.

The collected data were preprocessed, which included handling missing data, correcting outliers, and normalizing data to ensure accurate analyses.

The dataset was divided into two categories: in-sample (training data) and out-of-sample (testing data). The in-sample data were used for model estimation, while the out-of-sample data were used for evaluating the model's generalization ability and performance, as well as for model comparison.

The variables in this study include historical time series data and market infrastructure information, such as technical indicators and volatility measurements of the TSE index.



The primary research question is to propose a model for predicting stock market shocks using machine learning approaches.

The research consists of three phases:

1. Phase One: Data collection and calculation of the logarithmic returns for the TSE index. The data consist of 15-minute intra-day prices, and the ARMA-GARCH model is used to fit the data and estimate market shocks (new changes in the series).

2. Phase Two: A set of candidate features, including technical indicators, historical time series from the ARMA-GARCH model, and the newly estimated market shocks, are prepared. Using the mRMR (minimum redundancy, maximum relevance) algorithm, 10 optimal features, including the market shocks, are selected for input into the neural network.

3. Phase Three: After preprocessing, which includes normalization, data splitting into training (80%) and testing (20%) sets, and standardization, the selected features are fed into a neural network. Following training and validation, the final models for each index are predicted. Various performance metrics, such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE), are used to evaluate the models' performance and determine which one accurately predicts market shocks and effectively simulates stock market behavior.

Findings

The research aimed to develop a model for predicting stock market shocks in Iran. The study consists of three main phases:

1. Market Shock Estimation Using ARMA-GARCH Model:

In the first phase, descriptive statistics were used to analyze the intra-day returns of the market index. Various statistical measures like mean, median, standard deviation, skewness, and kurtosis were calculated. The results revealed that the market index had a near-zero mean, indicating balanced fluctuations, though the negative median suggested a tendency towards lower returns. The tests for stationarity, including the ADF and PP tests, confirmed that the time series was stationary, which is essential for modeling purposes. The ARMA-GARCH model with the best fit (ARMA(2,3)-GARCH(1,1)) was selected based on the Akaike Information Criterion (AIC). This model produced expected returns, conditional variances, and standardized residuals (market shocks).

2. Feature Selection:

The second phase involved selecting features relevant to predicting market shocks. The candidate features included historical series from the ARMA-GARCH model, along with technical indicators and volatility measurements. Using the Minimum Redundancy Maximum Relevance (mRMR) algorithm, 10 optimal features were chosen for input into the neural network.

3. Training, Validation, and Prediction Using a Convolutional Neural Network (CNN):

In the final phase, the data were preprocessed through normalization and standardization before being split into training (80%) and testing (20%) sets. A 1D Convolutional Neural Network (CNN) was built, with 128 filters in the first convolutional layer and multiple subsequent layers to capture time-series patterns. The model was trained and validated, showing a significant reduction in error metrics (MSE, RMSE, MAE). The results from the trained model were evaluated through regression and statistical distributions, showing that the model could accurately predict market shocks.

Key Evaluation Metrics:

For both training and testing datasets, the model demonstrated low error values, indicating high prediction accuracy.

The final model for predicting stock market shocks in Iran is expressed as:

ARMA(2,3)-GARCH(1,1)-NN(conv1D(9-8-1)).

Discussion and Conclusion

The proposed model offers a significant advancement in forecasting stock market shocks, with potential applications beyond the stock market to include other financial markets in Iran, such as gold and currency markets. Comparing the results across these markets could provide deeper insights into the dynamics of financial systems.

This model can serve as the foundation for an early warning system, assisting investors and fund managers in identifying potential market fluctuations and enabling better risk management strategies. By integrating this model into an AI-driven analytical tool, both retail and institutional investors could make more informed decisions. Such a tool would provide real-time market analysis and forecasts, supporting data-driven investment strategies.

Moreover, the Tehran Stock Exchange could utilize the proposed model to analyze market behavior and enhance its regulatory mechanisms. This application would help mitigate unnecessary market volatility and foster greater investor confidence, contributing to the overall stability and growth of financial markets.

تجزیه و تحلیل شوک‌های بازار با مدل هیبریدی ARMA-GARCH و یادگیری عمیق برای پیش‌بینی پیشرفته

ابراهیم رحیمی^۱، احمد محمدی^{۲*}، علی اصغر متقی^۳، سید علی پایتختی اسکویی^۳

۱. دانشجوی دکتری، گروه حسابداری و مهندسی مالی، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.

۲. استادیار، گروه حسابداری و مهندسی مالی، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران (نویسنده مسئول).

۳. استادیار، گروه حسابداری و مهندسی مالی، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.

*ایمیل نویسنده مسئول: iacp_a-mohammadi@yahoo.com

چکیده

اطلاعات مقاله

نوع مقاله

پژوهشی/اصیل

نحوه استناد به این مقاله:

رحیمی ا، محمدی ا، متقی ع، پایتختی اسکویی ع. (۱۴۰۳). تجزیه و تحلیل شوک‌های بازار با مدل هیبریدی ARMA-GARCH و یادگیری عمیق برای پیش‌بینی پیشرفته. مدیریت پویا و تحلیل کسب و کار، ۳(۵)، ۱۷۸-۱۹۹.



© ۱۴۰۳ تمامی حقوق انتشار این مقاله متعلق به نویسنده(گان) است. انتشار این مقاله به صورت دسترسی آزاد مطابق با گواهی (CC BY 4.0) صورت گرفته است.

هدف: هدف اصلی پژوهش حاضر ارائه مدلی جهت پیش‌بینی شوک بازار سهام با رویکرد یادگیری عمیق در بورس اوراق بهادار تهران می‌باشد. **روش شناسی:** پژوهش حاضر از نظر هدف، کاربردی است. به این معنا که در پی حل یک مشکل خاص در دنیای واقعی است، یعنی پیش‌بینی شوک‌های بازار سهام. از نظر ماهیت، این پژوهش توصیفی-تحلیلی است، زیرا هدف آن بررسی و تحلیل تغییرات بازار در طول زمان است و از روش‌های آماری برای تحلیل داده‌ها استفاده می‌کند، برای این منظور از داده‌های ۱۵ دقیقه‌ای درون روزی شاخص کل در بازه زمانی ۹۷/۳/۲۰ تا ۹۷/۱۲/۲۷ که شامل مقادیر آغازین، پایانی، بیشترین و کمترین شاخص مذکور بودند استفاده شده است. در مدل پیشنهادی ابتدا اطلاعات آماری سری زمانی توسط مدل ARMA-GARCH استخراج و سپس با استفاده از روش یادگیری عمیق و عصبی پیچشی (کانولوشن) یک بعدی برای مدل کردن رابطه غیر خطی مشاهدات سری زمانی استفاده می‌شود و در نهایت مدل‌های مناسب انتخاب و بر اساس معیارهای ریشه میانگین خطاها (MSE)، ریشه میانگین مجذور خطاها (RMSE) و میانگین قدر مطلق خطاها (MAE) دقت پیش‌بینی‌ها ارزیابی می‌شود. **یافته‌ها:** نتایج این پژوهش نشان داد که مدل ترکیبی ARMA-GARCH-CNN، با بهره‌گیری از نقاط قوت هر دو روش سنتی و مدرن، قادر به پیش‌بینی دقیق شوک‌های بازار بوده است. **نتیجه گیری:** این پژوهش با تمرکز بر بازار سهام ایران و استفاده از داده‌های درون‌روزی بورس تهران، توانسته است به‌طور خاص و دقیق به پیش‌بینی شوک‌های بازار ایران بپردازد. این تمرکز بومی‌سازی مدل برای بازار ایران به‌طور چشمگیری دقت نتایج را افزایش داده و امکان شبیه‌سازی شرایط خاص این بازار را فراهم کرده است. به‌طور کلی، این پژوهش یک مدل ترکیبی نوآورانه برای تحلیل بازار سهام ایران ارائه داده که می‌تواند برای تحلیل‌های آینده در بازارهای مشابه مفید واقع شود.

کلیدواژه‌گان: پیش‌بینی شوک بازار سهام، یادگیری عمیق، تحلیل سری‌های زمانی، مدل ARMA-GARCH-CNN.

مقدمه

بازارهای مالی، به‌ویژه بازار سهام، به دلیل ماهیت پیچیده و پویای خود، همواره تحت تأثیر نوسانات و شوک‌های غیرمنتظره قرار دارند. شوک‌های بازار ممکن است ناشی از عوامل متعددی از جمله تغییرات سیاست‌های اقتصادی، بحران‌های مالی جهانی، تغییرات ناگهانی قیمت کالاها، تحولات سیاسی و اجتماعی، یا رویدادهای غیرقابل پیش‌بینی مانند پاندمی‌ها باشند. این شوک‌ها اثرات گسترده‌ای بر تصمیم‌گیری‌های سرمایه‌گذاری، سیاست‌های مالی و اقتصادی و در نهایت ثبات اقتصاد کلان دارند و مدیریت آن‌ها از چالش‌های اساسی حوزه مالی محسوب می‌شود (He, 2024; Yasmin & Ferdaous, 2023; Yazdan Panah & Ahmadi Mousavi, 2023).

