



Predicting Index Trend Using Hybrid Neural Networks with a Focus on Multi-Scale Temporal Feature Extraction in the Tehran Stock Exchange

Mohammad Osoolian *

*Corresponding Author, Assistant Prof., Department of Financial Management and Insurance, Faculty of Management and Accounting, Shahid Beheshti University, Tehran, Iran. E-mail: m_osoolian@sbu.ac.ir

Ali Nikmaram

MSc., Department of Financial Management, Faculty of Management and Accounting, Shahid Beheshti University, Tehran, Iran. E-mail: nik_maram@mail.sbu.ac.ir

Mahdi Karimi

MSc., Department of Financial Management, Faculty of Management and Accounting, Shahid Beheshti University, Tehran, Iran. E-mail: mahdi.karimi2@mail.sbu.ac.ir

Abstract

Objective

Predicting the future trends in financial markets stands as a critical task for both investors and researchers, given its pivotal role in enabling well-informed decision-making processes and effective risk management strategies. Nevertheless, the realm of stock market dynamics is fraught with inherent complexities and uncertainties, posing a formidable challenge when it comes to achieving accurate predictions. A wide array of predictive modeling techniques have been meticulously investigated, spanning from conventional statistical methodologies to more sophisticated machine learning algorithms. The primary focus of this research endeavor revolves around the predictive analysis of the Tehran Stock Exchange (TSE)

Citation: Osoolian, Mohammad; Nikmaram, Ali & Karimi, Mahdi (2025). Predicting Index Trend Using Hybrid Neural Networks with a Focus on Multi-Scale Temporal Feature Extraction in the Tehran Stock Exchange. *Financial Research Journal*, 27(1), 85-113. <https://doi.org/10.22059/FRJ.2024.377816.1007611> (in Persian)



Composite Index, wherein a novel hybrid neural network framework is employed. This approach seamlessly integrates multiscale temporal features, with the ultimate objective of bolstering prediction precision and offering profound insights into prevailing market trends and dynamics.

Methods

The hybrid neural network architecture that has been put forward integrates the unique capabilities of convolutional neural networks (CNNs) in the realm of feature extraction with the effectiveness of long short-term memory (LSTM) networks in capturing temporal dependencies. The dataset used in this study consists of daily historical data pertaining to the TSE Composite Index, covering a substantial period from the year 1998 to 2022, which has been meticulously gathered, preprocessed, and subsequently partitioned into distinct sets for training and validation purposes. Within the framework of this hybrid neural network model, a sophisticated approach is adopted to harness multiscale temporal features derived from the input data, enabling the generation of highly accurate predictions regarding the future trends of the index. Moreover, to further enhance the performance and resilience of the model, sophisticated feature engineering methodologies are implemented to optimize its overall functionality.

Results

The results of the study reveal that while the hybrid neural network model, integrating CNN and LSTM components, demonstrates promising capabilities in predicting the TSE Composite Index, its accuracy falls short compared to competing models, particularly at weekly and monthly time scales. Conversely, the standalone CNN model exhibits superior performance, yielding more accurate predictions of the index's movements. These findings challenge the prevailing notion regarding the efficacy of hybrid neural network models in financial market prediction, highlighting the importance of evaluating alternative modeling approaches based on their specific strengths and limitations.

Conclusion

Despite the potential of hybrid neural network models, as demonstrated in previous research, the findings of this study suggest that simpler neural network architectures, such as CNNs, may offer better prediction performance in certain scenarios. To address the limitations identified, future research endeavors could explore alternative model configurations, ensemble methods, or hybrid architectures that combine the strengths of different predictive models. Additionally, incorporating additional market indicators and exploring intraday data sources could further enhance prediction accuracy and robustness. This abstract encapsulates the key findings and implications of the research, providing valuable insights for investors, researchers, and practitioners in the field of financial market prediction.

Keywords: Neural network, Convolution, Long-term memory, Modeling.

