



Shahid Bahonar
University of Kerman



Iranian E-Commerce Scientific
Association

Performance Comparison of Deep Learning Models for Forecasting Iran's Basic Metals Industry Stock Index

Pedram Mohammad Hassan¹, Mojtaba Bahmani², Mehran Hoseini³ and Sadegh Esmaeili⁴

1. **Corresponding Author**, Department of Economics, Faculty of Economics, University of Tehran, Tehran, Iran.
Email: p.mohamadhasan@ut.ac.ir
2. Department of Economics, Faculty of Management and Economics, Shahid Bahonar University of Kerman, Kerman, Iran.
Email: mbahmani@uk.ac.ir
3. Department of Industrial Engineering, Faculty of Management and Industrial Engineering, Malek Ashtar University of Technology, Tehran, Iran. **Email:** sm.hoseini@mut.ac.ir
4. Department of Economics, Faculty of Economics, Allameh Tabataba'i University, Tehran, Iran.
Email: esmaeili.sadegh13@gmail.com

ARTICLE INFO

A B S T R A C T

Article Type:

Research Article.

Article History:

Received: 9 April 2026

Received in revised form: 7 July 2026

Accepted: 20 July 2026

Available online:

Keywords:

Stock Price Index,
Macroeconomic Variables,
Machine Learning,
Recurrent Neural Networks.

JEL Classification:

C45, C53, E44.

Objective: The purpose of this study is to evaluate the predictive accuracy of different neural network models in forecasting the Iran Basic Metals Industry Stock Index.

Method: To achieve this objective, daily data on the Iran Basic Metals Industry Stock Index from April 2011 to March 2025 were analyzed, together with key macroeconomic variables, including the exchange rate, oil price, money supply, consumer price index, and gold price. Three neural network models—Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU)—were employed. Their forecasting performance was evaluated using the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Theil's Inequality Coefficient (Theil's U).

Results: The results indicate that the Gated Recurrent Unit (GRU) model outperformed the other models in terms of forecasting accuracy. Specifically, the GRU architecture demonstrated a superior ability to capture the nonlinear and temporal dynamics of financial time series, producing more accurate forecasts than both the Long Short-Term Memory (LSTM) and Recurrent Neural Network (RNN) models.

Conclusion: The findings highlight the importance of advanced machine learning techniques in improving the prediction of financial market indicators. More accurate forecasts can reduce investment risk, enhance decision-making, and improve the efficiency of capital markets. Therefore, developing machine learning-based forecasting tools and strengthening financial data infrastructures can support policymakers and investors in making more informed decisions while promoting capital market stability.

Cite this article: Mohammad Hassan, P., Bahmani, M., Hoseini, M., & Esmaeili, S. (2026). Performance comparison of deep learning models for forecasting Iran's basic metals industry stock index. *Journal of Development and Capital*, *(*)*, *-*



Publisher: Shahid Bahonar University of Kerman.

©The Author(s).

Introduction

Financial markets, particularly stock markets, play a crucial role in the efficient allocation of resources and the promotion of economic growth. Stock market indices reflect both corporate performance and investors' expectations regarding future economic conditions. Consequently, forecasting stock market indices has become an essential tool for reducing investment risk and improving decision-making for investors, policymakers, and financial institutions. By analyzing historical information and identifying underlying patterns, forecasting models facilitate more reliable estimates of future market movements.

Although the Efficient Market Hypothesis argues that stock prices fully reflect all available information and are therefore inherently unpredictable, numerous empirical studies suggest that historical price patterns and macroeconomic variables can improve forecasting performance. Traditionally, stock market prediction has relied on fundamental and technical analysis. However, recent advances in data science and artificial intelligence have led to the widespread application of machine learning techniques, particularly deep learning models, in financial forecasting.

Among these approaches, recurrent neural networks (RNNs) and their advanced variants—including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks—have demonstrated remarkable success in modeling nonlinear relationships and capturing complex temporal dependencies in financial time series. Compared with conventional statistical models, these architectures are better equipped to learn both short-term and long-term patterns, thereby improving forecasting accuracy.

Despite the increasing international application of deep learning techniques in financial forecasting, empirical evidence for the Iranian capital market remains relatively limited. To address this gap, the present study evaluates and compares the predictive performance of three recurrent neural network models—RNN, LSTM, and GRU—in forecasting the Iran Basic Metals Industry Stock Index. The findings are expected to provide useful insights for investors, financial analysts, and policymakers seeking more reliable forecasting tools.

Method

Recurrent Neural Networks (RNNs), originally introduced in the 1980s, were designed to model sequential data by capturing temporal dependencies across observations. Through recurrent hidden layers, these networks retain information from previous time steps, making them suitable for time series forecasting. However, conventional RNNs often struggle to capture long-term dependencies because of the vanishing gradient problem, which limits their predictive performance.

To overcome this limitation, the Long Short-Term Memory (LSTM) architecture was developed. By incorporating dedicated memory cells and gating mechanisms, LSTM networks effectively retain and update relevant information over long sequences, enabling them to model complex temporal patterns while mitigating the vanishing gradient problem.

The Gated Recurrent Unit (GRU), introduced by Cho et al (2014), represents a simplified version of the LSTM architecture. GRU employs only two gates—reset and update gates—to regulate information flow, eliminating the need for a separate memory cell. This simpler architecture reduces computational complexity while maintaining strong predictive performance, making GRU particularly attractive for financial time series forecasting.

The empirical analysis uses daily observations of the Iran Basic Metals Industry Stock Index from April 2011 to March 2025. In addition to the stock index, several macroeconomic variables—including the exchange rate, global oil prices, money supply, consumer price index, and global gold prices—were incorporated as explanatory variables to improve forecasting accuracy.

Results

Three recurrent neural network models—RNN, LSTM, and GRU—were implemented to forecast the Iran Basic Metals Industry Stock Index.

The baseline architecture consisted of two recurrent hidden layers followed by an output layer, with dropout regularization applied after each recurrent layer to reduce overfitting. The dataset was divided into training (80%) and testing (20%) subsets to evaluate the models' generalization performance. Initial model parameters included the ReLU activation function, the Adam optimizer, the Mean Squared Error (MSE) loss function, 128 and 64 neurons in the first and second hidden layers, respectively, a batch size of 64, a window size of 10, and 200 training epochs.

Each model was subsequently optimized by adjusting the number of neurons, batch size, window size, and training epochs to minimize forecasting errors. The optimal LSTM model consisted of 128 neurons in the first hidden layer, 64 neurons in the second hidden layer, a batch size of 32, a window size of 20, and 200 training epochs. Similar optimization procedures were conducted for the RNN and GRU models.

The empirical results demonstrate that the GRU model achieved the highest forecasting accuracy, with an RMSE of 0.1161. The LSTM model ranked second, with an RMSE of 0.1167, while the conventional RNN model produced the weakest performance, with an RMSE of 0.1271. Similar rankings were observed for the MAE, MAPE, and Theil's U statistics. Overall, the findings indicate that GRU provides the most effective framework for capturing both short-term and long-term dependencies in financial time series, resulting in superior forecasting performance.

Conclusions

This study compared the forecasting performance of three recurrent neural network architectures—RNN, LSTM, and GRU—in predicting the Iran Basic Metals Industry Stock Index. Daily observations covering the period from April 2011 to March 2025, together with major macroeconomic indicators, including the exchange rate, global oil prices, money supply, consumer price index, and global gold prices, were used as model inputs.

The empirical findings reveal that the GRU model consistently outperformed both the LSTM and conventional RNN models across all forecasting accuracy measures. These results demonstrate the superior capability of GRU networks in modeling nonlinear relationships and temporal dependencies within financial time series. The findings are also consistent with previous studies, including those of Kuang (2023) and Kamalov (2020), which reported the superior predictive performance of GRU-based architectures.

Based on these findings, several policy implications emerge. First, integrating deep learning-based forecasting models into the analytical frameworks of regulatory authorities and capital market institutions can improve forecasting accuracy and support more informed decision-making. Second, strengthening financial data infrastructure and enhancing transparency through the development of comprehensive financial databases can improve information efficiency and increase investor confidence. Finally, greater investment in research and development related to artificial intelligence in finance—including collaboration among universities, financial institutions, and policymakers—can facilitate the development of localized forecasting models and contribute to the long-term stability and efficiency of Iran's capital market.

4/ Performance comparison of deep learning models for forecasting Iran's basic metals industry stock index

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors would like to thank all participants in the present study.

Ethical Considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

authors declare no conflict of interest.



انجمن علمی تجارت الکترونیکی ایران

مجله توسعه و سرمایه

شماره چاپ: ۲۰۰۸-۲۴۲۸ شماره انگیزشی: ۲۴۴۵-۳۶۰۶

Homepage: <https://jdc.uk.ac.ir>



دانشگاه شهید باهنر کرمان

مقایسه عملکرد الگوهای یادگیری عمیق در پیش‌بینی شاخص سهام صنعت فلزات اساسی ایران

پدرام محمدحسن^۱✉، مجتبی بهمنی^۲، مهران حسینی^۳ و صادق اسمعیلی^۴

۱. نویسنده مسئول، گروه اقتصاد، دانشکده اقتصاد، دانشگاه تهران، تهران، ایران. p.mohamadhasan@ut.ac.ir **رایانامه:**

۲. گروه اقتصاد، دانشکده اقتصاد و مدیریت، دانشگاه شهید باهنر کرمان، کرمان، ایران. mbahmani@uk.ac.ir **رایانامه:**

۳. گروه مهندسی صنایع، دانشکده مدیریت و مهندسی صنایع، دانشگاه صنعتی مالک اشتر، تهران، ایران. sm.hoseini@mut.ac.ir **رایانامه:**

۴. گروه اقتصاد، دانشکده اقتصاد، دانشگاه علامه طباطبائی، تهران، ایران. esmaeili.sadegh13@gmail.com **رایانامه:**

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی.	هدف: هدف این پژوهش، ارزیابی دقت الگوهای مختلف شبکه عصبی در پیش‌بینی شاخص سهام صنعت فلزات اساسی ایران است.
تاریخ‌ها: تاریخ دریافت: ۱۴۰۵/۱/۲۰ تاریخ بازنگری: ۱۴۰۵/۴/۱۶ تاریخ پذیرش: ۱۴۰۵/۴/۲۹ تاریخ انتشار برخط:	روش: برای این منظور، از داده‌های روزانه شاخص سهام صنعت فلزات اساسی طی دوره فروردین ۱۳۹۰ تا اسفند ۱۴۰۳، همراه با متغیرهای کلان اقتصادی شامل نرخ ارز، قیمت جهانی نفت، حجم نقدینگی، شاخص قیمت مصرف‌کننده و قیمت طلا استفاده شده است. در این مطالعه، سه مدل شبکه عصبی شامل شبکه عصبی بازگشتی ساده (RNN)، حافظه طولانی کوتاه‌مدت (LSTM) و واحد بازگشتی دروازه‌دار (GRU) به کار گرفته شده و دقت پیش‌بینی آن‌ها با معیارهای ریشه دوم میانگین مربعات خطا (RMSE)، میانگین قدرمطلق خطا (MAE)، میانگین درصد خطای مطلق (MAPE) و ضریب نابرابری تیل (Theil's U) ارزیابی گردیده است.
واژه‌های کلیدی: شاخص قیمت سهام، متغیرهای اقتصادی کلان، یادگیری ماشین، شبکه‌های عصبی بازگشتی.	یافته‌ها: نتایج تحقیق نشان می‌دهد مدل GRU بهترین عملکرد را در میان الگوهای مورد بررسی داشته است. این امر نشان می‌دهد که معماری GRU می‌تواند در پیش‌بینی داده‌های سری زمانی مالی، در شرایط خاص، نتایج دقیق‌تری نسبت به سایر الگوهای مانند LSTM و RNN ارائه دهد.
طبقه‌بندی JEL: C45, C53, E44.	نتیجه‌گیری: اهمیت یافته‌ها در آن است که پیش‌بینی دقیق شاخص‌های مالی می‌تواند ریسک تصمیم‌گیری را کاهش داده و کارایی بازار سرمایه را افزایش دهد. بر این اساس، توسعه ابزارهای پیش‌بینی مبتنی بر یادگیری ماشین و تقویت زیرساخت‌های داده‌ای می‌تواند به بهبود تصمیم‌گیری سیاست‌گذاران و ارتقای ثبات بازار سرمایه کمک کند.
استناد: محمدحسن، پدرام؛ بهمنی، مجتبی؛ حسینی، مهران و اسمعیلی، صادق (۱۴۰۵). مقایسه عملکرد الگوهای یادگیری عمیق در پیش‌بینی شاخص سهام صنعت فلزات اساسی ایران. <i>مجله توسعه و سرمایه</i> ، (*،*)، *.	
ناشر: دانشگاه شهید باهنر کرمان. © نویسنده/ نویسندگان.	