شوک‌های بازار به طور مستقیم بر رفتار سرمایه‌گذاران، تصمیم‌گیری‌های اقتصادی، سیاست‌های مالی و اقتصادی و در نهایت بر سلامت کلی اقتصاد تأثیر می‌گذارند. در این راستا، پیش‌بینی دقیق این شوک‌ها به‌عنوان یکی از چالش‌های اصلی و پیچیده در حوزه مالی و اقتصادی به شمار می‌رود و می‌تواند به‌طور مؤثری در مدیریت ریسک و بهبود تصمیمات سرمایه‌گذاری و اقتصادی نقش ایفا کند (Mahdi & Shafiei, 2022).

پیش‌بینی شوک‌های بازار سهام به‌ویژه در بازارهای نوظهور، مانند بازار سهام ایران، اهمیت دوچندانی دارد. بازار سهام ایران به دلیل ویژگی‌هایی نظیر نوسانات شدید، تأثیرپذیری بالا از عوامل سیاسی-اقتصادی داخلی و خارجی، و نبود شفافیت کافی، شرایط پیچیده‌تری برای پیش‌بینی شوک‌ها فراهم کرده است. این موضوع نیاز به طراحی و توسعه مدل‌های پیش‌بینی دقیق و نوآورانه را دوچندان می‌کند (Mahmoudi et al., 2023; Salisu et al., 2024).

روش‌های پیش‌بینی کلاسیک نظیر مدل‌های ARIMA و GARCH، علی‌رغم موفقیت نسبی در مدل‌سازی الگوهای خطی و سری‌های زمانی، به دلیل محدودیت‌هایی نظیر عدم توانایی در پردازش حجم بالای داده‌ها و تحلیل الگوهای غیرخطی، در پیش‌بینی شوک‌های غیرمنتظره عملکرد ضعیفی نشان داده‌اند (Cai et al., 2022; Kim & Kim, 2022). در سال‌های اخیر، روش‌های یادگیری ماشین و یادگیری عمیق، به دلیل قابلیت تحلیل داده‌های حجیم و پیچیده و شناسایی الگوهای پنهان، به‌عنوان جایگزین‌های کارآمد در پیش‌بینی روندهای بازار معرفی شده‌اند (Messaoud et al., 2023; Xiao et al., 2017).

مدل‌های یادگیری ماشین نظیر جنگل‌های تصادفی، ماشین‌های بردار پشتیبان (SVM) و شبکه‌های عصبی پیچشی (CNN) به دلیل توانایی تحلیل روابط غیرخطی و شناسایی الگوهای پیچیده در داده‌ها، به‌عنوان ابزارهای مؤثر در پیش‌بینی شوک‌های بازار شناخته شده‌اند. به‌طور خاص، شبکه‌های عصبی عمیق نظیر LSTM و CNN در تحلیل داده‌های سری زمانی و پیش‌بینی دقیق‌تر شوک‌های بازار عملکرد چشمگیری داشته‌اند (Xiao et al., 2017).

با توجه به اهمیت پیش‌بینی شوک‌های بازار سهام در مدیریت ریسک، تصمیم‌گیری‌های سرمایه‌گذاری و جلوگیری از بحران‌های مالی و همچنین با توجه به اینکه بازار سهام ایران، به‌عنوان یک بازار نوظهور، با ویژگی‌های خاصی نظیر نوسانات بالا، تأثیرپذیری از عوامل سیاسی-اقتصادی داخلی و خارجی، و ساختار غیرشفاف شناخته می‌شود (Cai et al., 2022). در چنین بازاری، نبود ابزارهای پیش‌بینی دقیق و مناسب، منجر به تصمیم‌گیری‌های نادرست، افزایش ریسک سرمایه‌گذاری و خروج سرمایه از بازار شده است. این پژوهش با هدف ارائه مدلی ترکیبی شامل ARMA-GARCH و CNN طراحی شده است که با بهره‌گیری از نقاط قوت هر دو رویکرد، شوک‌های بازار سهام ایران را پیش‌بینی می‌کند. این مدل در سه مرحله طراحی شده است:

¹ Random Forest

- (۱) تخمین نوسانات با مدل ARMA-GARCH برای مدل‌سازی الگوهای خطی و واریانس شرطی.
 - (۲) انتخاب ویژگی‌های مؤثر با استفاده از الگوریتم حداقل افزونگی-حداکثر ارتباط (mRMR) برای شناسایی مهم‌ترین متغیرها و کاهش پیچیدگی داده‌ها.
 - (۳) پیش‌بینی شوک‌ها با استفاده از CNN برای شناسایی روابط غیرخطی و افزایش دقت پیش‌بینی.
- این پژوهش به‌طور خاص بر داده‌ها و شرایط بازار سهام ایران متمرکز است و با ارائه یک مدل بومی‌سازی شده، به دنبال طراحی و توسعه مدلی مبتنی بر یادگیری ماشین و یادگیری عمیق است که بتواند شوک‌های بازار سهام ایران را با دقت بالایی پیش‌بینی کرده و شکاف موجود در تحقیقات پیشین را پر کند. این پژوهش نه تنها به دنبال ارتقاء دقت پیش‌بینی‌ها است بلکه می‌خواهد چارچوبی جامع برای استفاده از فناوری‌های نوین در تحلیل و پیش‌بینی روندهای بازارهای مالی ارائه دهد.
- بازارهای مالی، به‌ویژه بازار سهام، نقش اساسی در تخصیص بهینه منابع و تأمین مالی فعالیت‌های اقتصادی ایفا می‌کنند. از آنجا که بازار سهام ابزار اصلی تأمین مالی شرکت‌ها و پروژه‌ها است، تغییرات و نوسانات آن می‌تواند به‌طور مستقیم بر روند رشد اقتصادی و حتی ثبات اقتصادی کشورها تأثیرگذار باشد (Fama, 1970). از سوی دیگر، شوک‌های ناگهانی بازار سهام می‌توانند تأثیرات قابل توجهی بر ثبات مالی، تصمیمات سرمایه‌گذاران، و سیاست‌گذاری‌های کلان اقتصادی و مالی داشته باشند (Herwiyanti et al., 2023; Mahdi & Shafiei, 2022; Yasmin & Ferdaous, 2023). از این رو، پیش‌بینی این شوک‌ها می‌تواند در مدیریت ریسک، بهبود کارایی بازار، و پیشگیری از بحران‌های مالی نقش حیاتی ایفا کند.
- پیش‌بینی شوک‌های بازار سهام نه تنها برای سرمایه‌گذاران، بلکه برای اقتصاد کلان نیز از اهمیت ویژه‌ای برخوردار است. شوک‌های بازار می‌توانند تأثیرات عمیقی بر پایداری اقتصادی و تصمیمات مالی داشته باشند. برای مثال، نوسانات شدید قیمت سهام ممکن است به بحران‌های مالی منجر شود و از آنجا که بازار سهام ایران به شدت تحت تأثیر عوامل اقتصادی و سیاسی داخلی و بین‌المللی قرار دارد، این نوسانات می‌توانند آثار منفی جدی بر اقتصاد کشور بگذارند (Eyshi Ravandi et al., 2024; Hasanzadeh et al., 2023; Yazdan, 2023; Panah & Ahmadi Mousavi, 2023).
- در چنین شرایطی، پیش‌بینی دقیق شوک‌های بازار، می‌تواند به سرمایه‌گذاران کمک کند تا از نوسانات شدید قیمت‌ها پیشگیری کرده و تصمیمات آگاهانه‌تری اتخاذ کنند. به علاوه، مدیریت صحیح ریسک در بازارهای مالی از جمله بازار سهام ایران، به‌طور مستقیم با سلامت سیستم مالی و اقتصادی کشور مرتبط است. از این رو، توسعه مدل‌های پیش‌بینی دقیق، می‌تواند به بهبود کیفیت تصمیمات سرمایه‌گذاری، کاهش ریسک‌های مالی، و افزایش کارایی بازار کمک کند (Bekaert & Harvey, 2003; Goodfellow et al., 2024).
- محدودیت‌های مدل‌های سنتی سری زمانی، مانند ARMA و GARCH، که معمولاً برای پیش‌بینی نوسانات بازار به کار گرفته می‌شوند، موجب کاهش دقت پیش‌بینی‌ها در شرایطی می‌شوند که داده‌ها دارای پیچیدگی‌های غیرخطی و ویژگی‌های جدیدی باشند. از این رو، استفاده از مدل‌های ترکیبی که شامل یادگیری ماشین و تکنیک‌های مدرن تحلیل داده هستند، می‌تواند به حل این مشکل کمک کند (Goodfellow et al., 2024).
- علاوه بر این، داده‌های تاریخی بازار سهام ایران از ویژگی‌های خاصی برخوردارند که باعث می‌شود مدل‌های پیش‌بینی به سادگی قادر به تحلیل آن‌ها نباشند. این داده‌ها شامل تغییرات ناگهانی در بازار، تأثیرات دولت در اقتصاد و نوسانات شدید ناشی از رویدادهای اقتصادی و سیاسی هستند. در این راستا، روش‌های یادگیری ماشین که قادر به شبیه‌سازی روابط پیچیده و غیرخطی هستند، می‌توانند مزیت‌های زیادی نسبت به مدل‌های سنتی داشته باشند (Goodfellow et al., 2024; Shiller, 2003).

ضرورت انجام این پژوهش در این است که با توجه به رشد سریع فناوری‌های یادگیری ماشین و نیاز به بهبود روش‌های پیش‌بینی بازار، می‌توان از این تکنیک‌ها برای تحلیل دقیق‌تر داده‌های بازار سهام ایران استفاده کرد. این پژوهش می‌تواند به توسعه مدلی جامع برای پیش‌بینی شوک‌های بازار سهام در ایران کمک کند و به عنوان ابزاری مؤثر برای سرمایه‌گذاران و تحلیلگران مالی به کار گرفته شود. به علاوه، این پژوهش می‌تواند به سیاست‌گذاران اقتصادی کمک کند تا از وقوع بحران‌های مالی جلوگیری کرده و تصمیمات بهتری برای هدایت اقتصاد کشور اتخاذ کنند.