پیش‌بینی روند شاخص کل با استفاده از شبکه‌های عصبی هیبریدی با تمرکز بر استخراج ویژگی مقیاس زمانی چندگانه در بورس اوراق بهادار تهران

محمد اصولیان*

* نویسنده مسئول، استادیار، گروه مدیریت مالی و بیمه، دانشکده مدیریت و حسابداری، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: m_osoolian@sbu.ac.ir

علی نیک‌مرام

کارشناس ارشد، گروه مدیریت مالی، دانشکده مدیریت و حسابداری، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: nik_maram@mail.sbu.ac.ir

مهدی کریمی

کارشناس ارشد، گروه مدیریت مالی، دانشکده مدیریت و حسابداری، دانشگاه شهید بهشتی، تهران، ایران. رایانامه: mahdi.karimi2@mail.sbu.ac.ir

چکیده

هدف: پیش‌بینی آینده در حوزه سرمایه‌گذاری اهمیت زیادی دارد؛ زیرا به سرمایه‌گذاران کمک می‌کند تا تصمیم‌های بهتری اتخاذ کنند و ریسک‌های خود را کاهش دهند. در این راستا با بهبود قدرت مدل‌های پیش‌بینی، می‌توان به بازدهی‌های بهتری در بازار دست یافت. با این حال، پیش‌بینی بازار سهام به دلیل نوسان قیمت‌ها و عدم قطعیت، دغدغه بزرگی است. به‌طور کلی، پیش‌بینی دقیق حرکت سهام بسیار دشوار است و بسیاری از پژوهشگران به بررسی روش‌هایی می‌پردازند که فقط جهت حرکت سهام را پیش‌بینی می‌کنند. از جمله این روش‌ها، می‌توان به گشت تصادفی، پروبیت و لاجیت اشاره کرد. روش‌های جدیدتری مانند ماشین بردار پشتیبان، الگوریتم نزدیک‌ترین همسایگی و شبکه عصبی مصنوعی، برای بهبود پیش‌بینی آینده معرفی شده‌اند. به‌علت اهمیت پیش‌بینی روند بازارهای مالی برای پژوهشگران و سرمایه‌گذاران، این پژوهش با هدف پیش‌بینی روند شاخص کل بورس اوراق بهادار تهران با استفاده از یک شبکه عصبی هیبریدی با تمرکز بر استخراج ویژگی مقیاس زمانی چندگانه انجام شده است. هدف از اجرای این پژوهش، بررسی قدرت پیش‌بینی روش معرفی شده و مقایسه آن با روش‌های رقیب است.

روش: در این پژوهش، از یک شبکه عصبی هیبریدی که شامل شبکه عصبی کانولوشن (CNN) برای استخراج ویژگی‌ها و سه شبکه عصبی حافظه طولانی - کوتاه‌مدت (LSTM) برای یادگیری وابستگی‌های زمانی است، استفاده شده است. داده‌های استفاده‌شده، مقادیر روزانه شاخص کل بورس اوراق بهادار تهران از سال ۱۳۷۷ تا ۱۴۰۱ بود که پس از جمع‌آوری و نرمال‌سازی، به دو بخش آموزش و اعتبارسنجی تقسیم شد. این شبکه عصبی هیبریدی با بهره‌گیری از ویژگی مقیاس زمانی چندگانه، تلاش می‌کند تا پیش‌بینی دقیقی از

استناد: اصولیان، محمد؛ نیک‌مرام، علی و کریمی، مهدی (۱۴۰۴). پیش‌بینی روند شاخص کل با استفاده از شبکه‌های عصبی هیبریدی با تمرکز بر استخراج ویژگی مقیاس زمانی چندگانه در بورس اوراق بهادار تهران. *تحقیقات مالی*، ۲۷(۱)، ۸۵-۱۱۳.

روند شاخص ارائه دهد. همچنین، از روش‌های مهندسی استخراج برای بهبود دقت این شبکه‌ها استفاده شده که عبارت است از: ترکیب شبکه‌های عصبی مختلف در یک شبکه جامع.