۱- مقدمه

بازارهای مالی در هر اقتصاد، به‌عنوان نهادهای سرمایه‌گذاری و واسطه‌های اصلی میان پس‌اندازکنندگان و واحدهای تولیدی عمل می‌کنند و نقشی حیاتی در تخصیص بهینه منابع دارند. در میان این بازارها، بازار سهام جایگاهی ویژه دارد و اغلب به‌عنوان یکی از مهم‌ترین شاخص‌های سلامت و پویایی اقتصادی یک کشور شناخته می‌شود. بازار سهام با جذب سرمایه‌های خرد و کلان، به توسعه کسب‌وکارها و افزایش تولید ملی کمک کرده و از این طریق نقش مهمی در رشد پایدار اقتصادی ایفا می‌کند. ارزش و نوسانات شاخص سهام نه تنها بازتابی از عملکرد جاری شرکت‌هاست، بلکه انتظارات سرمایه‌گذاران از رشد و تحولات آینده اقتصاد را نیز منعکس می‌کند. به‌طور کلی، نوسانات شاخص‌های مالی، تحت‌تأثیر مجموعه‌ای از عوامل قرار دارند؛ از جمله متغیرهای کلان اقتصادی (نظیر سیاست‌های پولی و مالی)، شرایط سیاسی کشور و جهان، ریسک‌های ژئوپلیتیک، انتظارات و احساسات سرمایه‌گذاران، ساختار رقابتی صنایع فعال بورسی، و کیفیت مدیریت و عملکرد شرکت‌ها (کاویانی و همکاران، ۱۴۰۴).

پیش‌بینی شاخص‌ها و متغیرهای اقتصادی در بازارهای مالی، ابزاری ارزشمند برای کاهش ریسک تصمیم‌گیری و افزایش سودآوری شرکت‌ها و سرمایه‌گذاران خصوصی است. مدیران و سیاستگذاران بخش‌های مختلف اقتصادی و مالی، به‌منظور دستیابی به این اهداف، همواره به دنبال روش‌هایی هستند که پیش‌بینی‌های آنان را به واقعیت نزدیک‌تر، و خطاهای آنان را کمتر کند. پیش‌بینی شاخص‌های بازار سهام، فرآیندی است که با استفاده از داده‌های گذشته، سعی در تخمین شاخص، در آینده دارد. این فرآیند با استفاده از الگوهای ریاضی، الگوهای رفتاری شاخص‌های مالی را کشف می‌کند. هرچه مدل مورد استفاده، بهتر باشد و بتواند الگوی موجود در داده‌ها را بهتر کشف کند، پیش‌بینی‌هایش دقیق‌تر خواهد بود. سرمایه‌گذاری در بورس از ریسک و نااطمینانی بالایی برخوردار است که نتیجه‌ی دخالت متغیرهای اقتصادی و غیر اقتصادی فراوان، بر رفتار متغیرهای این بازار است. در این شرایط پر ریسک و نااطمینان، نیاز به استفاده از الگوهای دقیق به منظور پیش‌بینی شاخص‌های بازار سرمایه احساس می‌شود؛ زیرا با این کار، ریسک سرمایه‌گذاری کاهش می‌یابد و فعالان اقتصادی با اطمینان بیشتری به سرمایه‌گذاری در این بازار می‌پردازند. در این صورت است که بازار سرمایه به مکانی برای رشد و توسعه اقتصاد کشور تبدیل می‌شود (نبی‌پور و همکاران، ۱۴۰۰).

پیش‌بینی شاخص‌های بازار سهام یکی از موضوعات محوری و چالش‌برانگیز در حوزه مالی است که همواره توجه پژوهشگران، تحلیل‌گران و فعالان بازار را به خود جلب کرده است. بر اساس فرضیه بازار کارا، پیش‌بینی دقیق مقادیر آتی سهام غیرممکن دانسته می‌شود، زیرا قیمت‌ها رفتاری تصادفی دارند. با این حال، شواهد تجربی و تحلیل‌های فنی نشان می‌دهند که بخش قابل توجهی از رفتار قیمتی سهام در داده‌های تاریخی نهفته است و بررسی الگوها و روندها می‌تواند به بهبود فرایند پیش‌بینی کمک کند. پژوهش‌های موجود در حوزه بازارهای مالی، در دو مسیر اصلی قرار می‌گیرند: ارزیابی عملکرد بازار، و توسعه الگوهای پیش‌بینی. رویکردهای سنتی عمدتاً بر تحلیل بنیادی و تحلیل تکنیکال تمرکز داشته‌اند. تحلیل بنیادی به بررسی ارزش ذاتی شرکت‌ها از طریق متغیرهایی مانند درآمد، سودآوری و ارزش منصفانه می‌پردازد، در حالی که تحلیل تکنیکال بر الگوهای قیمتی، حجم معاملات و روندهای بازار اتکا دارد. با پیشرفت علوم داده و گسترش روش‌های یادگیری ماشین، این تکنیک‌ها به‌طور فزاینده‌ای در پیش‌بینی بازار سهام مورد استفاده قرار گرفته‌اند. ویژگی‌های استخراج‌شده از شاخص‌های بنیادی و تکنیکال می‌توانند ورودی مناسبی برای الگوهای یادگیری ماشین و یادگیری عمیق فراهم کنند. در ابتدا، الگوهای خطی مانند الگوی

خودرگسیون با میانگین متحرک معرفی شدند، اما به مرور، الگوهای پیچیده‌تری همچون شبکه‌های عصبی مصنوعی توسعه یافتند. در سال‌های اخیر، یادگیری عمیق به جایگاه ویژه‌ای در این حوزه دست یافته است. الگوهایی مانند الگوهای شبکه عصبی بازگشتی توانسته‌اند با استخراج الگوهای غیرخطی و روابط زمانی پیچیده، دقت پیش‌بینی سری‌های زمانی مالی را به‌طور قابل توجهی بهبود دهند. این الگوها اکنون به‌عنوان ابزارهای قدرتمند در پیش‌بینی بازار سهام و سایر بازارهای مالی شناخته می‌شوند. با وجود افزایش محبوبیت جهانی شبکه‌های عصبی مصنوعی، در ایران پژوهش‌های اندکی در زمینه استفاده از الگوهای یادگیری ماشین برای پیش‌بینی متغیرهای مالی انجام شده است. یادگیری ماشین می‌تواند ابزاری ارزشمند برای سرمایه‌گذاران در این بازار باشد (آتسالاکیس و والاونیس^۱، ۲۰۰۹ و براون و همکاران^۲، ۲۰۱۳).

صنعت فلزات اساسی یکی از بخش‌های حیاتی تولیدی در اقتصاد است و به دلیل تولید محصولات با کاربرد گسترده در بخش‌های مختلف اقتصادی، نقشی کلیدی و واسطه‌ای در رشد و توسعه اقتصادی ایفا می‌کند. فلزات اساسی به عنوان کالاهای واسطه‌ای در تولید محصولات مانند خودرو، لوازم خانگی، ساختمان، تجهیزات صنعتی و سایر صنایع مورد استفاده قرار می‌گیرند. بنابراین، توسعه و افزایش ظرفیت تولید این صنعت می‌تواند تأثیر مستقیم و قابل توجهی بر رشد سایر بخش‌های اقتصاد داشته باشد. توسعه روزافزون صنعت و پیشرفت‌های فناوری موجب افزایش تقاضای جهانی برای فلزات اساسی شده و چشم‌انداز روشنی را برای بازار فلزات اساسی در ایران و جهان ترسیم کرده است. بسیاری از کشورها با هدف کسب جایگاه رقابتی در اقتصاد جهانی، تلاش می‌کنند صنعت فلزات اساسی خود را تقویت کنند. ایران نیز با برخورداری از منابع معدنی غنی و نیروی کار ماهر، پتانسیل بالایی برای توسعه این صنعت دارد. با این حال، تحقق این پتانسیل مستلزم تأمین مالی مناسب و پایدار است. بازار سرمایه می‌تواند نقشی کلیدی در این زمینه ایفا کند؛ با جذب سرمایه‌های خرد و کلان، امکان تأمین مالی مورد نیاز شرکت‌های فعال در صنعت فلزات اساسی فراهم می‌شود. این امر به شرکت‌ها اجازه می‌دهد تا ظرفیت تولید خود را افزایش دهند، کیفیت محصولات را ارتقا دهند و در نهایت، سهم بیشتری از بازار داخلی و بین‌المللی کسب کنند (عسگری، ۱۳۹۹ و صمدی و همکاران، ۱۳۸۸).

بر این اساس، هدف اصلی پژوهش حاضر، ارزیابی دقت الگوهای مختلف شبکه عصبی در پیش‌بینی شاخص قیمت سهام صنعت فلزات اساسی ایران است. برای دستیابی به این هدف، از داده‌های روزانه شاخص قیمت سهام صنعت فلزات اساسی طی دوره فروردین ۱۳۹۰ تا اسفند ۱۴۰۳ استفاده شده و متغیرهای کلان اقتصادی اثرگذار بر عملکرد بازار سهام شامل نرخ ارز، قیمت جهانی نفت، حجم نقدینگی، شاخص قیمت مصرف‌کننده و قیمت جهانی طلا به‌عنوان متغیرهای توضیحی در مدل لحاظ شده‌اند. در این مطالعه، سه مدل پرکاربرد شبکه عصبی بازگشتی شامل شبکه عصبی بازگشتی ساده (RNN)، شبکه حافظه طولانی کوتاه‌مدت (LSTM) و شبکه بازگشتی دروازه‌دار (GRU) به‌منظور پیش‌بینی و مقایسه عملکرد، مورد استفاده قرار گرفته‌اند.

ساختار مقاله به‌اختصار چنین است: در بخش دوم، مبانی نظری بررسی می‌گردد؛ سپس در بخش سوم، پیشینه ادبیات داخلی و خارجی مرور می‌شود؛ در بخش چهارم، داده‌ها و متغیرهای پژوهش، و همچنین شبکه‌های عصبی مورد استفاده به منظور پاسخ به سوال پژوهش، معرفی می‌گردند؛ بخش پنجم، به گزارش و تحلیل نتایج اختصاص دارد؛ و در بخش ششم، جمع‌بندی و پیشنهادات سیاستی ارائه می‌گردد.

^۱ Atsalakis & Valavanis^۲ Brown

۲- مبانی نظری پژوهش

دیدگاه ارزش فعلی، مبنای اصلی تحلیل بازار سهام است و بیان می‌کند که قیمت سهام، برابر ارزش فعلی جریان‌های نقدی آتی شرکت است که با نرخ تنزیل، تنزیل می‌شود. در چارچوب شهودی مدل گوردون^۱، هر عاملی که نرخ تنزیل را افزایش دهد یا رشد مورد انتظار سود را کاهش دهد، ارزش ذاتی سهام و در نتیجه شاخص قیمت را پایین می‌آورد. بنابراین، متغیرهای کلان از سه مسیر بر شاخص قیمت سهام اثر می‌گذارند: نرخ تنزیل، جریان‌های نقدی بنگاه و جانشینی سبد دارایی‌ها. در ادامه، سازوکار اثرگذاری هر یک از متغیرهای اقتصادی کلان بررسی می‌شود (ورما و بانسال^۲، ۲۰۲۱).