در نهایت، ارائه مدلی که از توانایی‌های ترکیبی مدل‌های آماری و یادگیری ماشین بهره‌برداری کند، می‌تواند به عنوان یک گام مهم در پیش‌بینی دقیق‌تر نوسانات و شوک‌های بازار سهام در ایران و دیگر بازارهای نوظهور شناخته شود. هدف اصلی پژوهش حاضر ارائه یک مدل ترکیبی برای پیش‌بینی شوک‌های بازار سهام در بورس تهران با استفاده از داده‌های درون‌روزی و الگوریتم‌های یادگیری ماشین می‌باشد. بنابراین سوال اصلی پژوهش حاضر این است که مدل بهینه برای پیش‌بینی شوک بازار سهام بر اساس رویکرد یادگیری ماشین در ایران چگونه است؟

روش پژوهش

پژوهش حاضر از نظر هدف، کاربردی است. به این معنا که در پی حل یک مشکل خاص در دنیای واقعی است، یعنی پیش‌بینی شوک‌های بازار سهام. از نظر ماهیت، این پژوهش توصیفی-تحلیلی است، زیرا هدف آن بررسی و تحلیل تغییرات بازار در طول زمان است و از روش‌های آماری برای تحلیل داده‌ها استفاده می‌کند (Sekarani & Bougie, 2016).

جامعه آماری یا مجموعه داده استفاده شده در این پژوهش داده‌های شاخص کل در بورس تهران در بازه‌ی زمانی ۱۳۹۷/۳/۲۰ تا ۱۳۹۷/۱۲/۲۷ می‌باشد که شامل مقدار آغازین شاخص، مقدار پایانی آن، بیشترین و کمترین مقدار آن برای داده‌های ۱۵ دقیقه‌ای درون‌روزی از ساعت ۹ تا ۱۲:۳۰ هستند که برای تحلیل‌های درون‌روزی و پیش‌بینی شوک‌های بازار مورد استفاده قرار می‌گیرند.

برای نمونه‌گیری، از داده‌های ۱۵ دقیقه‌ای برای این شاخص استفاده شده است که در مجموع شامل ۲۵۲۱ مشاهده می‌باشد. این داده‌ها از وب‌سایت شرکت مدیریت فناوری بورس تهران و نرم‌افزار ره‌آورد نوین استخراج شده‌اند. برای نمونه‌گیری از نوع نمونه‌گیری تصادفی استفاده شده است تا نمایندگی خوبی از جامعه آماری حاصل شود (Babbie, 2013).

ابزار اصلی جمع‌آوری داده‌ها در این پژوهش شامل نرم‌افزارهای Excel (برای سازماندهی و مدیریت داده‌ها و انجام محاسبات ابتدایی)، نرم‌افزار Python 3.9 (برای تحلیل‌های آماری پیشرفته و پیاده‌سازی مدل‌های ARMA-GARCH و شبکه عصبی) و نرم‌افزار ره‌آورد نوین و سایت مدیریت فناوری بورس تهران (برای دریافت داده‌های درون‌روزی سهام) بود.

در این پژوهش، داده‌ها از سایت بورس تهران دریافت شده و پس از پیش‌پردازش، برای تجزیه و تحلیل بیشتر مورد استفاده قرار گرفته‌اند. پیش‌پردازش شامل حذف داده‌های گم‌شده، اصلاح مقادیر پرت و همسان‌سازی داده‌ها برای تجزیه و تحلیل‌های دقیق‌تر است.

کل مجموعه داده‌ها به دو دسته درون نمونه‌ای (داده‌های آموزش) و برون نمونه‌ای (داده‌های آزمایش) تقسیم می‌شوند. از داده‌ای درون نمونه‌ای برای برآورد و تخمین مدل استفاده می‌شود و از داده‌ای برون نمونه‌ای برای ارزیابی توانایی تعمیم و بررسی کارایی مدل و مقایسه مدل‌ها استفاده می‌شود.



متغیرهای این پژوهش شامل عناصر سری زمانی تاریخی و اطلاعات زیر ساختار بازار از جمله شاخص‌های فنی یا تکنیکال و اندازه‌گیری تلاطم (نوسانات) شاخص کل در بورس اوراق بهادار تهران می‌باشد.

مسئله اصلی این پژوهش ارائه مدلی برای پیش‌بینی شوک‌های بازار سهام با رویکرد یادگیری ماشین است.

پژوهش حاضر، شامل سه مرحله می‌باشد:

در مرحله اول، ابتدا داده‌ها جمع‌آوری و بازده لگاریتمی برای هر شاخص کل محاسبه می‌گردد. داده‌های جمع‌آوری شده شامل داده‌های درون‌روزی ۱۵ دقیقه‌ای قیمت‌های باز، بسته، بیشینه، و کمینه برای شاخص کل می‌باشد. سپس بازده لگاریتمی توسط مدل ARMA-GARCH برازش می‌شود و شوک‌های بازار (سری جدید تغییرات) تخمین زده می‌شوند.

در مرحله دوم، ابتدا مجموعه‌ای از ویژگی‌های کاندید شامل شاخص‌های فنی، سری زمانی تاریخی مدل ترکیبی آرما-گارچ و همچنین سری جدید تغییرات استخراج شده در مرحله قبل تهیه و با استفاده از روش فیلتر و بر اساس معیار اندازه‌گیری اطلاعات متقابل و الگوریتم حداقل افزونگی - حداکثر ارتباط (mRMR)، ۱۰ ویژگی بهینه همراه با شوک بازار تخمین زده شده جهت ورود به شبکه عصبی انتخاب می‌شود.

جدول ۱

ویژگی‌های کاندید (متغیرها)

طبقه بندی	ویژگی کاندید (متغیرهای ورودی)	فرمول
عناصر سری زمانی	نوآوری یا شوک بازار (Z_t)	$z_t = \frac{\varepsilon_t}{\sigma_t}$
	باقیمانده (پسماند)	$\varepsilon_{t-1}^2 = \alpha_1 \varepsilon_{t-1}^2$
	میانگین شرطی	$\mu_t = c + \rho_1 r_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$
	علامت نوآوری $\text{sign}(Z)$	آیا نوآوری مثبت است یا نه؟
	بازده	$r_{t,i} = \log\left(\frac{P_{t,i}}{P_{t,i-1}}\right)$
	علامت بازده	آیا بازده مثبت است یا نه؟
شاخص‌های فنی	دامنه	$RN = \frac{1}{4\log(2)} (High_t - Low_t)$
	دامنه صحیح	$TR_t = High_t - Low_t$
	اطلاعات نمودار ستونی ۱t	$BI1_t = \frac{(High_t - Open_t)}{(High_t - Low_t)}$
	اطلاعات نمودار ستونی ۲t	$BI2_t = \frac{(Open_t - Close_t)}{(High_t - Low_t)}$
	اطلاعات نمودار ستونی ۳t	$BI3_t = \frac{(Close_t - Low_t)}{(High_t - Low_t)}$
	واریانس تحقق یافته	$RV_t = \sum_{i=1}^M r_{t,i}^2$
	تلاطم دو قدرته	$BPV_t = \frac{\pi}{2} \left(\frac{M}{M-2} \right) \sum_{j=3}^M r_{t,j} t_{t,j-2} $

در مرحله سوم، پس از انتخاب مجموعه ویژگی‌های جدید، این داده‌ها شامل ۱۰ ویژگی انتخاب شده (۹ ویژگی به علاوه شوک بازار (متغیر هدف)) پس از پیش پردازش که شامل نرمالسازی، تقسیم داده‌ها به آموزش (۸۰٪ داده‌ها) و آزمایش (۲۰٪ داده‌ها) و استانداردسازی می‌باشد، جهت ایجاد مدل، وارد شبکه عصبی شده و پس از پایان یادگیری شبکه (آموزش و اعتبار سنجی)، مدل‌های نهایی برای شاخص کل

پیش بینی می‌شوند. برای ارزیابی عملکرد مدل‌ها، از معیارهای مختلفی همچون میانگین مربعات خطا (MSE)، ریشه میانگین مربعات خطا (RMSE) و میانگین قدر مطلق خطا (MAE) استفاده می‌شود. این ارزیابی کمک می‌کند تا مشخص شود که کدام مدل در پیش‌بینی شوک‌های بازار دقت بیشتری دارد و به‌طور مؤثری عملکرد بازار سهام را شبیه‌سازی می‌کند.

یافته‌ها

پژوهش حاضر که هدف آن ارائه مدلی جهت پیش‌بینی شوک بازار سهام می‌باشد مشتمل بر سه مرحله می‌باشد: مرحله اول؛ تخمین شوک بازار سهام با استفاده از مدل ARMA-GARCH، مرحله دوم؛ استخراج ویژگی‌های کاندید و انتخاب ویژگی می‌باشد و نهایتاً مرحله سوم، آموزش و اعتبار سنجی و پیش‌بینی شوک بازار برای شاخص مورد نظر می‌باشد. جهت ارزیابی مدل از معیارهای ریشه میانگین خطاها (MSE)، ریشه میانگین مجذور خطاها (RMSE) و میانگین قدر مطلق خطاها (MAE) استفاده خواهد شد.

کلیه مراحل پژوهش با استفاده از نرم افزار Python3.9 انجام گرفته است.

مرحله اول: تخمین شوک بازار (سری جدید تغییرات)

در ابتدا باید توسط آمار توصیفی بازده درون روزی مربوط به شاخص کل مورد بررسی قرار گیرد و روند متغیرها از لحاظ مرکزیت و پراکندگی مشخص شود. آمار توصیفی شامل شاخص‌های مرکزی (میانگین، میانه)، شاخص‌های پراکندگی (بیشینه، کمینه، انحراف معیار، ضریب چولگی و ضریب کشیدگی) می‌باشد. **جدول ۲** نتایج آمار توصیفی سری زمانی بازده لگاریتمی برای شاخص ذکر شده را نمایش می‌دهد.

جدول ۲

آمار توصیفی داده‌های بازده درون روزی

میانگین	میانه	بیشینه	کمینه	انحراف معیار	ضریب کشیدگی ^۱	ضریب چولگی ^۲
۰/۰۰۰۱۰۴	-۰/۰۰۰۰۲۳	۰/۰۱۴۲۷۳	-۰/۰۱۵۶۱۱	۰/۰۰۱۳۳۵	۳۸/۸۱۲۰۳۷۹۷	۲/۰۶۱۰۲۹

میانگین شاخص کل نزدیک به صفر است، که نشان می‌دهد تغییرات کلی شاخص کل در طول زمان متعادل بوده است. با این حال، میانه در مقداری منفی است، که ممکن است نشان‌دهنده تمایل بیشتر به کاهش بازده باشد.

مقادیر مثبت ضریب چولگی نشان‌دهنده این است که توزیع بازده به سمت مقادیر مثبت تمایل دارد (دم سمت راست بلندتر است). شاخص کل کشیدگی بسیار بالایی دارند، که نشان‌دهنده توزیع دم‌سنگین است. این ویژگی در بازارهای مالی رایج است و به معنای احتمال وقوع شوک‌های شدید است.

جهت بررسی مانایی بازده لگاریتمی درون روزی از آزمون ریشه واحد استفاده می‌شود. جهت انجام این آزمون، از آزمون دیکی فولر تعمیم یافته^۳ و فیلیپس پرون^۴ استفاده شده است. نتایج در **جدول ۳** ارائه شده است.

^۱ Kurtosis

^۲ Skewness

^۳ Augmented Dickey-Fuller Test|ADF

^۴ Philips-Perron |PP



جدول ۳

نتایج آزمون ریشه واحد شاخص کل

شاخص	آماره ADF	آماره PP
کل	-۶/۹۱۷	-۴۳/۵۵۷
	p val ue = (۰/۰۰۰)	p val ue = (۰/۰۰۰)
مقدار بحرانی آزمون در سطح (Critical Values)	-۳/۴۳ (%۱)	-۲/۵۷ (%۱۰) -۳/۴۳ (%۱) -۲/۸۶ (%۵)

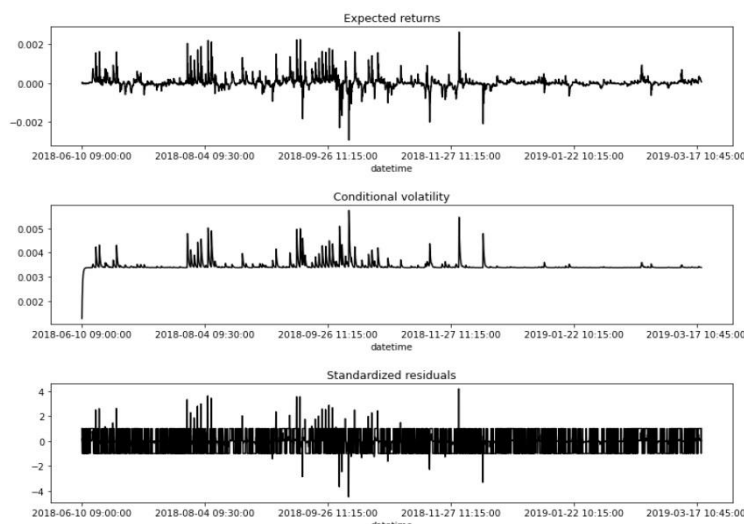
مطابق با نتایج ارائه شده، بر اساس آزمون‌های ADF و PP، مقدار قدر مطلق آماره آزمون در شاخص کل از مقدار بحرانی در سطح احتمال ۱ درصد، ۵ درصد و ۱۰ درصد بیشتر بوده و $\text{prob} < 0.05$ ، لذا سری زمانی بازده در شاخص فاقد ریشه واحد بوده و مانا می‌باشند. این نتیجه از اهمیت بالایی برخوردار است، زیرا مدلسازی سری‌های زمانی غیر مانا (مانند رگرسیون) می‌تواند نتایج گمراه کننده تولید کند.

برای انتخاب مرتبه مدل ترکیبی ARMA-GARCH الگوهای مختلف ARMA-GARCH با مراتب مختلف را آزمون کرده که در این پژوهش از معیارهای آکائیک^۱ برای انتخاب مدل مناسب استفاده شده است. کمترین میزان برای معیارهای AIC مناسب‌ترین مدل می‌باشد. نتیجه آزمون مقدار کمترین آکائیک را (۲۳۵۷۳/۳۲۶۱۳-) و مرتبه مدل ARMA-GARCH را (AR=2, MA=3, P=1, q=1) بدست آورد. بنابراین مناسب ترین مدل ترکیبی ARMA-GARCH برای شاخص کل ARMA(2,3)-GARCH(1,1) با درجه آزادی ۸ است. خروجی حاصل از مدل ARMA-GARCH شامل بازده مورد انتظار، واریانس شرطی و باقیمانده‌های استاندارد (شوک بازار) می‌باشد که برای شاخص کل در شکل ۱ نمایش داده شده است.

شکل ۱

خروجی مدل ARMA(2,3)-GARCH(1,1) برای شاخص کل

^۱ AIC



بازده مورد انتظار نشان می‌دهد نوسانات نسبتاً بالا در برخی دوره‌ها وجود دارد. این امر ممکن است به تأثیر عوامل کلان اقتصادی و شرایط عمومی بازار مربوط باشد.

نمودار نوسانات شرطی^۲ نشان می‌دهد که نوسانات در ابتدا به سرعت افزایش می‌یابد و سپس کاهش می‌یابد. این رفتار می‌تواند نشان‌دهنده واکنش بازار به یک رویداد خاص یا دوره بی‌ثباتی باشد.

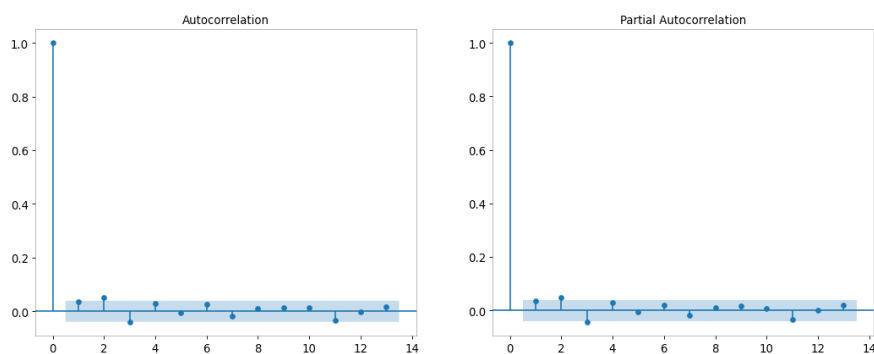
در نمودار باقیمانده‌های استاندارد شده^۳ وجود برخی پرش‌های بزرگ در داده‌های باقی‌مانده نشان می‌دهد که مدل ARMA- GARCH ممکن است برخی شوک‌ها را به‌طور کامل مدل‌سازی نکرده باشد.

در شکل ۲ ضرایب خود همبستگی و خود همبستگی جزئی نشان داده شده است. ضریب خود همبستگی^۴ نشان می‌دهد که یک مقدار در سری زمانی تا چه حد با مقادیر قبلی خودش در تاخیرهای مختلف^۵ مرتبط است. تمام وابستگی‌ها (مستقیم و غیر مستقیم) را محاسبه میکند و ضریب خود همبستگی جزئی^۶ تأثیر مستقیم یک مقدار با مقدار قبلی خودش را در یک تاخیر مشخص نشان می‌دهد، بدون اینکه تأثیر تاخیرهای میانی را در نظر بگیرد. فقط وابستگی مستقیم بین متغیرها را نشان می‌دهد.

شکل ۲

ضرایب AC و PAC برای سری جدید تغییرات شاخص کل

^۱ Expected Returns
^۲ Conditional Volatility
^۳ Standardized Residuals
^۴ AC
^۵ Lag
^۶ PAC



مقادیر نزدیک صفر برای این ضرایب نشان دهنده تصادفی بودن سری جدید تغییرات است.

این امر نشان می‌دهد که مدل ARMA-GARCH به خوبی توانسته الگوهای سری زمانی را مدلسازی کند.

چنانچه جدول ۴ آمار توصیفی سری جدید تغییرات را نکایش می‌دهد. نتایج نشان می‌دهند که میانگین و میانه شاخص نزدیک صفر

است که مطابق با خصوصیت داده‌های تصادفی است.