نرخ بهره، هزینه استقراض پول یا بازده سرمایه‌گذاری است. افزایش آن با بالا بردن نرخ تنزیل، ارزش حال جریان‌های نقدی را کاهش می‌دهد و هم‌زمان هزینه تأمین مالی بنگاه‌ها را افزایش داده و حاشیه سودشان را محدود می‌کند. همچنین با جذاب‌تر شدن اوراق بدهی و سپرده‌ها، سرمایه از بازار سهام به دارایی‌های کم‌ریسک‌تر منتقل می‌شود و اثر جانشینی به زیان بازار سهام عمل می‌کند (لی و براهماسرنه^۳، ۲۰۱۸ و انور و همکاران^۴، ۲۰۱۹).

رابطه تورم و بازار سهام یکی از مباحث چالش‌برانگیز اقتصاد مالی است. دو دیدگاه نظری، مطرح می‌شود: اثر فشار که سهام را محافظتی در برابر تورم می‌داند و معتقد است افزایش قیمت‌ها، به درآمد اسمی و ارزش دارایی‌های واقعی شرکت منتقل می‌شود؛ و در مقابل، نظریه توهم پولی و نااطمینانی تورمی، که بیان می‌کند تورم، به‌ویژه نوع پیش‌بینی‌نشده آن، پریمیوم ریسک را افزایش داده و ارزش حال سودهای آتی شرکت را کاهش می‌دهد و در نتیجه، اثری منفی بر قیمت سهام دارد. این اثر در سطح صنایع، متفاوت است؛ صنایع دارایی‌محور با قدرت قیمت‌گذاری، ممکن است بهره‌مند شوند، اما صنایع کاربر، بیشتر در معرض آسیب هستند (برکات و همکاران^۵، ۲۰۱۵ و مگاراولی و سامپاگنارو^۶، ۲۰۱۸).

نرخ ارز از مسیرهای گوناگون بر شاخص قیمت سهام اثر می‌گذارد. طبق الگوی دورنبوش فیشر، تضعیف پول ملی با تقویت رقابت‌پذیری صادرات، درآمد و سود شرکت‌های صادرات‌محور را افزایش داده و قیمت سهام آن‌ها را بالا می‌برد؛ اما برای شرکت‌های واردات‌محور، فشار هزینه ایجاد می‌کند. بنابراین، اثر نهایی نرخ ارز بر بازار سهام، تابع میانگین مواجهه ارزی صنایع مختلف است (هاید^۷، ۲۰۰۷؛ داهیر و همکاران^۸، ۲۰۱۸ و بالونزادنوری و شهیدی، ۱۴۰۴).

از دیدگاه نظری، افزایش نقدینگی از طریق کاهش نرخ‌های کوتاه‌مدت، گسترش اعتبار و تعدیل ترکیب سبد دارایی، می‌تواند به رشد قیمت دارایی‌ها منجر شود. در کانال اعتباری سیاست پولی، سهام شرکت‌های با محدودیت مالی بالا، نسبت به شوک‌های نقدینگی، حساس‌ترند. با این حال، رشد سریع نقدینگی در صورت افزایش انتظارات تورمی و تضعیف نرخ واقعی، می‌تواند موجب خروج سرمایه از بازار سهام و همچنین افزایش پریمیوم ریسک شود. از این رو، رابطه نقدینگی و بازار سهام لزوماً خطی و یکنواخت نیست (تیریکی و همکاران^۹، ۲۰۱۹ و اتاله و ازه^{۱۰}، ۲۰۱۹).

به‌طور کلی، قیمت جهانی طلا به‌عنوان دارایی امن، نقش مهمی در رفتار سرمایه‌گذاران دارد. افزایش قیمت طلا معمولاً نشان‌دهنده رشد عدم‌اطمینان اقتصادی است و می‌تواند سرمایه‌ها را از بازار سهام به سوی خود جذب کند. در چنین شرایطی، کاهش اشتها برای ریسک، موجب افت تقاضا برای سهام و فشار قابل توجهی بر بازدهی بازار سهام می‌شود. از سوی

^۱ Gordon Model^۲ Verma & Bansal^۳ Lee & Brahmaasrene^۴ Anwer^۵ Barakat^۶ Megaravalli & Sampagnaro^۷ Hyde^۸ Dahir^۹ Tiryaki^{۱۰} Etale & Eze

دیگر، ثبات یا کاهش قیمت طلا، می تواند سرمایه گذاران را به سمت دارایی های ریسکی تر هدایت کرده و بر بهبود عملکرد بورس اثر گذار باشد (سینگ و شارما^۱، ۲۰۱۸ و الامیر و همکاران^۲، ۲۰۱۸).

در اقتصادهای صادرکننده نفت، این کالا هم نهادهای اساسی در تولید، و هم منبع اصلی درآمدهای ارزی است؛ از این رو، اثر نوسانات قیمت نفت بر بازار سهام به ماهیت شوک بستگی دارد. شوک های منفی عرضه، با افزایش هزینه انرژی، حاشیه سود بنگاه های غیرانرژی را کاهش داده و معمولاً شاخص سهام را تضعیف می کنند. در مقابل، شوک های مثبت تقاضای جهانی که بیانگر رونق اقتصادی اند، می توانند ضمن افزایش قیمت نفت، تقاضا برای کالاهای صادراتی را نیز تقویت کرده و اثری خنثی یا حتی مثبت بر بازار سهام داشته باشند. اما در شرایط محدودیت صادرات یا تحریم، افزایش قیمت جهانی نفت الزاماً به درآمد ارزی تبدیل نشده و به دلیل تضعیف کانال مالی، اثر نهایی آن بر بازار سهام می تواند مبهم باشد (دلگادو و همکاران^۳، ۲۰۱۸ و جین و بیسوال^۴، ۲۰۱۶).

علاوه بر مبانی ارزش فعلی، درک رفتار شاخص های بازار سهام مستلزم بررسی دو دیدگاه نظری کلیدی دیگر نیز است که بر چگونگی شکل گیری قیمت ها تاکید دارند. نظریه بازار کارا^۵ بیان می کند که در بازارهای کارا، قیمت ها همواره تمامی اطلاعات موجود را به طور کامل منعکس می کنند؛ لذا به دست آوردن بازدهی مازاد بر ریسک بازار از طریق تحلیل های فنی یا بنیادی ممکن نیست. بر اساس این نظریه، قیمت های سهام از گام تصادفی پیروی می کنند و تغییرات قیمت تنها در واکنش به اطلاعات و اخبار غیرمنتظره رخ می دهد (مالکیل^۶، ۲۰۰۳). با این حال، شواهد تجربی در بسیاری از بازارهای نوظهور نشان داده است که قیمت ها همواره به طور کامل کارا نیستند و الگوهای قابل شناسایی وجود دارند که الگوهای یادگیری عمیق در این پژوهش، درصدد استخراج آن ها هستند.

در مقابل، نظریه مالی رفتاری^۷ با نقد فروض عقلانیت کامل نظریه های کلاسیک، بر محدودیت های شناختی و سوگیری های روانی سرمایه گذاران تمرکز دارد. این نظریه معتقد است سرمایه گذاران تحت تاثیر عواملی همچون بیش اعتمادی، زیان گریزی و رفتار گله ای قرار دارند که منجر به واکنش های بیش از حد یا کمتر از حد لازم به اطلاعات شده و باعث انحراف قیمت از ارزش ذاتی می شود. افراد در شرایط عدم قطعیت، به جای حداکثرسازی مطلوبیت مورد انتظار، از میان برهای ذهنی استفاده می کنند که می تواند به خطاهای سیستماتیک منجر شود (شیلر^۸، ۲۰۱۷). در بازار سهام ایران، به دلیل نوسانات متغیرهای کلان و نااطمینانی های سیاسی، رفتارهای توده ای و واکنش های احساسی سرمایه گذاران نقش پررنگی در تغییرات شاخص دارد. در واقع، الگوهای یادگیری عمیق به دلیل توانایی در مدل سازی روابط غیرخطی و پیچیده، ابزاری قدرتمند برای پیش بینی قیمت در محیط های پیچیده مالی محسوب می شوند.

۳- پیشینه تجربی پژوهش

۳-۱- مطالعات داخلی

تسلیم پور و همکاران (۱۴۰۳) به پیش بینی شاخص کل بورس اوراق بهادار تهران، با بهره گیری از مدل شبکه عصبی خودرگرسیون غیرخطی با ورودی های برونزا، و مدل خودرگرسیون با میانگین متحرک پرداختند. داده های روزانه شاخص کل بورس، طی سال های ۱۳۸۸ تا ۱۴۰۲ برای آموزش و ارزیابی الگوها به کار گرفته شد. نتایج پژوهش نشان داد

¹ Singh & Sharma

² Al-Ameer

³ Delgado

⁴ Jain & Biswal

⁵ Efficient Market Hypothesis

⁶ Malkiel

⁷ Behavioral Finance Theory

⁸ Shiller

مدل شبکه عصبی خودرگرسیون غیرخطی با ورودی‌های برون‌زا، با تعیین وقفه‌های بهینه و معماری مناسب، عملکرد بهتری نسبت به مدل خودرگرسیون با میانگین متحرک دارد و خطای پیش‌بینی آن، به‌طور قابل توجهی کاهش یافته است. همچنین در میان الگوریتم‌های آموزشی، الگوریتم لونبرگ-مارکوات بالاترین دقت را ایجاد کرده است.

حیدری و امیری (۱۴۰۲) به بررسی و مقایسه الگوهای مختلف در حوزه یادگیری ماشین به منظور پیش‌بینی قیمت سهام در بورس اوراق بهادار تهران پرداخته‌اند. در این مطالعه، از داده‌های سری زمانی ۱۵۰ شرکت پذیرفته شده در بورس اوراق بهادار تهران، طی سال‌های ۱۳۹۰ تا ۱۳۹۹ به صورت روزانه استفاده شده است. الگوهای به کار رفته به منظور پیش‌بینی قیمت سهام، شامل الگوهای خودرگرسیون با میانگین متحرک، جنگل تصادفی، و حافظه طولانی کوتاه‌مدت بوده است. نتایج به دست آمده نشان می‌دهد پیش‌بینی‌های شبکه عصبی حافظه طولانی کوتاه‌مدت از دقت بالاتری نسبت به سایر الگوها برخوردار است.

اسدی و همکاران (۱۴۰۱) در مطالعه‌ای عوامل موثر بر پیش‌بینی سود شرکت‌های حاضر در بورس اوراق بهادار تهران را با بهره‌گیری از الگوهای یادگیری ماشین بررسی کردند. در این پژوهش، از الگوهای شبکه عصبی و الگوهای آماری کلاسیک، به همراه داده‌های سال‌های ۱۳۹۳ تا ۱۳۹۸ استفاده شده است. نتایج تحلیل‌ها نشان داد نتایج الگوهای شبکه عصبی، نسبت به سایر روش‌ها عملکرد دقیق‌تری داشته و توانسته خطای پیش‌بینی سود شرکت‌ها را به‌طور معناداری کاهش دهد.

شریف‌فر و همکاران (۱۴۰۱) با هدف پیش‌بینی قیمت سهام، از الگوریتم‌های مبتنی بر شبکه عصبی پیچشی بهره گرفته‌اند. در این پژوهش، داده‌های روزانه قیمت سهام همراه با ده شاخص منتخب تکنیکال مربوط به سهام شرکت ذوب آهن اصفهان، به عنوان ورودی مدل استفاده شد، و چندین معماری مختلف شبکه عصبی پیچشی مورد ارزیابی قرار گرفت. نتایج تحقیق نشان داد که معماری شامل لایه ادغام بیشینه (پارامترها: اندازه دسته ۶۴، تعداد فیلتر ۲۵۶، و تابع فعال‌سازی ReLU)، کمترین میزان خطا را ایجاد کرده است. این یافته‌ها بیانگر برتری معماری مذکور نسبت به سایر ساختارهای شبکه عصبی پیچشی در پیش‌بینی قیمت سهام است.