بالا بودن ضرایب کشیدگی و چولگی نشاندهنده غیر نرمال بودن توزیع در داده‌های شوک بازار میباشد.

جدول ۴

آمار توصیفی داده‌های شوک بازار (سری جدید تغییرات)

میانگین	میانه	بیشینه	کمینه	انحراف معیار	ضریب کشیدگی ^۱	ضریب چولگی ^۲
۰/۱۳۱۱۰۵	۰/۰۸۹۱۱۷	۵/۶۶۵۶۹۷	-۴/۸۷۵۲۳۷	۰/۴۴۰۴۷۶	۴۶/۰۴۸۹۷۵۴۲	۱/۷۰۶۶۷۶۱۴

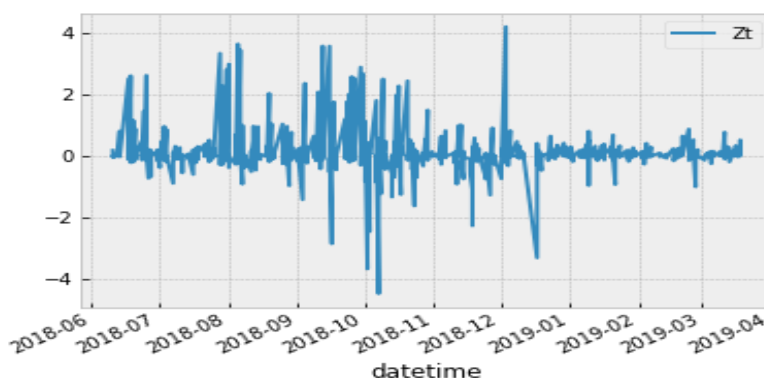
داده‌های سری زمانی تغییرات (شوک بازار) برای مدل ترکیبی ARMA(2,3)-GARCH(1,1) در شکل ۳، ارائه شده است.

شکل ۳

سری جدید تغییرات برای شاخص کل

^۱ Kurtosis

^۲ Skewness



از نمودار سری جدید تغییرات می‌توان نتیجه‌گیری کرد که رفتار این نمودار با پارامترهای مدل و توصیف آماری آنها سازگار است. لذا در ادامه مراحل، انتخاب ویژگی، نرمال سازی و استاندارد سازی داده‌ها را قبل از ورود به شبکه عصبی انجام خواهیم داد.

مرحله دوم: انتخاب ویژگی

در این پژوهش ویژگی‌های کاندید که شامل عناصر سری زمانی تاریخی مدل ARMA-GARCH و اطلاعات زیر ساخت بازار شامل شاخص‌های فنی، اندازه‌گیری‌های نوسانات (تلاطم‌های) با فراوانی زیاد و... هستند از مدل سان و همکاران (۲۰۱۹) برگرفته شده است که در آن شوک بازار متغیر وابسته (هدف) و دیگر ویژگی‌ها متغیر مستقل هستند.

در این پژوهش از روش فیلتر و بر اساس عملکرد اطلاعات متقابل و الگوریتم حداقل افزونگی - حداکثر ارتباط^۱ ویژگی‌های بهینه انتخاب می‌شوند.

الگوریتم حداقل افزونگی - حداکثر ارتباط یک روش انتخاب ویژگی در یادگیری ماشین است که برای پیدا کردن زیر مجموعه‌ای از ویژگی‌ها استفاده می‌شود. هدف این الگوریتم، بهبود عملکرد مدل‌ها با انتخاب ویژگی‌هایی است که بیشترین ارتباط را با متغیر هدف دارند و کمترین افزونگی یا وابستگی را با یکدیگر نشان می‌دهد.

مزایای این الگوریتم؛ کاهش ابعاد داده و جلوگیری از بیش‌برازش، بهبود کارایی و سرعت مدل‌های یادگیری ماشین و انتخاب ویژگی‌هایی که هم با هدف مرتبط اند و هم غیر افزونه (غیر همبسته) هستند.

از میان ۱۳ ویژگی مورد نظر، ۱۰ ویژگی برتر انتخاب و جهت بهینه‌سازی مدل وارد شبکه عصبی می‌شوند.

ویژگی‌های انتخاب شده شامل: $\text{sign}(r_t)$, Z_t , $BI3$, RN , $BI2$, TR , $BI1$, et , BPV , ε_y می‌باشند.

پس از انتخاب مجموعه ویژگی‌های جدید، این داده‌ها پس از نرمالیزاسیون و استاندارد سازی، به داده‌های آموزشی و آزمایشی تقسیم شده و جهت بهینه‌سازی مدل وارد شبکه عصبی می‌شوند.

مرحله سوم: آموزش و اعتبار سنجی

ابتدا کل داده‌ها بارگذاری می‌شوند. پس از آن مرحله پیش پردازش داده‌هاست که به شرح ذیل صورت می‌گیرد:

پیش‌پردازش داده‌ها معمولاً شامل نرمالیزاسیون، تقسیم داده‌ها به مجموعه‌های آموزش و تست و استاندارد سازی می‌شود.

داده‌ها با استفاده از تابع MinMaxScaler داده‌ها در دامنه $[-1,1]$ نرمال سازی می‌شوند.

^۱- mMRMR



این روش معمولاً برای داده‌های سری زمانی استفاده می‌شود، زیرا داده‌ها معمولاً دارای مقیاس‌های مختلف هستند و نرمالیزاسیون باعث می‌شود که مدل به راحتی از داده‌ها یاد بگیرد.

داده‌ها به مجموعه‌های آموزش (۸۰٪ داده‌ها) و آزمایش (۲۰٪ داده‌ها) تقسیم می‌شوند. این تقسیم‌بندی به‌طور استاندارد برای ارزیابی عملکرد مدل‌های یادگیری ماشین انجام می‌شود تا از احتمال بیش‌برازش جلوگیری شود.

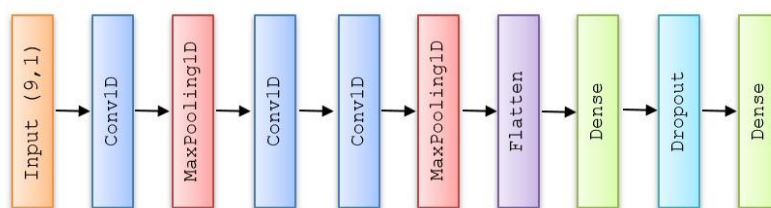
مجموعه داده (هم آموزش و هم آزمایش) را با استفاده از StandardScaler کتابخانه Sklearn استاندارد سازی می‌کنیم و پس از آن داده‌های آموزش را وارد شبکه عصبی می‌کنیم.

پس از پیش پردازش داده‌ها شبکه عصبی کانولوشن یک بعدی ساخته می‌شود. این شبکه مدل پیچیده‌ای است که از لایه‌های مختلف برای استخراج ویژگی‌ها و پیش‌بینی استفاده می‌کند. در این مدل، ورودی‌ها به‌طور خاص طراحی شده‌اند تا از اطلاعات سری زمانی (مانند داده‌های مالی یا بازار سهام) به‌طور موثر استخراج شود. در ادامه، بخش‌های مختلف این مدل را تشریح می‌کنم:

اولین لایه شبکه‌های عصبی مصنوعی پیچشی لایه ورودی می‌باشد که بایستی ورودی آن سائز داده ورودی باشد؛ از آنجا که ورودی‌های شبکه عصبی شامل ۱۰ ویژگی (۹ ویژگی به عنوان داده‌های ورودی و یک ویژگی (شوک بازار) به عنوان داده هدف) می‌باشد لذا ابعاد لایه ورودی (None,9,1) می‌باشد. اولین لایه کانولوشن با تعداد فیلتر ۱۲۸ و طول پنجره ۵ و قدم یک و با تابع فعالساز relu تشکیل و سپس لایه Maxpooling با سائز ۲ و قدم یک و سپس دو لایه کانولوشن با تعداد فیلتر به ترتیب ۶۴ و ۱۶ و اندازه پنجره به ترتیب ۳ و ۲ و تابع فعالساز relu تشکیل و سپس لایه Maxpooling با سائز ۲ و قدم یک و لایه faltten جهت مسطح سازی داده‌ها و سپس لایه‌های Dense با ورودی و خروجی ۳۲ و تابع فعالساز relu و Dropout(0.5) و در نهایت لایه Dense با ورودی ۳۲ و یک خروجی که شوک بازار می‌باشد تشکیل داده میشود تا مدل شبکه عصبی پیچشی تشکیل گردد. شکل ۴ معماری شبکه عصبی کانولوشن یک بعدی را برای شاخص کل نشان می‌دهد.

شکل ۴

معماری شبکه عصبی کانولوشن یک بعدی



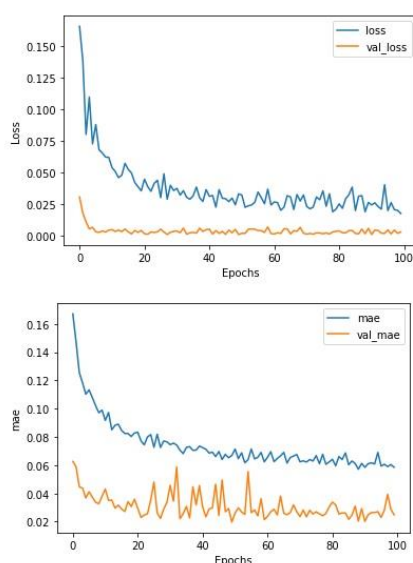
پس از ساخت شبکه عصبی مجموعه داده‌های آموزش در شبکه عصبی به آموزش (۸۰٪ داده‌های آموزشی) و اعتبار سنجی (۲۰٪ داده‌های آموزشی) تقسیم می‌شوند و فرآیند آموزش و اعتبار سنجی روی داده‌ها انجام می‌شود.