سیف و همکاران (۱۴۰۰) به پیش‌بینی شاخص کل بورس اوراق بهادار تهران، با استفاده از نوسان‌نمای موج الویت و شاخص قدرت نسبی پرداخته‌اند. در این تحقیق، از داده‌های سری زمانی شاخص کل بورس برای سال‌های ۱۳۸۷ تا ۱۳۹۹ به صورت روزانه استفاده شده است. همچنین از روش‌های مبتنی بر یادگیری ماشین، شامل سه الگوریتم درخت تصمیم، بیز ساده، و ماشین بردار پشتیبان استفاده شد. نتایج پژوهش نشان داد که در شاخص بورس اوراق بهادار تهران، شناسایی امواج الیوت امکان‌پذیر است؛ همچنین الگوریتم‌های ماشین بردار پشتیبان و درخت تصمیم، قادرند شاخص کل را با دقت خوبی پیش‌بینی کنند.

یراقی و ربیعی (۱۳۹۹) در مطالعه‌ای به بررسی و مقایسه عملکرد دو شبکه عصبی بازگشتی حافظه طولانی کوتاه‌مدت، و واحد بازگشتی دروازه‌دار، به منظور پیش‌بینی نرخ ارز در ایران، پرداخته‌اند. داده‌های مورد استفاده در این تحقیق، مربوط به داده‌های روزانه سال‌های ۱۳۷۱ تا ۱۳۹۶ است. یافته‌های پژوهش نشان داد واحد بازگشتی دروازه‌دار، از نظر دقت و کارایی محاسباتی، عملکرد بهتری نسبت به شبکه حافظه طولانی کوتاه‌مدت دارد. واحد بازگشتی دروازه‌دار، در کاهش شاخص‌های خطا، از جمله ریشه میانگین مربعات خطا و میانگین قدر مطلق خطا، نتایج مطلوب‌تری (۲۰ درصد کاهش خطا، نسبت به شبکه حافظه طولانی کوتاه‌مدت) ارائه داد.

فخاری و همکاران (۱۳۹۶) در پژوهشی به بررسی عملکرد شبکه عصبی بیزین و الگوریتم لونبرگ-مارکوات در مقایسه با الگوهای کلاسیک در پیش‌بینی قیمت سهام (مدل خودرگرسیون با میانگین متحرک)، پرداخته‌اند. در این تحقیق

از داده‌های شاخص قیمت روزانه برای سال‌های ۱۳۹۰ تا ۱۳۹۳ استفاده شده است. نتایج حاصل از استفاده از شبکه عصبی بیزین نشان داد این شبکه، خطای کمتر و قدرت پیش‌بینی بیشتری نسبت به مدل خودرگرسیون با میانگین متحرک دارد.

۳-۲- مطالعات خارجی

کوانگ^۱ (۲۰۲۳) در پژوهشی به مقایسه روش‌های مختلف پیش‌بینی قیمت سهام در بورس انگلستان پرداخته است. در این مطالعه از داده‌های سال‌های ۲۰۱۰ تا ۲۰۱۹ استفاده شده است. همچنین از روش‌های پیش‌بینی، شامل رگرسیون خطی، خودرگرسیون با میانگین متحرک، و شبکه عصبی حافظه طولانی کوتاه‌مدت استفاده شد. نتایج به‌دست آمده در این پژوهش حاکی از آن است که میزان دقت پیش‌بینی قیمت سهام توسط شبکه عصبی حافظه طولانی کوتاه‌مدت، دارای کمترین خطا و روش رگرسیون خطی، دارای بیشترین خطا بوده است.

چن و همکاران^۲ (۲۰۲۲) به منظور پیش‌بینی شاخص سهام صنعت بانکداری در کشور چین، از روش یادگیری عمیق استفاده کردند. بدین منظور ابتدا از الگوریتم خوشه‌بندی k-means استفاده شد تا بانک‌هایی که رفتار قیمتی مشابهی با یکدیگر دارند، در گروه‌های مشابه قرار گیرند. سپس برای پیش‌بینی ایستا و پویای قیمت سهام، از شبکه عصبی حافظه طولانی کوتاه‌مدت به منظور آموزش مدل استفاده شد. نتایج این تحقیق نشان می‌دهد الگوهای یادگیری عمیق، عملکرد و دقت بسیار بالایی در پیش‌بینی متغیرهای مالی مانند قیمت سهام دارند. همچنین در این مطالعه مشخص شد عملکرد الگوهای پویا از الگوهای ایستا بهتر بوده است.

گویی و یون وو^۳ (۲۰۲۲) به مقایسه روش‌های پیش‌بینی رایج در اقتصادسنجی سری‌های زمانی، با روش‌های مبتنی بر یادگیری ماشین، پرداخته‌اند. بدین منظور از داده‌های شاخص قیمت سهام صنعت دارو در سال‌های ۲۰۲۰ تا ۲۰۲۱ به صورت روزانه استفاده شده است. در این تحقیق، از الگوهای اقتصادسنجی شامل رگرسیون چند متغیره، و خودرگرسیون با میانگین متحرک؛ و همچنین الگوهای شبکه عصبی، شامل مدل پس‌انتشار خطا، و الگوریتم جنگل تصادفی استفاده شده است. نتایج به‌دست آمده حاکی از آن است که از بین الگوهای استفاده شده به منظور پیش‌بینی قیمت سهام، مدل پس‌انتشار خطا از عملکرد و دقت بهتری برخوردار بوده است.

کمالو^۴ (۲۰۲۰) در مطالعه‌ای به بررسی و پیش‌بینی شاخص قیمت سهام با استفاده از روش‌های یادگیری ماشین، پرداخته است. بدین منظور از سه مدل اصلی در شبکه‌های عصبی، شامل پرسپترون چندلایه، شبکه عصبی پیچشی و شبکه عصبی حافظه طولانی کوتاه‌مدت استفاده شده است. داده‌های مورد استفاده در این پژوهش، داده‌های روزانه ۴ شرکت بزرگ کشور آمریکا به مدت ۱۰ سال است. نتیجه این پژوهش آن است که روش شبکه عصبی حافظه طولانی کوتاه‌مدت، از دقت بسیار بالایی برخوردار بوده و برای مدل‌سازی‌های مالی از درجه اطمینان بالایی برخوردار است.

ویدال و کریست جانپولر^۵ (۲۰۲۰) به بررسی و پیش‌بینی نوسانات قیمت طلا با استفاده از روش‌های یادگیری ماشین، پرداخته‌اند. بدین منظور در این پژوهش به‌طور همزمان از دو روش یادگیری عمیق، شامل شبکه عصبی حافظه طولانی کوتاه‌مدت و شبکه عصبی پیچشی استفاده شده است. نتایج به دست آمده حاکی از پیشرفت چشمگیر در میزان دقت پیش‌بینی‌ها بوده است؛ به عنوان مثال، نتایج، بهبود ۳۷ درصدی در شاخص میانگین مربعات جملات خطا را نسبت به مدل

¹ Kuang

² Chen

³ Gui & Wu

⁴ Kamalov

⁵ Vidal & Krist janpoller

خودرگرسین شرطی تعمیم‌یافته ثبت کرده است.

گالشچوک^۱ (۲۰۱۶) به پیش‌بینی نرخ ارز به کمک الگوهای شبکه عصبی مصنوعی پرداخته است. در این مطالعه، از نرخ‌های ارز دلار/یورو، ین/دلار، دلار/پوند برای دوره زمانی ژانویه ۲۰۱۴ تا آوریل ۲۰۱۴ میلادی به صورت روزانه استفاده شده است. همچنین از شبکه عصبی پرسپترون چندلایه به منظور پیش‌بینی‌های کوتاه‌مدت و یک قدم رو به جلو استفاده شده است. نتایج به دست آمده، نشان می‌دهد که شبکه عصبی پرسپترون چندلایه از دقت بالایی در خصوص پیش‌بینی‌های کوتاه‌مدت برخوردار است.

۳-۳- جمع‌بندی ادبیات، ضرورت و نوآوری پژوهش

مرور ادبیات پژوهش نشان می‌دهد که بخش قابل توجهی از مطالعات داخلی و خارجی، به پیش‌بینی قیمت یا شاخص سهام با استفاده از روش‌های کلاسیک اقتصادسنجی یا الگوهای منفرد یادگیری ماشین پرداخته‌اند. در بسیاری از پژوهش‌ها، تمرکز اصلی بر تحلیل‌های تک‌متغیره یا همراه با متغیرهای محدود اقتصادی بوده و نقش هم‌زمان متغیرهای کلان اقتصادی در چارچوب الگوهای یادگیری عمیق، به‌ویژه برای بازار سهام ایران، کمتر مورد توجه قرار گرفته است. همچنین اغلب مطالعات داخلی، یا از داده‌های دوره‌های زمانی کوتاه استفاده کرده‌اند، یا به مقایسه محدود چند مدل پرداخته و ارزیابی جامعی از شبکه‌های عصبی بازگشتی پیشرفته ارائه نکرده‌اند. نتایج برخی مطالعات در خصوص برتری نسبی الگوهای RNN، LSTM و GRU نیز یکسان نبوده و شواهد تجربی در این زمینه، به‌ویژه برای اقتصادهای در حال توسعه، همچنان پراکنده و گاه متناقض است.

با توجه به اهمیت فزاینده پیش‌بینی دقیق شاخص‌های مالی در کاهش ریسک سرمایه‌گذاری و بهبود کارایی بازار سرمایه، ضرورت انجام پژوهشی جامع با بهره‌گیری از داده‌های به‌روز و روش‌های پیشرفته یادگیری ماشین احساس می‌شود. پژوهش حاضر با استفاده از داده‌های روزانه متغیرهای کلان اقتصادی طی یک بازه زمانی بلندمدت، تلاش می‌کند این خلأ علمی را پر کند. نوآوری‌های اصلی این پژوهش، در به کارگیری هم‌زمان مجموعه‌ای از متغیرهای کلان اقتصادی اثرگذار، همراه با مقایسه سه مدل پیشرفته شبکه عصبی بازگشتی شامل RNN، LSTM و GRU برای پیش‌بینی شاخص قیمت سهام صنعت فلزات اساسی است. در این مطالعه، به جای پیش‌بینی شاخص کل بازار سهام، تمرکز بر پیش‌بینی شاخص قیمت سهام یک صنعت خاص (فلزات اساسی) در ایران دارد؛ چرا که صنایع مختلف، اثرپذیری متفاوتی از متغیرهای کلان اقتصادی دارند و پیش‌بینی در سطح صنعت می‌تواند بینش دقیق‌تری ارائه کند. استفاده از داده‌های روزانه (طی یک بازه زمانی بلندمدت)، ارزیابی عملکرد الگوها با معیارهای خطای معتبر، و تمرکز بر مقایسه دقت پیش‌بینی شبکه‌های بازگشتی، این پژوهش را از مطالعات پیشین متمایز می‌سازد. نتایج این تحقیق می‌تواند ضمن غنای ادبیات علمی، ابزار تحلیلی مناسبی برای سرمایه‌گذاران، تحلیل‌گران و سیاست‌گذاران بازار سرمایه فراهم کرده و به بهبود تصمیم‌گیری و افزایش ثبات بازار سهام کمک کند.