¹ overfitting

تنظیمات آموزش شبکه عصبی به این ترتیب زیر تنظیم شده است که تعداد دوره‌ها ۱۰۰، اندازه دسته‌ها^۱ ۶۴ و تعداد داده‌ها جهت اعتبار سنجی^۲ ۲۰ درصد باشد. معمولاً برای شبکه‌های عصبی استفاده می‌شوند. این تنظیمات به شبکه عصبی اجازه می‌دهند تا بر روی داده‌ها به خوبی آموزش ببیند و در عین حال ارزیابی مناسبی از عملکرد مدل انجام گیرد. پس از آموزش مدل، مدل آموزش داده شده ارزیابی خواهد شد.

شکل ۵

توابع زیان و میانگین قدر مطلق خطاها برای شاخص کل



نمودارهای توابع زیان و میانگین قدر مطلق خطاها در شکل ۵ نشان می‌دهند که مدل در حال نزدیک شدن به خطای صفر است. هرچه مقادیر خطا به صفر نزدیک‌تر باشند، عملکرد مدل بهتر است.

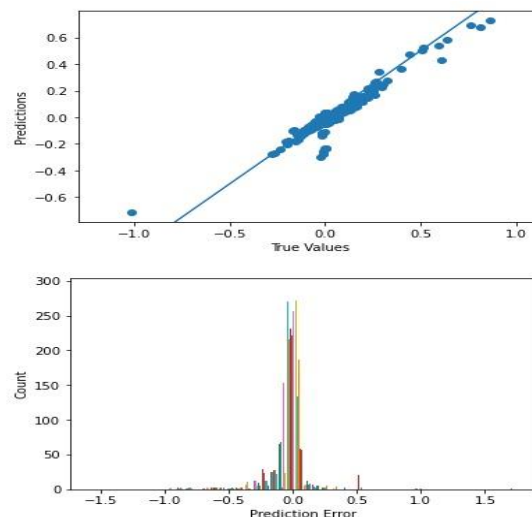
پس از ارزیابی مدل آموزش داده شده، داده‌های آزمایش مورد ارزیابی قرار می‌گیرند.

خروجی نمودار رگرسیون خطی و توزیع آماری بر اساس فرضیه صفر و همچنین خروجی نمودار پیش بینی شوک بازار در مقایسه با داده‌های آزمایش، مطابق شکل ۶ استخراج شده است.

شکل ۶

نمودار رگرسیون خطی و توزیع آماری بر اساس فرضیه صفر

¹ epoch
² batch size
³ validation split

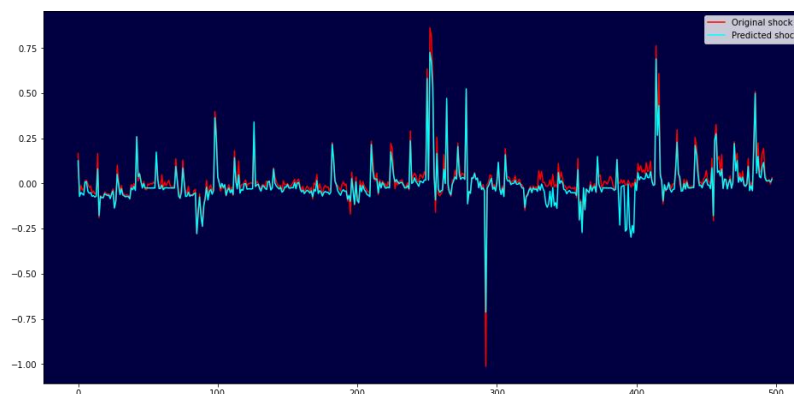


نمودار رگرسیون خطا نشان می‌دهند که رابطه بین شوک‌های پیش‌بینی شده و تخمین زده شده یک رابطه مستقیم می‌باشد. به عبارت دیگر، پیش‌بینی‌ها هم‌راستا با داده‌های واقعی هستند، که می‌تواند نشان‌دهنده دقت مدل در پیش‌بینی شوک‌ها باشد. توزیع نرمال خطای پیش‌بینی نشان می‌دهد که مدل از نظر آماری به‌طور صحیح شوک‌ها را پیش‌بینی کرده است. توزیع نرمال خطاهای پیش‌بینی، نشانه‌ای از دقت بالا و نبود انحراف سیستماتیک در پیش‌بینی‌هاست.

شکل ۷ نمودار پیش‌بینی شوک بازار سهام (خطوط سبز رنگ) در مقایسه با شوک‌های تخمین زده شده (خطوط قرمز رنگ) در مرحله اول حاصل از مدل ARMA-GARCH را نمایش می‌دهد. این نمودار کمک می‌کند تا تفاوت‌های پیش‌بینی مدل در مقایسه با داده‌های واقعی مشاهده شوند و عملکرد کلی مدل در پیش‌بینی شوک‌ها ارزیابی گردد.

شکل ۷

نمودار پیش‌بینی شوک بازار در مقایسه با داده‌های آزمایش برای شاخص کل



پس از ارزیابی مدل با استفاده از داده‌های آزمایشی، مشاهده می‌شود که مدل توانسته است به‌خوبی پیش‌بینی شوک‌ها را انجام دهد. از نمودارهای رگرسیون خطی و توزیع آماری می‌توان نتیجه گرفت که پیش‌بینی‌های مدل با داده‌های آزمایشی مطابقت خوبی دارند و دقت پیش‌بینی بالا است.

نتایج ارزیابی مدل برای در **جدول ۵** نمایش داده شده است.

جدول ۵

خلاصه نتایج ارزیابی مدل برای شاخص کل (برای داده‌های آموزشی و آزمایشی)

داده‌های آموزشی			داده‌های آزمایشی		
MAE	RMSE	MSE	MAE	RMSE	MSE
۰/۰۲۷۴	۰/۰۴۵۸	۰/۰۰۲۱	۰/۰۴۰۵	۰/۱۴۶۹	۰/۰۲۱۶

همانطور که در **جدول ۵** مشاهده می‌گردد، مقادیر خطای مدل برای داده‌های آموزشی و آزمایشی به‌طور کلی پایین و در سطح مناسبی قرار دارند که نشان‌دهنده دقت بالای پیش‌بینی‌هاست. همچنین مقدار MAE در داده‌های آزمایش خیلی نزدیک به داده‌های آموزشی می‌باشد و تقریباً مشابه است که این نشان‌دهنده دقت خوب آزمایش مدل می‌باشد. این نشان می‌دهد که مدل توانسته است عملکرد خوبی در پیش‌بینی داده‌های جدید (داده‌های آزمایشی) داشته باشد.

در انتها در پاسخ به سوال اصلی پژوهش که مدل بهینه برای پیش‌بینی شوک بازار سهام بر اساس رویکرد یادگیری ماشین در ایران چگونه است، میتوان مدل زیر را به عنوان مدل پیش‌بینی شوک بازار سهام در ایران ارائه داد:

$$\text{ARMA}(2,3)\text{-GARCH}(1,1)\text{-NN}(\text{conv1D}(9\text{-}8\text{-}1))$$

بحث و نتیجه‌گیری

هدف اصلی پژوهش ارائه یک مدل ترکیبی برای پیش‌بینی شوک‌های بازار سهام در بورس تهران با استفاده از داده‌های درون‌روزی و الگوریتم‌های یادگیری ماشین بود. با توجه به پیچیدگی‌های بازار سهام ایران، از جمله نوسانات بالا، تأثیرپذیری از عوامل سیاسی و اقتصادی، و نبود ابزارهای پیش‌بینی دقیق، ارائه مدلی جامع که بتواند دقت پیش‌بینی‌ها را افزایش دهد، ضروری به نظر می‌رسید. نتایج این پژوهش نشان داد که مدل ترکیبی ARMA-GARCH-CNN، با بهره‌گیری از نقاط قوت هر دو روش سنتی و مدرن، قادر به پیش‌بینی دقیق شوک‌های بازار بوده است. این مدل در سه مرحله طراحی و اجرا شد: در مرحله اول نوسانات خطی تحلیل شد، بدین صورت که با استفاده از مدل ARMA-GARCH، نوسانات شرطی بازار به دقت تحلیل و روابط خطی موجود در داده‌ها استخراج شدند. در مرحله بعد الگوریتم mRMR ویژگی‌هایی با بیشترین ارتباط با شوک‌های بازار و کمترین همبستگی با یکدیگر را شناسایی کرد. این انتخاب، علاوه بر بهبود کارایی مدل، باعث کاهش پیچیدگی و جلوگیری از بیش‌برازش شد. در انتها شبکه عصبی کانولوشن (CNN) با تحلیل ویژگی‌های منتخب، توانست الگوهای پیچیده و غیرخطی موجود در داده‌ها را شناسایی کند و پیش‌بینی شوک‌ها را با دقت بالایی انجام دهد. یافته‌های این پژوهش تأیید کرد که مدل ترکیبی ARMA-GARCH-CNN، در مقایسه با مدل‌های سنتی مانند ARMA یا GARCH به تنهایی، عملکرد بهتری داشته و توانسته است با معیارهای خطای کمتر (MSE, RMSE, MAE) دقت بیشتری در پیش‌بینی شوک‌ها ارائه دهد.