۴- روش‌شناسی پژوهش

شبکه‌های عصبی بازگشتی^۲، طی چند دهه گذشته یکی از پایه‌های اصلی پژوهش در یادگیری عمیق بوده‌اند. مفهوم شبکه عصبی بازگشتی، نخستین بار در دهه ۱۹۸۰ توسط دیوید روملهارت^۳، جفری هینتن^۴ و رونالد ویلیامز^۵ معرفی شد. آن‌ها استفاده از اتصالات بازگشتی را برای مدل‌سازی روابط زمانی در داده‌ها پیشنهاد کردند. الگوهای RNN برای داده‌های

^۱ Galeshchuk

^۲ Recurrent Neural Network

^۳ David Rumelhart

^۴ Geoffrey Hinton

^۵ Ronald Williams

چند متغیره‌ی سری زمانی مناسب بوده و قادرند وابستگی‌های زمانی در دوره‌های مختلف را ثبت کنند. در یک RNN، مقادیر ورودی، وارد شبکه می‌شوند و خروجی‌ها بر اساس این ورودی‌ها و پارامترهای شبکه تولید می‌گردند. الگوریتم پس‌انتشار خطا^۱ را می‌توان برای آموزش شبکه‌های بازگشتی به کار برد؛ این کار با باز کردن شبکه در طول زمان و محدود کردن برخی اتصالات برای ثابت نگه داشتن وزن‌ها در همه‌ی گام‌های زمانی انجام می‌شود. بنابراین، RNN‌ها قادرند تا اطلاعات پیشین را در خلال فرآیند یادگیری به لایه‌ی پنهان فعلی منتقل کنند و همین امر، یادگیری مؤثر داده‌های سری زمانی را امکان‌پذیر می‌سازد. در شبکه‌های عصبی بازگشتی ساده (RNNs)، خروجی y_t به صورت زیر محاسبه می‌گردد:

$$y_t = f(W_y h_t) \quad (۱)$$

$$h_t = \sigma(W_h h_{t-1} + W_x x_t) \quad (۲)$$

در اینجا، W_y ، W_h و W_x ماتریس‌های مربوط به خروجی لایه‌ی پنهان h_t ، فعالیت لایه‌ی پنهان قبلی h_{t-1} و ورودی x_t هستند. «بازگشت» در معادله‌ی (۲)، فعالیت لایه‌ی پنهان فعلی h_t را به فعالیت لایه‌ی پنهان قبلی h_{t-1} مرتبط می‌کند. این وابستگی به دلیل استفاده از تابع سیگموید $\sigma(\cdot)$ غیرخطی است (هانسن و راستمی^۲، ۲۰۱۹).

به صورت نظری، شبکه‌های RNN باید قادر به یادگیری دنباله‌هایی با هر میزان پیچیدگی باشند. با این حال، مشاهدات تجربی نشان می‌دهد که این شبکه‌ها در حفظ اطلاعات مرتبط از ورودی‌های گذشته در بازه‌های زمانی طولانی با مشکل مواجه‌اند. این ناتوانی در ذخیره‌سازی و حفظ اطلاعات گذشته، منجر به ناپایداری در هنگام تولید دنباله‌ها می‌شود. برطرف کردن این ناپایداری، مستلزم داشتن ظرفیت حافظه بلندمدت است؛ زیرا شبکه‌ای با حافظه‌ی گسترده، می‌تواند حتی با وجود فاصله‌ی زمانی زیاد بین رویدادهای مهم، همچنان از ورودی‌ها یاد بگیرد. این نیاز، منجر به معرفی مدل حافظه‌ی طولانی کوتاه‌مدت^۳ شد؛ مدلی که نوعی RNN با توانایی‌های ذخیره‌سازی و دسترسی پیشرفته‌تر نسبت به نسخه‌های سنتی محسوب می‌شود. در نتیجه، LSTM‌ها می‌توانند ویژگی‌های مهم را در دنباله‌ی ورودی در مراحل اولیه تشخیص داده و آن‌ها را در یک مسیر طولانی منتقل کنند، در حالی که وابستگی‌های بلندمدت را در طول فرآیند، حفظ می‌کنند (چن و همکاران^۴، ۲۰۲۲).

هوخرایتر^۵ و اشمیدهور^۶ نخستین پژوهشگرانی بودند که مدل LSTM را معرفی کردند؛ مدلی که بعداً در پژوهش‌های گریوز^۷ توسعه و بهبود یافت. LSTM که به عنوان جایگزینی اصلی برای RNN‌های سنتی طراحی شد، مشکلاتی مانند ناپدید شدن گرادیان‌ها را برطرف کرده و محاسبات را کارآمدتر می‌سازد. در بسیاری از مطالعات اخیر، LSTM‌ها عملکرد برتری از خود نشان داده‌اند و آموزش آن‌ها نیز نسبتاً آسان‌تر است. در نتیجه، LSTM‌ها به معماری‌های پایه برای پردازش داده‌های ترتیبی همراه با اطلاعات زمانی، تبدیل شده‌اند (یوزفویچ و همکاران^۸، ۲۰۱۵).

یک واحد LSTM از یک سلول و سه دروازه تشکیل شده است: دروازه‌ی ورودی، دروازه‌ی فراموشی، و دروازه‌ی خروجی. سلول حافظه، نقش بخش حافظه‌ی LSTM را ایفا می‌کند و برای ذخیره‌سازی مقادیر در بازه‌های زمانی دلخواه به کار می‌رود. در LSTM‌ها، یک وابستگی خطی میان سلول‌های حافظه c_t و حالت قبلی آن‌ها c_{t-1} برقرار می‌شود. علاوه بر این، LSTM شامل دروازه‌های ورودی و خروجی است که توابع غیرخطی را بر ورودی و خروجی LSTM اعمال می‌کنند. روابط موجود در یک LSTM به صورت زیر بیان می‌شود:

^۱ Backpropagation Algorithm

^۲ Hansson & Rostami

^۳ Long Short-Term Memory Network

^۴ Chen

^۵ Hochreiter

^۶ Schmidhuber

^۷ Graves

^۸ Jozefowicz

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1}) \quad (۳)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1}) \quad (۴)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1}) \quad (۵)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t) \quad (۶)$$

$$h_t = o_t \odot \tanh(c_t) \quad (۷)$$

در این روابط، i_t ، f_t و o_t به ترتیب دروازه‌ی ورودی، دروازه‌ی فراموشی، و دروازه‌ی خروجی را نشان می‌دهند (چو و همکاران، ۲۰۱۴).

واحد بازگشتی دروازه‌دار^۲ که توسط چو و همکاران (۲۰۱۴) معرفی شد، شباهت‌های بسیاری با LSTM دارد اما در نحوه‌ی به کارگیری دروازه‌ها متفاوت است. برخلاف LSTM که از سه دروازه استفاده می‌کند، GRU تنها از دو دروازه بهره می‌برد: دروازه‌ی بازنشانی و دروازه‌ی به‌روزرسانی. دروازه‌ی بازنشانی تعیین می‌کند که ورودی جدید چگونه با حافظه‌ی قبلی ترکیب شود و همچنین کنترل می‌کند که چه میزان از حافظه‌ی قبلی، نادیده گرفته شود. مقادیر کوچک برای دروازه‌ی بازنشانی نشان می‌دهد که حافظه‌ی پیشین تا حد زیادی کنار گذاشته می‌شود. دروازه‌ی به‌روزرسانی، میزان حفظ حافظه‌ی قبلی را مشخص می‌کند (سونگ^۳، ۲۰۱۸).

یکی از تفاوت‌های اساسی میان GRU و LSTM این است که GRU فاقد سلول حافظه‌ی مجزا است؛ بنابراین عملیات آن به شکل ساده‌تری تعریف می‌شود. روابط ریاضی در GRU به صورت زیر هستند:

$$z_t = \sigma(W_zx_t + U_zh_{t-1}) \quad (۹) \quad h_t = (1 - z_t)h_{t-1} + z_t\tilde{h}_t \quad (۸)$$

$$r_t = \sigma(W_rx_t + U_rh_{t-1}) \quad (۱۱) \quad \tilde{h}_t = \tanh(W_hx_t + U(r_t \odot h_{t-1})) \quad (۱۰)$$

در اینجا، h_t خروجی GRU در زمان t را نشان می‌دهد، در حالی که z_t و r_t به ترتیب نمایانگر دروازه‌ی به‌روزرسانی و دروازه‌ی بازنشانی هستند. ماتریس‌های W_z ، W_r ، W_h ، U_z و U_r نیز وزن‌های به کاررفته در معماری GRU هستند. ساختار ساده‌شده‌ی GRU موجب کاهش پیچیدگی محاسباتی می‌شود، در حالی که همچنان عملکرد قوی‌ای در ثبت وابستگی‌های زمانی از خود نشان می‌دهد (هانسن و راستمی^۴، ۲۰۱۹).

متغیر هدف در این مطالعه، شاخص قیمت سهام صنعت فلزات اساسی است که از پایگاه اطلاعاتی بورس و نرم‌افزار ره‌آورد نوین استخراج شده است. متغیرهای توضیحی، شامل نرخ ارز، قیمت جهانی نفت، حجم نقدینگی، شاخص قیمت مصرف‌کننده و قیمت جهانی طلا می‌باشند که از پایگاه اطلاعاتی بانک جهانی و گزارش نماگرهای اقتصادی بانک مرکزی به دست آمده است. دوره‌ی زمانی پژوهش، داده‌های روزانه دوره فروردین ۱۳۹۰ تا اسفند ۱۴۰۳ است. از آنجا که برخی از متغیرهای اقتصادی کلان نظیر شاخص قیمت مصرف‌کننده و نقدینگی با بسامد ماهانه منتشر می‌شوند و متغیر هدف در این مطالعه روزانه است، برای هم‌بسامدسازی داده‌ها از رویکرد انتشارمحور استفاده شده است. بدین صورت که در هر روز معاملاتی، تنها آخرین مقدار منتشرشده و در دسترس بازار از هر متغیر اقتصادی کلان به‌عنوان ورودی مدل، لحاظ گردیده است. برای تبدیل متغیرهای کم‌بسامد به بسامد روزانه، از روش نگه‌داشتن آخرین مشاهده مشروط به تاریخ انتشار رسمی استفاده شده است. بدین ترتیب، از بروز تورش نگاه به آینده، جلوگیری شده است. لازم به ذکر است که متغیرهای

^۱ Cho^۲ Gated Recurrent Unit^۳ Song^۴ Hansson & Rostami

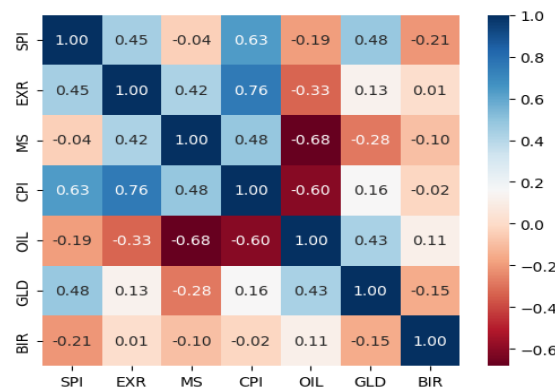
قیمت نفت، طلا و نرخ ارز که به صورت روزانه در دسترس بوده‌اند، بدون نیاز به هم‌بسامدسازی وارد مدل شده‌اند.

۵- یافته‌های پژوهش

پس از انجام مراحل اولیه آماده‌سازی داده‌ها، به منظور رفع ناهمگونی میان واحدها و مقیاس‌های متفاوت متغیرهای مورد استفاده، نرمال‌سازی داده‌ها انجام شده است. برای این منظور از روش Min-Max Normalization و کلاس MinMaxScaler استفاده گردید تا تمامی متغیرها بر اساس مقادیر حداقل و حداکثر خود در بازه استاندارد [۰، ۱] مقیاس‌بندی شوند. نرمال‌سازی داده‌ها نه تنها موجب جلوگیری از غلبه متغیرهای با مقدار بزرگ‌تر بر مدل می‌شود، بلکه روند آموزش شبکه‌های LSTM، RNN و GRU را پایدارتر کرده و به بهبود دقت پیش‌بینی‌ها کمک می‌کند.