پژوهش حاضر با ارائه یک مدل ترکیبی ARMA-GARCH-CNN که از CNN به‌عنوان بخشی از مدل برای تحلیل روابط غیرخطی استفاده کرده است، توانسته است نسبت به بسیاری از تحقیقات مشابه که از مدل‌های تک‌متغیره یا الگوریتم‌های ساده‌تر استفاده کرده‌اند، مزایای قابل توجهی به‌دست آورد. ترکیب CNN با ARMA و GARCH توانسته است تحلیل دقیق‌تری از روندها، نوسانات و شوک‌های بازار ارائه دهد. در حالی که در بسیاری از پیشینه‌ها از مدل‌های مانند GARCH-SVM، ARIMA-LSTM و SVM استفاده شده است که بیشتر به تحلیل روندهای خطی یا تحلیل داده‌های با پیچیدگی کمتر محدود هستند، پژوهش حاضر از توانایی‌های CNN برای شبیه‌سازی روابط غیرخطی پیچیده بهره برده است که موجب افزایش دقت پیش‌بینی‌های کوتاه‌مدت و شوک‌های بازار شده است.

یکی دیگر از ویژگی‌های برجسته پژوهش حاضر، استفاده از داده‌های درون‌روزی برای تحلیل شاخص کل بورس تهران است. این داده‌ها به‌ویژه برای تحلیل نوسانات لحظه‌ای و پیش‌بینی شوک‌های بازار سهام ایران انتخاب شده‌اند و موجب دقت بالاتر پیش‌بینی‌ها در مقایسه با داده‌های روزانه یا ماهانه می‌شود. این رویکرد با استفاده از داده‌های دقیق و کوتاه‌مدت به‌ویژه در بازار سهام ایران که تحت تأثیر نوسانات و واکنش‌های سریع است، مزیتی قابل توجه دارد.

همچنین، استفاده از الگوریتم mRMR برای انتخاب ویژگی‌های بهینه یکی از نوآوری‌های پژوهش حاضر است که به کاهش پیچیدگی‌های اضافی و افزایش دقت مدل کمک کرده است. به‌طور خاص، این روش به‌طور مؤثری ویژگی‌های غیرمرتبط را حذف کرده و ویژگی‌های کلیدی برای پیش‌بینی شوک‌های بازار را نگه می‌دارد. در تحقیقات پیشین، این روش به‌طور گسترده استفاده نشده است و این می‌تواند نقطه قوت این پژوهش باشد.

در نهایت، این پژوهش با تمرکز بر بازار سهام ایران و استفاده از داده‌های درون‌روزی بورس تهران، توانسته است به‌طور خاص و دقیق به پیش‌بینی شوک‌های بازار ایران بپردازد. این تمرکز بومی‌سازی مدل برای بازار ایران به‌طور چشمگیری دقت نتایج را افزایش داده و امکان شبیه‌سازی شرایط خاص این بازار را فراهم کرده است. به‌طور کلی، این پژوهش یک مدل ترکیبی نوآورانه برای تحلیل بازار سهام ایران ارائه داده که می‌تواند برای تحلیل‌های آینده در بازارهای مشابه مفید واقع شود.

در مقایسه با تحقیقات مشابه، پژوهش حاضر از یک مدل ترکیبی ARMA-GARCH-CNN استفاده کرده است که شامل سه بخش اصلی است: مدل ARMA برای تحلیل روندهای کوتاه‌مدت داده‌های زمانی، مدل GARCH برای مدل‌سازی واریانس شرطی و نوسانات بازار، و CNN به‌عنوان بخشی از یادگیری عمیق برای تحلیل الگوهای غیرخطی و پیچیده. این مدل ترکیبی به‌ویژه برای پیش‌بینی شوک‌های بازار سهام ایران طراحی شده است. این رویکرد به‌طور قابل توجهی نسبت به بسیاری از تحقیقات پیشین که بیشتر به استفاده از مدل‌های تک‌بعدی مانند ARIMA-LSTM، GARCH-SVM یا SVM پرداخته‌اند، مزیت خاصی دارد (Goodfellow et al., 2024; He, 2024; Salisu et al., 2024).

همچنین، تحقیقات پیشین (Eyshi Ravandi et al., 2024; Goodfellow et al., 2024; Mahmoudi et al., 2023) بیشتر به ترکیب مدل‌های آماری و الگوریتم‌های یادگیری ماشین پرداخته‌اند، در حالی که مدل شما با ترکیب ARMA، GARCH و CNN به‌طور هم‌زمان توانسته است قابلیت‌های مختلف مدل‌ها را در یک سیستم واحد به خدمت بگیرد. این ترکیب مدل‌های آمار سنتی با یادگیری عمیق، قدرت تحلیل هم‌زمان روندها، نوسانات و الگوهای پیچیده را فراهم کرده است.

داده‌های مورد استفاده در پژوهش حاضر شامل داده‌های درون‌روزی ۱۵ دقیقه‌ای برای شاخص کل در بازار سهام ایران بوده است. این داده‌ها به‌ویژه برای پیش‌بینی شوک‌های کوتاه‌مدت و تحلیل نوسانات لحظه‌ای بازار انتخاب شده‌اند. استفاده از داده‌های درون‌روزی در

مقایسه با داده‌های روزانه یا ماهانه که در بسیاری از تحقیقات مشابه استفاده شده، مزیت قابل توجهی ایجاد کرده است (Hasanzadeh et al., 2023; Kim & Kim, 2022; Mahdi & Shafiei, 2022; Xiao et al., 2017).

این استفاده از داده‌های درون‌روزی، علاوه بر افزایش دقت پیش‌بینی، به مدل این امکان را داده است که نوسانات سریع و واکنش‌های لحظه‌ای بازار را نیز شبیه‌سازی کند. همچنین، استفاده از داده‌های خاص بورس تهران باعث شده تا پژوهش حاضر با شرایط خاص بازار ایران هماهنگ‌تر باشد و نتایج مرتبط‌تری به‌دست دهد. یکی از ویژگی‌های برجسته پژوهش حاضر استفاده از CNN به‌عنوان بخش اصلی مدل برای تحلیل روابط غیرخطی پیچیده است. پژوهش حاضر با استفاده از CNN به‌عنوان یک شبکه عصبی کانولوشنی، توانسته است روابط پیچیده و غیرخطی در داده‌ها را به‌طور مؤثرتری شبیه‌سازی کند. CNN به‌ویژه برای تحلیل داده‌های ساختارمند و نیمه‌ساختارمند مانند داده‌های مالی و بازار سهام بسیار مؤثر است. ترکیب این مدل با ARMA و GARCH نیز امکان تحلیل هم‌زمان روندها، نوسانات و الگوهای پیچیده را فراهم کرده است، که در پیشینه پژوهش‌ها کمتر دیده می‌شود. همچنین، برای ارزیابی عملکرد مدل در پژوهش حاضر، از معیارهای RMSE، MAE و MSE استفاده شده است. استفاده از معیارهای RMSE و MAE برای ارزیابی دقت پیش‌بینی به‌ویژه در تحلیل نوسانات کوتاه‌مدت و شوک‌های بازار بسیار مؤثر است و این معیارها در پیشینه پژوهش‌های مشابه کمتر به این‌طور دقیق به‌کار گرفته شده‌اند. در پژوهش حاضر، از الگوریتم mRMR برای انتخاب ویژگی‌های بهینه استفاده شده است. این روش به‌طور مؤثری ویژگی‌های غیرمرتبط را حذف کرده و ویژگی‌های مرتبط برای پیش‌بینی شوک‌های بازار را نگه می‌دارد. این امر موجب افزایش دقت پیش‌بینی و کاهش پیچیدگی مدل شده است.

این تفاوت در استفاده از روش‌های انتخاب ویژگی نشان‌دهنده نوآوری پژوهش حاضر است، زیرا انتخاب ویژگی‌های بهینه موجب بهبود عملکرد مدل و جلوگیری از پیچیدگی‌های اضافی شده است.

یکی از ویژگی‌های منحصر به فرد تحقیق حاضر، تمرکز ویژه بر بازار سهام ایران است. در این پژوهش، از داده‌های درون‌روزی بورس تهران برای پیش‌بینی شوک‌ها و نوسانات بازار استفاده شده است. تمرکز بر بازار ایران به مدل اجازه داده تا نتایج دقیق‌تری در خصوص نوسانات و شوک‌های بازار ایران ارائه دهد. این بومی‌سازی مدل و تمرکز بر شاخص کل بورس تهران، علاوه بر اینکه به مدل اجازه می‌دهد که نتایج دقیق‌تری برای شرایط خاص بازار سهام ایران به‌دست آید، همچنین موجب افزایش کاربردپذیری مدل در تحلیل و پیش‌بینی شوک‌های بازار ایران می‌شود. به‌ویژه در کشوری مانند ایران که بازار سهام آن ویژگی‌ها و چالش‌های خاص خود را دارد، مدل‌های بومی و خاص می‌توانند عملکرد بهتری نسبت به مدل‌های جهانی داشته باشند. در تحقیقات پیشین، استفاده از داده‌های کلی‌تر مانند داده‌های کلان اقتصادی یا داده‌های روزانه به جای داده‌های دقیق و درون‌روزی، منجر به تحلیل‌های محدودتر و دقیق‌تر کمتری شده است. علاوه بر این، با توجه به اینکه بازار سهام ایران تحت تأثیر عوامل سیاسی، اقتصادی، و اجتماعی خاص خود است، مدل بومی‌سازی شده قادر است پیش‌بینی‌های دقیق‌تری را ارائه دهد که در مدل‌های مبتنی بر داده‌های جهانی به‌سختی قابل دستیابی است.