در ادامه تحلیل، ماتریس هم‌بستگی بین شاخص قیمت سهام صنعت فلزات اساسی (SPI) و متغیرهای توضیحی، شامل نرخ ارز (EXR)، قیمت جهانی نفت (OIL)، حجم نقدینگی (MS)، شاخص قیمت مصرف‌کننده (CPI)، نرخ بهره (BIR) و قیمت جهانی طلا (GLD) محاسبه و به صورت نمایش حرارتی شکل (۱) ارائه شده است. این نمودار، شدت و جهت روابط خطی میان متغیرها را نمایش و به درک بهتر تأثیرگذاری متغیرها بر رفتار شاخص قیمت سهام صنعت فلزات اساسی کمک می‌کند. نتایج نشان می‌دهد برخی متغیرها مانند شاخص قیمت مصرف‌کننده، نرخ ارز و قیمت جهانی طلا دارای هم‌بستگی مثبت نسبتاً قوی با شاخص قیمت سهام صنعت فلزات اساسی هستند؛ این موضوع می‌تواند ناشی از ماهیت دلاری درآمدهای شرکت‌های صادرات‌محور و همچنین اثر انتقال قیمت‌ها در شرایط تورمی باشد. در حالی که قیمت جهانی نفت، حجم نقدینگی و نرخ بهره رابطه‌ای منفی با شاخص قیمت سهام صنعت فلزات اساسی دارند؛ این موضوع می‌تواند ناشی از نقش هزینه انرژی در افزایش هزینه تولید و همچنین ساختار غیرخطی انتقال شوک‌های نفتی به سودآوری شرکت‌ها باشد. همچنین با افزایش نرخ بهره، از طریق افزایش هزینه تامین مالی، افزایش نرخ تنزیل جریان‌های نقدی و کاهش جذابیت نسبی بازار سهام در مقایسه با دارایی‌های بدون ریسک، منجر به کاهش قیمت سهام می‌شود. با وجود این، هیچ‌یک از ضرایب هم‌بستگی، در سطح بسیار بالا یا نزدیک به مقدار یک قرار ندارند؛ بنابراین، مشکل هم‌خطی شدید میان متغیرها مشاهده نمی‌شود و هیچ متغیری به‌طور کامل بر دیگری منطبق نیست. نبود هم‌بستگی کامل، امکان استفاده هم‌زمان از این متغیرها را بدون ایجاد اختلال در فرایند مدل‌سازی فراهم می‌کند.

در این مرحله، با توجه به ماهیت سری‌زمانی داده‌ها و نیاز به استخراج الگوهای وابستگی کوتاه‌مدت و بلندمدت، معماری پیشنهادی پایه، مبتنی بر شبکه‌های بازگشتی، شامل LSTM، RNN و GRU مورد استفاده قرار گرفتند. طراحی این معماری که در جدول (۱) ارائه شده است، بر اساس مجموعه‌ای از آزمایش‌های تجربی اولیه و بررسی تعداد لایه‌ها و تعداد بهینه نورون‌ها برای دستیابی به بهترین عملکرد پیش‌بینی انجام شد. هر سه مدل به صورت Sequential پیاده‌سازی شده است و شامل دو لایه بازگشتی متوالی و یک لایه خروجی هستند. به منظور کنترل بیش‌برازش، پس از هر لایه بازگشتی از Dropout استفاده شده است که با حذف تصادفی بخشی از نورون‌ها، توانایی مدل را در تعمیم‌دهی افزایش می‌دهد. برای ارزیابی عملکرد الگوها، داده‌های موجود به دو بخش آموزش و آزمون تقسیم می‌شوند؛ در این فرآیند، ۸۰ درصد از کل داده‌ها به عنوان داده‌های آموزش مورد استفاده قرار گرفته است تا مدل بتواند الگوهای اصلی و روابط میان متغیرها را فراگیرد؛ همچنین ۲۰ درصد باقیمانده نیز به عنوان داده‌های آزمون در نظر گرفته شد تا توانایی مدل در تعمیم‌دهی به داده‌های مشاهده‌نشده، سنجیده شود.



شکل ۱. ماتریس هم‌بستگی شاخص قیمت سهام صنعت فلزات اساسی و متغیرهای توضیحی (منبع: یافته‌های پژوهش)

جدول ۱. پارامترهای استفاده شده در مدل پایه

پارامتر	مقدار
تابع فعال‌سازی	ReLU
بهینه‌ساز	Adam
تابع هزینه	MSE
تعداد نورون در لایه اول	۱۲۸
تعداد نورون در لایه دوم	۶۴
تعداد دوره آموزش	۲۰۰
اندازه دسته	۶۴
اندازه پنجره	۱۰

منبع: یافته‌های پژوهش

ارزیابی عملکرد الگوهای پیش‌بینی، یکی از مراحل بنیادین در پژوهش‌های مرتبط با یادگیری ماشین و شبکه‌های عصبی است. برای سنجش میزان دقت مدل، استفاده از شاخص‌های خطا ضروری است؛ زیرا این شاخص‌ها میزان انحراف مقادیر پیش‌بینی شده از مقادیر واقعی را کمی‌سازی می‌کنند. در میان معیارهای رایج، میانگین قدرمطلق خطا (MAE)، میانگین مربعات خطا (MSE)، مجذور میانگین مربعات خطا (RMSE)، میانگین درصد خطای مطلق (MAPE) و همچنین ضریب نابرابری تایل (Theil U)، از مهم‌ترین شاخص‌هایی هستند که در مسائل پیش‌بینی سری‌های زمانی مالی به کار می‌روند. شاخص MAE میانگین قدرمطلق خطاها را محاسبه کرده و تفسیر ساده و مستقیمی از میانگین فاصله پیش‌بینی‌ها از مقادیر واقعی ارائه می‌دهد. در مقابل، شاخص‌های MSE و RMSE نسبت به خطاهای بزرگ حساسیت بیشتری دارند، زیرا انحرافات بزرگ را به صورت مربعی در نظر می‌گیرند و بنابراین خطاهای شدید تأثیر بیشتری بر مقدار نهایی شاخص می‌گذارند. این ویژگی در حوزه پیش‌بینی بازارهای مالی اهمیت دارد، زیرا خطاهای بزرگ می‌توانند به تصمیمات اشتباه و زیان مالی قابل توجه منجر شوند. شاخص MAPE میانگین درصد خطای مطلق پیش‌بینی را نسبت به مقادیر واقعی نشان می‌دهد. شاخص Theil U خطای مدل را به صورت نرمال شده بیان می‌کند؛ این شاخص در بازه صفر تا یک قرار دارد و هرچه به صفر نزدیک‌تر باشد، مدل از دقت بیشتری برخوردار است. روابط ریاضی این شاخص‌ها در بخش زیر ارائه شده است:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2} \quad (۱۳)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2 \quad (۱۲)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{A_t} \times 100 \quad (۱۵)$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |A_t - F_t| \quad (۱۴)$$

$$Theil U = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n A_t^2} + \sqrt{\frac{1}{n} \sum_{t=1}^n F_t^2}} \quad (16)$$

پس از تعیین مدل پایه بر اساس پارامترهای ارائه شده در جدول (۱)، قدم بعدی شامل بهینه سازی معماری هر یک از شبکه های LSTM، RNN و GRU است. بدین منظور، مجموعه ای از آزمایش های تجربی با تغییر سیستماتیک پارامترهای مختلف از جمله تعداد نوروها در لایه های بازگشتی، تعداد دوره های آموزش، اندازه دسته و اندازه پنجره انجام شده است. هدف از این فرایند، یافتن معماری بوده است که کمترین میزان خطا را بر اساس شاخص های ارزیابی عملکرد مدل داشته باشد و بیشترین پایداری را در پیش بینی سری زمانی فراهم کند. در هر مرحله، معماری جدید، آموزش داده شد و عملکرد آن با سایر معماری ها مقایسه گردید. نتیجه این فرایند بهینه سازی، استخراج معماری های نهایی و منتخب برای هر یک از الگوها بود که در جدول (۲) ارائه شده اند. به عنوان نمونه، در معماری بهینه شبکه LSTM، تعداد ۱۲۸ نورو در لایه نخست، و ۶۴ نورو در لایه دوم به عنوان ترکیب مناسب تعیین شد. همچنین تعداد دوره های آموزش ۲۰۰، اندازه دسته ۳۲ و اندازه پنجره ۲۰ انتخاب گردید. فرایند مشابهی برای الگوهای RNN و GRU نیز انجام شد و پارامترهای نهایی آن ها بر اساس کمترین مقدار خطا و بیشترین دقت پیش بینی تعیین گشت. بدین ترتیب، معماری های ارائه شده در جدول (۲) بهترین عملکرد را در مقایسه با سایر ساختارهای آزمایشی نشان داده اند و به عنوان الگوهای بهینه انتخاب شدند.

جدول ۲. پارامترهای بهینه برای هر یک از الگوهای شبکه عصبی

پارامتر	RNN	LSTM	GRU	مدل شبکه عصبی
تابع فعال سازی	ReLU	ReLU	ReLU	
بهینه ساز	Adam	Adam	Adam	
تابع هزینه	MSE	MSE	MSE	
تعداد نورو در لایه اول	۶۴	۱۲۸	۱۲۸	
تعداد نورو در لایه دوم	۳۲	۶۴	۳۲	
تعداد دوره آموزش	۱۱۰	۲۰۰	۱۵۰	
اندازه دسته	۳۲	۳۲	۳۲	
اندازه پنجره	۲۰	۲۰	۲۰	

منبع: یافته های پژوهش

نتایج ارائه شده در جدول (۳) عملکرد الگوهای شبکه عصبی مختلف، شامل LSTM، RNN و GRU را بر اساس شاخص های MAPE، MAE، RMSE، MSE و Theil U به تفکیک معماری پایه و معماری بهینه، نشان می دهد. همان گونه که مشاهده می شود، مدل GRU با مقدار RMSE برابر با ۰/۱۱۶۱ بهترین عملکرد را در میان الگوهای مورد بررسی داشته است. این امر نشان می دهد معماری GRU می تواند در پیش بینی داده های سری زمانی مالی، در شرایط خاص، نتایج دقیق تری نسبت به سایر الگوهایی مانند LSTM و RNN ارائه دهد. پس از GRU، مدل LSTM با مقدار RMSE برابر با ۰/۱۱۶۷، در رتبه دوم قرار گرفته که نشان دهنده توانایی بالای این مدل (با ساختار پیچیده تر نسبت به مدل GRU) در یادگیری وابستگی های زمانی است. مدل RNN نیز با مقدار RMSE برابر با ۰/۱۲۷۱ در رتبه سوم قرار گرفته و نشان می دهد شبکه بازگشتی ساده RNN، در مقایسه با شبکه های بازگشتی پیچیده تر، توان کمتری در مدل سازی پویایی های سری های زمانی دارد.

نتایج به دست آمده نشان می دهد بهینه سازی پارامترهای شبکه های عصبی، اثر معنادار و قابل توجهی در بهبود عملکرد این الگوها داشته است. در هر سه معماری LSTM، RNN و GRU، مقادیر شاخص های خطا در مدل بهینه نسبت به مدل

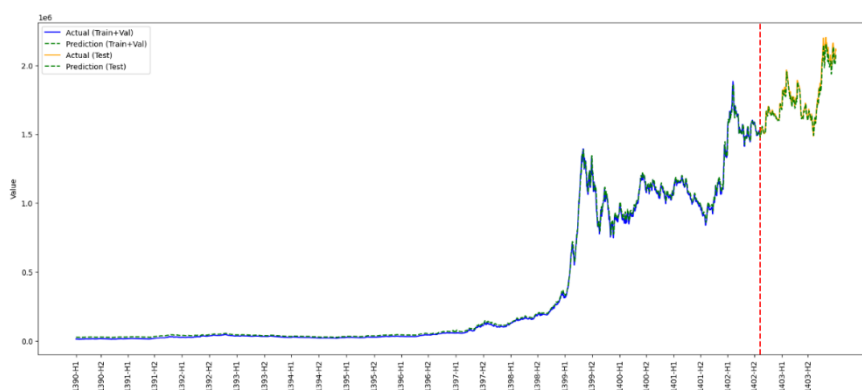
پایه، کاهش یافته که بیانگر افزایش دقت پیش‌بینی و بهبود توان تعمیم‌پذیری الگوها است. این امر نشان می‌دهد که تنظیم بهینه پارامترهایی نظیر اندازه پنجره، تعداد نوروها، تعداد دوره‌های آموزش و سایر تنظیمات ساختاری، نقش مهمی در کاهش بیش‌برازش و ارتقای کیفیت پیش‌بینی ایفا می‌کند. از سوی دیگر، مقایسه نتایج الگوهای یادگیری عمیق با مدل مبنای ARIMAX حاکی از برتری شبکه‌های عصبی در مدل‌سازی رفتار شاخص قیمت سهام است. کاهش مقادیر شاخص‌های خطا در الگوهای بهینه نسبت به مدل کلاسیک ARIMAX نشان می‌دهد که شبکه‌های عصبی، توانایی بیشتری در استخراج الگوهای پیچیده، روابط غیرخطی و وابستگی‌های زمانی پنهان در داده‌های مالی دارند. این یافته‌ها بیانگر آن است که در محیط‌های مالی که با نوسانات، نویز و روابط پویا همراه‌اند، بهره‌گیری از معماری‌های یادگیری عمیق می‌تواند دقت پیش‌بینی بالاتری نسبت به الگوهای کلاسیک فراهم آورد.

جدول ۳. مقایسه عملکرد الگوهای شبکه عصبی و مدل مبنای

مدل	مدل بهینه			مدل پایه			مدل
	RNN	LSTM	GRU	ARIMAX	RNN	LSTM	
MSE	۰/۰۱۶۱	۰/۰۱۳۶	۰/۰۱۳۴	۰/۰۲۱۵	۰/۰۲۲۵	۰/۰۱۸۳	GRU
RMSE	۰/۱۲۷۱	۰/۱۱۶۷	۰/۱۱۶۱	۰/۱۴۶۸	۰/۱۵۰۱	۰/۱۳۵۲	
MAE	۰/۱۰۱۹	۰/۰۹۱۷	۰/۰۹۱۰	۰/۱۳۲۱	۰/۱۴۳۶	۰/۱۲۱۸	
MAPE	۱۱/۸۷۴۵	۱۱/۲۵۱۵	۱۱/۲۱۰۸	۱۴/۹۵۵۹	۱۵/۹۱۱۸	۱۴/۳۹۱۲	
Theil U	۰/۰۸۲۸	۰/۰۷۳۱	۰/۰۷۲۷	۰/۱۱۲۵	۰/۱۱۲۶	۰/۰۹۴۲	

منبع: یافته‌های پژوهش

شکل (۲) روند پیش‌بینی شاخص قیمت سهام صنعت فلزات اساسی را برای شبکه عصبی واحد بازگشتی دروازه‌دار نشان می‌دهد. همان‌طور که مشاهده می‌شود، مدل توانسته است الگوی کلی تغییرات شاخص را با دقت بالایی دنبال کند و هم‌پوشانی مناسبی میان مقادیر واقعی و پیش‌بینی شده در هر دو بخش آموزش و آزمون ایجاد نماید. در بخش آزمون، برآزش بسیار نزدیک میان سری واقعی و پیش‌بینی شده بیانگر توانایی مدل در یادگیری ساختارهای غیرخطی داده‌ها است. هرچند انحرافات محدودی در برخی نقاط مشاهده می‌شود، اما مدل همچنان روند کلی شاخص و نقاط تغییر مسیر را به‌خوبی بازتولید می‌کند. این نتایج نشان‌دهنده قدرت شبکه عصبی واحد بازگشتی دروازه‌دار در استخراج وابستگی‌های زمانی و تعمیم‌پذیری مناسب آن در پیش‌بینی سری‌های زمانی مالی است.



شکل ۲. پیش‌بینی با مدل شبکه عصبی GRU (منبع: یافته‌های پژوهش)

۶- نتیجه‌گیری و پیشنهادها

هدف اصلی پژوهش حاضر، ارزیابی دقت الگوهای مختلف شبکه عصبی در پیش‌بینی شاخص قیمت سهام صنعت فلزات اساسی بوده است. برای دستیابی به این هدف، از داده‌های روزانه شاخص قیمت سهام صنعت فلزات اساسی طی

دوره فروردین ۱۳۹۰ تا اسفند ۱۴۰۳ استفاده شد و متغیرهای اقتصادی کلان اثرگذار بر عملکرد بازار سهام شامل نرخ ارز، قیمت جهانی نفت، حجم نقدینگی، شاخص قیمت مصرف کننده و قیمت جهانی طلا به عنوان متغیرهای توضیحی در مدل لحاظ شدند. در این مطالعه، سه مدل پر کاربرد شبکه عصبی بازگشتی شامل شبکه عصبی بازگشتی ساده (RNN)، شبکه حافظه طولانی کوتاه مدت (LSTM) و واحد بازگشتی دروازه دار (GRU) به منظور پیش بینی و مقایسه عملکرد، مورد استفاده قرار گرفتند. در ارزیابی دقت پیش بینی شاخص قیمت، از معیارهای ارزیابی ریشه دوم میانگین مربعات خطا، میانگین قدرمطلق خطا، میانگین درصد خطای مطلق و ضریب نابرابری تایل استفاده گردید. بر اساس نتایج، شبکه عصبی واحد بازگشتی دروازه دار بهترین عملکرد را در پیش بینی شاخص قیمت، ثبت کرده است و پس از آن، به ترتیب حافظه طولانی کوتاه مدت، و شبکه بازگشتی ساده قرار دارند. به طور کلی، مزیت اصلی الگوهای شبکه عصبی بازگشتی، توانایی یادگیری وابستگی های زمانی است؛ قابلیتی که در داده های مالی، اهمیت ویژه ای دارد. نتیجه به دست آمده، همسو با مطالعاتی همچون کوانگ^۱ (۲۰۲۳)، گویی و یون وو^۲ (۲۰۲۲)، کمالو^۳ (۲۰۲۰)، و حیدری و امیری (۱۴۰۲) است.

یافته های پژوهش نشان دادند واحد بازگشتی دروازه دار، کمترین خطای پیش بینی را ثبت کرده است؛ با این حال، تفسیر این نتیجه صرفاً «بهتر بودن عددی» نیست؛ بلکه از منظر سازوکار یادگیری نیز قابل توضیح است. داده های مالی معمولاً پویا، نوسانی و دارای روابط غیرخطی اند و علاوه بر شوک های کوتاه مدت، نوعی حافظه ای زمانی را نیز در خود دارند. در چنین شرایطی، واحد بازگشتی دروازه دار نسبت به شبکه عصبی بازگشتی ساده، مزیت دارد زیرا مشکل محوشدگی گرادیان را کاهش داده و توان یادگیری وابستگی های زمانی بلندمدت را حفظ می کند. واحد بازگشتی دروازه دار، با وجود بهره گیری از مکانیزم دروازه ها، نسبت به شبکه حافظه طولانی کوتاه مدت، پارامترهای کمتری دارد و ساختار ساده تری ارائه می کند؛ بنابراین در داده های سری زمانی مالی، می تواند ریسک بیش برآزش را کاهش داده و تعمیم پذیری بالاتری فراهم کند. برای اطمینان از پایداری نتایج، در طراحی الگوها از نرمال سازی داده ها، Dropout و بهینه سازی پارامترهای کلیدی استفاده شد. همچنین بررسی هم بستگی متغیرهای مستقل، نبود هم خطی شدید را تایید کرد؛ بنابراین مدل می تواند از اطلاعات متغیرهای کلان اقتصادی به صورت بهینه بهره گیرد. از این رو، واحد بازگشتی دروازه دار، علاوه بر کمترین خطا، به دلیل سازگاری با ماهیت سری های زمانی مالی، گزینه ای مناسب برای پیش بینی شاخص قیمت سهام صنعت فلزات اساسی ایران است. مطابق یافته های پژوهش، توصیه های سیاستی زیر ارائه می گردد:

- نهادینه سازی استفاده از الگوهای یادگیری عمیق در نهادهای سیاست گذار و نظارتی بازار سرمایه: با توجه به برتری شبکه عصبی واحد بازگشتی دروازه دار در پیش بینی شاخص قیمت سهام، پیشنهاد می شود نهادهای سیاست گذار و نظارتی از الگوهای یادگیری عمیق به عنوان ابزار مکمل تحلیل و پیش بینی استفاده کنند. به کارگیری این الگوها می تواند دقت پیش بینی، کیفیت تصمیم گیری و مدیریت ریسک بازار سرمایه را بهبود بخشد.
- تقویت زیرساخت های داده ای و شفافیت اطلاعاتی در بازار سرمایه: از آنجا که عملکرد الگوهای یادگیری ماشین به کیفیت داده ها وابسته است، پیشنهاد می شود با توسعه پایگاه های داده یکپارچه، بهبود دسترسی به اطلاعات و ارتقای شفافیت، زمینه توسعه ابزارهای پیش بینی هوشمند و افزایش کارایی اطلاعاتی بازار سرمایه فراهم شود.

¹ Kuang³ Kamalov² Gui & Wu

- حمایت هدفمند از پژوهش و توسعه در حوزه هوش مصنوعی مالی: پیشنهاد می‌شود سیاست‌گذاران با حمایت از پژوهش‌های مرتبط، توسعه الگوهای بومی، آموزش نیروی انسانی متخصص و تقویت همکاری دانشگاه‌ها و نهادهای مالی، زمینه گسترش کاربرد هوش مصنوعی در بازار سرمایه و ارتقای توان تحلیل و پیش‌بینی را فراهم کنند.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

نویسندگان اصول اخلاقی را در انجام و انتشار این پژوهش علمی رعایت و این موضوع مورد تأیید همه آنهاست.

مشارکت نویسندگان

همه نویسندگان در مقاله سهم و نقش یکسان داشته‌اند.

تعارض منافع

بنا بر اظهار نویسندگان این مقاله تعارض منافع ندارد.

حامی مالی

نویسندگان هیچگونه حمایت مالی برای تحقیق، تألیف و انتشار این مقاله دریافت نکرده‌اند.

تقدیر و تشکر

از تمامی افرادی که در انجام این تحقیق مؤلفین را مساعدت نمودند، قدردانی می‌شود.

منابع

- اسدی، مسعود؛ میربرگ کار، سید مظفر و چیرانی، ابراهیم (۱۴۰۱). ارائه یک مدل شبکه عصبی برای پیش‌بینی سود شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران و مقایسه دقت آن با الگوهای HDZ و ARIMA. *مجله حسابداری مدیریت*، ۱۵(۵۴)، ۳۳-۴۷. DOI: 10.30495/jma.2022.21080
- بالونزادنوری، روزبه و شهیدی، آمنه (۱۴۰۴). بررسی رابطه غیرخطی مستقیم و معکوس نرخ ارز و بازدهی بازار سهام در ایران. *مجله توسعه و سرمایه*، ۱۰(۲)، ۱۴۶-۱۲۵. DOI: 10.22103/jdc.2024.23602.1474
- تسلیم‌پور، علی؛ عسگرزاده، غلامرضا؛ قالمق، کریم و نصیری، هومن (۱۴۰۴). پیش‌بینی شاخص قیمت و بازده نقدی بورس اوراق بهادار تهران (TEDPIX) با استفاده از مدل شبکه عصبی NARX. *فصلنامه مدیریت دارایی و تأمین مالی*، ۱۴(۱)، ۴۸-۲۱. DOI: 10.22108/amf.2025.143461.1940
- حیدری، مهدی و امیری، حمیدرضا (۱۴۰۱). بررسی توان الگوهای مبتنی بر هوش مصنوعی در پیش‌بینی روند قیمت سهام بورس اوراق بهادار تهران. *پژوهش‌های مالی*، ۲۴(۴)، ۶۲۳-۶۰۲. DOI: 10.22059/frj.2022.320064.1007149
- کاویانی، میثم؛ فخرحسینی، سید فخرالدین و حسن‌پور، امین (۱۴۰۴). تحلیل مقایسه‌ای پویایی شاخص‌های بازار سرمایه تهران با استفاده از مدل خودرگرسیون در چارچوب رگرسیون کوانتایل. *توسعه و سرمایه*، مقالات آماده انتشار. DOI: 10.22103/jdc.2025.25330.1561
- سیف، سمیرا؛ جمشیدی نوید، بابک؛ قنبری، مهرداد و اسماعیل‌پور، منصور (۱۴۰۰). پیش‌بینی روند بازار سهام ایران با استفاده از نوسان‌نمای موج الیوت و شاخص قدرت نسبی. *پژوهش‌های مالی*، ۲۳(۱)، ۱۵۷-۱۳۴. DOI: 10.22059/frj.2020.310664.1007072
- شریف‌فر، امیر؛ خلیلی عرقی، مریم؛ رئیسی ونانی، ایمان و فلاح شمس، میرفیض (۱۴۰۱). کاربرد معماری‌های یادگیری عمیق در پیش‌بینی قیمت سهام: رویکرد شبکه عصبی پیچشی. *فصلنامه مدیریت دارایی و تأمین مالی*، ۱۰(۳)، ۲۰-۱. DOI: 10.22108/amf.2022.129205.1673
- صمدی، سعید؛ شریفی، علیمراد؛ احمدزاده، عزیز و خانزادی، آزاد (۱۳۸۸). بررسی امکان جانشینی بین نهادهای انرژی با سرمایه در بخش فلزات اساسی. *فصلنامه تحقیقات اقتصادی*، ۴۴(۴)، ۱۲۹-۱۵۵. https://jte.ut.ac.ir/article_20344.html
- عسگری، منصور (۱۳۹۹). عوامل مؤثر بر اشتغال در صنایع کارخانه‌ای کشور. *مجله توسعه و سرمایه*، ۵(۲)، ۴۷-۶۵. DOI: 10.22103/jdc.2020.11705.1040
- فخاری، حسین؛ ولی‌پور خطیر، محمدولی و موسوی، سیده مانده (۱۳۹۶). ارزیابی عملکرد شبکه‌های عصبی بیزین و لونبرگ-مارکوارت در مقایسه با الگوهای کلاسیک در پیش‌بینی قیمت سهام شرکت‌های سرمایه‌گذاری. *پژوهش‌های مالی*، ۱۹(۲)، ۳۱۸-۲۹۹. DOI: 10.22059/jfr.2017.214203.1006264