یکی محدودیت‌های این پژوهش تلفیق مدل ARMA-GARCH و شبکه عصبی کانولوشن بود که نیازمند تنظیم دقیق پارامترها و اجرای چندین مرحله بهینه‌سازی است که زمان‌بر و پیچیده است (هیاتون و همکاران، ۲۰۱۷). یکی دیگر از محدودیت‌ها تأثیر عوامل خارجی بود؛ بازار ایران به‌شدت تحت تأثیر عوامل خارجی نظیر تحریم‌ها و تغییرات سیاست‌های بین‌المللی است که پیش‌بینی دقیق این عوامل نیازمند تحلیل‌های بیشتری است.



پیشنهاد می‌شود در تحقیقات آتی از مدل‌های پیشرفته‌تر نظیر شبکه‌های عصبی بازگشتی (RNN) و مدل‌های مبتنی بر توجه (Transformer) استفاده شود تا دقت پیش‌بینی افزایش یابد. همچنین گنجاندن متغیرهای کلان اقتصادی مانند نرخ بهره، نرخ ارز، و نرخ تورم و یا استفاده از سایر شاخص‌های تکنیکال به عنوان ورودی‌های اضافی می‌تواند پیش‌بینی‌ها را به شرایط اقتصادی نزدیک‌تر کند. مدل پیشنهادی را می‌توان برای پیش‌بینی شوک‌های سایر بازارهای مالی ایران، نظیر بازار طلا و ارز، به کار گرفت و نتایج آن را با بازار سهام مقایسه کرد. همچنین مدل ارائه شده می‌تواند به عنوان هسته اصلی یک سیستم هشدار زودهنگام برای سرمایه‌گذاران و مدیران صندوق‌ها عمل کند. این سیستم می‌تواند نوسانات احتمالی بازار را پیش‌بینی کرده و هشدارهای لازم را برای مدیریت ریسک ارائه دهد. با استفاده از مدل پیشنهادی، یک ابزار تحلیلی مبتنی بر هوش مصنوعی طراحی شود که سرمایه‌گذاران خرد و کلان بتوانند با استفاده از آن تصمیمات بهتری بگیرند. این ابزار می‌تواند اطلاعات بازار را به صورت زنده تحلیل کرده و پیش‌بینی‌های لازم را ارائه دهد. سازمان بورس و اوراق بهادار تهران می‌تواند از این مدل برای تحلیل رفتار بازار و بهبود مکانیزم‌های نظارتی خود استفاده کند. این امر به کاهش نوسانات غیرضروری بازار و افزایش اعتماد سرمایه‌گذاران کمک خواهد کرد.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافی وجود ندارد.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

موازن اخلاقی

در انجام این پژوهش تمامی موازن و اصول اخلاقی رعایت گردیده است.

شفافیت داده‌ها

داده‌ها و مآخذ پژوهش حاضر در صورت درخواست از نویسنده مسئول و ضمن رعایت اصول کپی رایت ارسال خواهد شد.

حامی مالی

این پژوهش حامی مالی نداشته است.

References

- Babbie, E. (2013). *The Practice of Social Research*. Cengage Learning. [https://books.google.com/books?hl=en&lr=&id=IFvjDwAAQBAJ&oi=fnd&pg=PP1&dq=Babbie,+E.+\(2013\).+The+Practice+of+Social+Research.+Cengage+Learning.+&ots=I4zT3G3OPa&sig=rd5UQ-MVdmRU7OWQy3DSVPGoYTA](https://books.google.com/books?hl=en&lr=&id=IFvjDwAAQBAJ&oi=fnd&pg=PP1&dq=Babbie,+E.+(2013).+The+Practice+of+Social+Research.+Cengage+Learning.+&ots=I4zT3G3OPa&sig=rd5UQ-MVdmRU7OWQy3DSVPGoYTA)
- Bekaert, G., & Harvey, C. R. (2003). Emerging Equity Market Volatility. *Journal of Financial Economics*, 43(1), 29-77. [https://doi.org/10.1016/S0304-405X\(96\)00889-6](https://doi.org/10.1016/S0304-405X(96)00889-6)
- Cai, D., Zhang, T., Han, K., & Liang, J. (2022). Economic Policy Uncertainty Shocks and Chinese Stock Market Volatility: An Empirical Analysis With SVAR. *Complexity*, 2022(1). <https://doi.org/10.1155/2022/6944318>

- Eyshi Ravandi, M., Moeinaddin, M., Taftiyan, A., & Rostami Bashmani, M. (2024). Investigating the Impact of Investor Sentiment and Liquidity on Stock Returns of the Iranian Stock Exchange. *Dynamic Management and Business Analysis*, 3(1), 40-52. <https://doi.org/10.22034/dmbaj.2024.2038046.1068>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2024). Generative adversarial nets. *Advances in Neural Information Processing Systems*,
- Hasanzadeh, I., Sheikh, M. J., Arabzadeh, M., & Farzinfar, A. A. (2023). The Role of Economic Policy Uncertainty in Relation to Financial Market Instability and Stock Liquidity in Tehran Stock Exchange Companies. *Dynamic Management and Business Analysis*, 2(3), 163-178. <https://doi.org/10.22034/dmbaj.2024.2031971.2315>
- He, Y. (2024). A Study of Factors Influencing Investor Expectations in the Capital Market. *SHS Web of Conferences*, 188, 01020. <https://doi.org/10.1051/shsconf/202418801020>
- Herwiyanti, E., Suwito, G., & Astuti, T. (2023). Investment Decisions in the Capital Market: Perceptions of Christian Students. *Journal of Economics Finance and Management Studies*, 06(07). <https://doi.org/10.47191/jefms/v6-i7-49>
- Kim, J., & Kim, Y.-S. (2022). Market-wide Shocks and the Predictive Power for the Real Economy in the Korean Stock Market. *Pacific Economic Review*, 27(4), 380-399. <https://doi.org/10.1111/1468-0106.12405>
- Mahdi, M., & Shafiei, G. (2022). New Capital Market in the Light of Sixth Development Plan Law (Functionality of blockchain technology in the capital market). *Journal of Securities Exchange*, 14(56), 197-224. <https://www.sid.ir/paper/955559/fa>
- Mahmoudi, R., Talebnia, G. A., Vakilifard, H., Ahmadi, F., & Moradi Shahdadi, K. (2023). Providing a model of sustainability reporting with the foundation data theory approach in companies listed on the Tehran Stock Exchange. *Financial Economics*, 64(17), 207-222. <https://doi.org/10.30495/fed.2023.705598>
- Messaoud, D., Ben Amar, A., & Boujelbene, Y. (2023). Investor sentiment and liquidity in emerging stock markets. *Journal of Economic and Administrative Sciences*, 39(4), 867-891. <https://doi.org/10.1108/JEAS-11-2020-0198>
- Salisu, A. A., Demirer, R., Gupta, R., Sangeetha, J. M., & Alfia, K. J. (2024). Technological shocks and stock market volatility over a century Financial stock market forecast using evaluated linear regression based machine learning technique. *Journal of Empirical Finance*, 79, 101561. <https://doi.org/10.1016/j.measen.2023.100950>
- Sekaran, U., & Bougie, R. (2016). *Research Methods for Business: A Skill-Building Approach*. Wiley & Sons. [https://books.google.com/books?hl=en&lr=&id=Ko6bCgAAQBAJ&oi=fnd&pg=PA19&dq=Sekaran,+U.,+%26+Bougie,+R.+\(2016\).+Research+Methods+for+Business:+A+Skill-Building+Approach.+Wiley+%26+Sons.+&ots=2D0P_6JWIT&sig=jeDfgg_r2wGkz8fnV0QJj41bts8](https://books.google.com/books?hl=en&lr=&id=Ko6bCgAAQBAJ&oi=fnd&pg=PA19&dq=Sekaran,+U.,+%26+Bougie,+R.+(2016).+Research+Methods+for+Business:+A+Skill-Building+Approach.+Wiley+%26+Sons.+&ots=2D0P_6JWIT&sig=jeDfgg_r2wGkz8fnV0QJj41bts8)
- Shiller, R. J. (2003). *The New Financial Order: Risk in the 21st Century*. Princeton University Press. <https://doi.org/10.1515/9781400825479>
- Xiao, K., Liu, Q., Liu, C., & Xiong, H. (2017). Price Shock Detection With an Influence-Based Model of Social Attention. *ACM Trans. Manage. Inf. Syst.*, 9(1), Article 2. <https://doi.org/10.1145/3131781>
- Yasmin, F., & Ferdaous, J. (2023). Behavioral Biases Affecting Investment Decisions of Capital Market Investors in Bangladesh: A Behavioral Finance Approach. *Investment Management and Financial Innovations*, 20(2), 149-159. [https://doi.org/10.21511/imfi.20\(2\).2023.13](https://doi.org/10.21511/imfi.20(2).2023.13)
- Yazdan Panah, M., & Ahmadi Mousavi, S. M. (2023). Legal Supervision Styles for Capital Market Health. *Dynamic Management and Business Analysis*, 2(3), 196-207. <https://doi.org/10.22034/dmbaj.2024.2037208.1063>