براقی، مرضیه و ربیعی، اعظم (۱۳۹۹). مروری و مقایسه الگوریتم‌های شبکه عصبی بازگشتی عمیق LSTM و GRU در مدل‌سازی داده‌های سری زمانی نرخ ارز. *علوم محاسباتی*، ۵(۴)، ۴۰-۵۰. <https://civilica.com/doc/1901544>

References

- Al-Ameer, M., Hammad, W., Ismail, A., & Hamdan, A. (2018). The relationship of gold price with the stock market: The case of Frankfurt Stock Exchange. *International Journal of Energy Economics and Policy*, 8(5), 357–371. <https://www.econjournals.com/index.php/ijeeep/article/view/6720>
- Anwer, Z., Azmi, W., & Ramadili Mohd, S. M. (2019). How do Islamic equities respond to monetary actions? *International Journal of Emerging Markets*, 14(4), 503–522. <https://doi.org/10.1108/IJOEM-11-2017-0459>
- Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques—Part II: Soft computing methods. *Expert Systems with Applications*, 36(3), 5932–5941. [10.1016/j.eswa.2008.07.006](https://doi.org/10.1016/j.eswa.2008.07.006)
- Bermúdez Delgado, N. A., Bermúdez Delgado, E., & Saucedo, E. (2018). The relationship between oil prices, the stock market, and the exchange rate: Evidence from Mexico. *The North American Journal of Economics and Finance*, 45, 266–275. <https://doi.org/10.1016/j.najef.2017.10.006>
- Asadi, M., Mirbargkar, S. M., & Chirani, E. (2022). Presenting a neural network model for predicting the profits of companies listed on the Tehran Stock Exchange and comparing its accuracy with HDZ and ARIMA models. *Management Accounting Journal*, 15(54), 33–47. <https://doi.org/10.30495/jma.2022.21080> [In Persian].
- Asgari, M. (2021). Factors affecting employment in manufacturing industries of Iran. *Journal of Development and Capital*, 5(2), 47–65. <https://doi.org/10.22103/jdc.2020.11705.1040> [In Persian].
- Balounejad Nouri, R., & Shahidi, A. (2025). Investigating the direct and inverse nonlinear relationship between exchange rate and stock market return in Iran. *Journal of Development and Capital*, 10(2), 125–146. <https://doi.org/10.22103/jdc.2024.23602.1474> [In Persian].
- Barakat, M. R., Elgazzar, S. H., & Hanafy, K. M. (2015). Impact of macroeconomic variables on stock markets: Evidence from emerging markets. *International Journal of Economics and Finance*, 8(1), 195–207. <https://doi.org/10.5539/ijef.v8n1p195>
- Brown, M. S., Pelosi, M. J., & Dirska, H. (2013). Dynamic-radius species-conserving genetic algorithm for the financial forecasting of Dow Jones index stocks. In *International Workshop on Machine Learning and Data Mining in Pattern Recognition* (pp. 27–41). Springer. https://doi.org/10.1007/978-3-642-39712-7_3
- Chen, X., Wei, L., & Xu, J. (2022). House price prediction using LSTM. *arXiv*. [10.48550/arXiv.1709.08432](https://arxiv.org/abs/10.48550/arXiv.1709.08432)
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734. <https://doi.org/10.3115/v1/D14-1179>
- Dahir, A. M., Mahat, F., Ab Razak, N. H., & Bany-Arifin, A. N. (2018). Revisiting the dynamic relationship between exchange rates and stock prices in BRICS countries: A wavelet analysis. *Borsa Istanbul Review*, 18(2), 101–113. <https://doi.org/10.1016/j.bir.2017.10.001>
- Etale, L. M., & Eze, G. P. (2019). Analysing stock market reaction to macroeconomic variables: Evidence from Nigerian Stock Exchange (NSE). *Global Journal of Arts, Humanities and Social Sciences*, 7(3), 14–28. <https://ejournals.org/wp-content/uploads/Analysing.pdf>
- Fakhari, H., Valipour Khateer, M., & Mousavi, S. M. (2017). Evaluating the performance of Bayesian and Levenberg–Marquardt neural networks compared to classical models in predicting stock prices of investment companies. *Financial Research*, 19(2), 299–318. [10.22059/jfr.2017.214203.1006264](https://doi.org/10.22059/jfr.2017.214203.1006264) [In Persian].
- Galeshchuk, S. (2016). Neural networks performance in exchange rate prediction. *Neurocomputing*, 172, 446–452. <https://doi.org/10.1016/j.neucom.2015.03.100>
- Gui, J., & Wu, X. (2022). Forecasting the stock price of vaccine manufacturers in China using machine learning and econometrics model. In *Proceedings of the International Conference on Cyber Security, Artificial Intelligence, and Digital Economy (CSAIDE 2022)* (Vol. 12330, pp. 577–584). SPIE. [10.1117/12.2647506](https://doi.org/10.1117/12.2647506)
- Hansson, F., & Rostami, J. (2019). *Time series forecasting of house prices: An evaluation of a support vector machine and a recurrent neural network with LSTM cells* (Bachelor's thesis, Uppsala University). <https://www.maklarstatistik.se/wp-content/uploads/kandidatuppsats2019.pdf>
- Heydari, M., & Amiri, H. R. (2022). Investigating the power of artificial intelligence-based models in predicting stock price trends of the Tehran Stock Exchange. *Financial Research*, 24(4), 602–623. <https://doi.org/10.22059/frj.2022.320064.1007149> [In Persian].

- Hyde, S. (2007). The response of industry stock returns to market, exchange rate, and interest rate risks. *Managerial Finance*, 33(9), 693–709. <https://doi.org/10.1108/03074350710776244>
- Jain, A., & Biswal, P. C. (2016). Dynamic linkages among oil price, gold price, exchange rate, and stock market in India. *Resources Policy*, 49, 179–185. <https://doi.org/10.1016/j.resourpol.2016.06.001>
- Jozefowicz, R., Zaremba, W., & Sutskever, I. (2015). An empirical exploration of recurrent network architectures. In *Proceedings of the 32nd International Conference on Machine Learning* (pp. 2342–2350). PMLR. <https://proceedings.mlr.press/v37/jozefowicz15.html>
- Kamalov, F. (2020). Forecasting significant stock price changes using neural networks. *Neural Computing and Applications*, 32(23), 17655–17667. <https://doi.org/10.1007/s00521-020-04942-3>
- Kaviani, M., Fakhrehosseini, S. F., & Hasanpour, A. (2025). A comparative analysis of the dynamics of Tehran stock market indexes using an autoregressive model within the quantile regression framework. *Journal of Development and Capital*. Advance online publication. [10.22103/jdc.2025.25330.1561](https://doi.org/10.22103/jdc.2025.25330.1561) [In Persian].
- Lee, J. W., & Brahmasrene, T. (2018). An exploration of dynamical relationships between macroeconomic variables and stock prices in Korea. *Journal of Asian Finance, Economics and Business*, 5(3), 7–17. <https://doi.org/10.13106/jafeb.2018.vol5.no3.7>
- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17(1), 59–82. <https://doi.org/10.1257/089533003321164958>
- Megaravalli, A. V., & Sampagnaro, G. (2018). Macroeconomic indicators and their impact on stock markets in Asian 3: A pooled mean group approach. *Cogent Economics & Finance*, 6(1), Article 1432450. <https://doi.org/10.1080/23322039.2018.1432450>
- Samadi, S., Sharifi, A., Ahmadzadeh, A., & Khanzadi, A. (2010). Substitutability between energy and capital in basic metals manufacturing. *Journal of Economic Research*, 44(4), 129–155. jte.ut.ac.ir/article_20344.html [In Persian].
- Seif, S., Jamshidi Navid, B., Ghanbari, M., & Esmailpour, M. (2021). Predicting Iran's stock market trend using the Elliott Wave oscillator and the Relative Strength Index. *Financial Research*, 23(1), 134–157. <https://doi.org/10.22059/frj.2020.310664.1007072> [In Persian].
- Shariffar, A., Khaliliaraghi, M., Raeesi Vanani, I., & Fallahshams, M. (2022). Application of deep learning architectures in stock price forecasting: A convolutional neural network approach. *Journal of Asset Management and Financing*, 10(3), 1–20. <https://doi.org/10.22108/amf.2022.129205.1673> [In Persian].
- Shiller, R. J. (2017). Narrative economics. *American Economic Review*, 107(4), 967–1004. [10.1257/aer.107.4.967](https://doi.org/10.1257/aer.107.4.967)
- Kuang, S. (2023). A comparison of linear regression, LSTM model and ARIMA model in predicting stock price: A case study of HSBC stock price. *BCP Business & Management*, 44, 478–488. [10.54691/bcpbm.v44i.4858](https://doi.org/10.54691/bcpbm.v44i.4858)
- Singh, N. P., & Sharma, S. (2018). Phase-wise analysis of dynamic relationship among gold, crude oil, U.S. dollar, and stock market. *Journal of Advances in Management Research*, 15(4), 480–499. [10.1108/JAMR-12-2017-0124](https://doi.org/10.1108/JAMR-12-2017-0124)
- Song, Y. (2018). *Stock trend prediction: Based on machine learning methods* (Doctoral dissertation, University of California, Los Angeles). ProQuest. <https://www.proquest.com/docview/2031586457>
- Taslimpour, A., Askarzadeh, G., Ghalmegh, K., & Nasiri, H. (2026). Forecasting the Tehran Stock Exchange Dividend and Price Index (TEDPIX) using a NARX neural network model. *Journal of Asset Management and Financing*, 14(1), 21–48. <https://doi.org/10.22108/amf.2025.143461.1940> [In Persian].
- Tiryaki, A., Ceylan, R., & Erdogan, L. (2019). Asymmetric effects of industrial production, money supply, and exchange rate changes on stock returns in Turkey. *Applied Economics*, 51(20), 2143–2154. <https://doi.org/10.1080/00036846.2018.1540850>
- Verma, R. K., & Bansal, R. (2021). Impact of macroeconomic variables on the performance of stock exchange: A systematic review. *International Journal of Emerging Markets*, 16(7), 1291–1329. [10.1108/IJOEM-11-2019-0993](https://doi.org/10.1108/IJOEM-11-2019-0993)
- Vidal, A., & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, 157, Article 113481. <https://doi.org/10.1016/j.eswa.2020.113481>
- Yaraqi, M., & Rabiei, A. (2020). A review and comparison of deep recurrent neural network algorithms LSTM and GRU in modeling exchange rate time-series data. *Computational Sciences*, 5(4), 40–50. <https://search.isc.ac/dl/search/defaultta.aspx?DTC=8&DC=1328953> [In Persian].