یافته‌ها: نتایج این پژوهش نشان داد که مدل شبکه عصبی هیبریدی پیشنهادی که ترکیبی از شبکه‌های عصبی CNN و LSTM است، برای پیش‌بینی شاخص کل بورس اوراق بهادار تهران توانایی کافی را ندارد. دقت این مدل در مقیاس‌های زمانی هفتگی و ماهانه، کمتر از مدل‌های رقیب بود. در مقابل، مدل شبکه عصبی CNN که به‌عنوان یکی از مدل‌های رقیب بررسی شد، عملکرد بهتری داشت و توانست نتایج دقیق‌تری در پیش‌بینی شاخص کل بورس ارائه دهد. این نتایج با مطالعات قبلی که موفقیت مدل‌های هیبریدی در پیش‌بینی بازارهای مختلف را نشان داده بودند، در تضاد است.

نتیجه‌گیری: مدل شبکه عصبی هیبریدی پیشنهادی نتوانست شاخص کل بورس اوراق بهادار تهران را به‌درستی پیش‌بینی کند؛ در حالی که مدل CNN به تنهایی نتایج بهتری ارائه داد. این یافته‌ها نشان می‌دهد که شبکه‌های عصبی ساده‌تر، مانند CNN، ممکن است در مواردی عملکرد بهتری داشته باشند. برای پژوهش‌های آتی، پیشنهاد می‌شود با تغییر داده‌های روزانه به داده‌های بین‌روزی (مانند داده‌های دقیقه‌ای)، مدل شبکه عصبی هیبریدی مبتنی بر استخراج ویژگی زمانی چندگانه بار دیگر بررسی شود. همچنین، استفاده از شاخص‌های بیشتری مانند مقادیر آغازین، حجم، حداقل، حداکثر، میانگین متحرک و شاخص قدرت نسبی، می‌تواند بهبود دقت مدل‌های پیش‌بینی را به همراه داشته باشد.

کلیدواژه‌ها: شبکه عصبی، کانولوشن، حافظه طولانی مدت، مدل‌سازی.

مقدمه

در حوزه سرمایه‌گذاری پیش‌بینی آینده، موضوعی بسیار مهم است. پیش‌بینی آینده به سرمایه‌گذاران کمک می‌کند تا در خصوص سرمایه‌گذاری‌ها و کاهش ریسک تصمیم‌های بهتری اتخاذ کنند (اومر، آوایس و موزامول^۱، ۲۰۱۹). پیش‌بینی اندازه دقیق حرکت سهم بسیار دشوار است و بسیاری از پژوهشگران به بررسی روش‌هایی می‌پردازند که فقط جهت حرکت سهام را پیش‌بینی می‌کند.

در حال حاضر مدل‌های هوش مصنوعی^۲ از جدیدترین روش‌هایی هستند که برای پیش‌بینی روند حرکتی سهام استفاده می‌شوند. یادگیری ماشینی^۳ شاخه‌ای از هوش مصنوعی است که رایانه‌ها را قادر می‌سازد تا از داده‌ها بیاموزند و پیش‌بینی‌ها و تصمیم‌هایی را بر اساس آن داده‌ها اتخاذ کنند (یائو و ژنگ^۴، ۲۰۲۳). شبکه عصبی مصنوعی یکی از بهترین الگوریتم‌های زیر مجموعه یادگیری ماشینی است. شبکه عصبی مصنوعی مدل‌های محاسباتی‌ای هستند که از ساختار و عملکرد مغز انسان الهام گرفته شده‌اند و هدف آن‌ها تقلید از قابلیت‌های یادگیری مغز است (تورات، پاندیت و بالوت^۵، ۲۰۲۲).

یکی از موضوعاتی که توجه پژوهشگران را در حوزه شبکه‌های عصبی مصنوعی جلب کرده است، افزایش دقت این مدل است. یکی از روش‌های بهبود دقت شبکه‌های عصبی مصنوعی، استفاده از رویکردهای مهندسی استخراج^۶ است که متدولوژی‌های مختلفی را برای بهینه‌سازی فرایندها و افزایش شفافیت در برمی‌گیرد. یکی از رویکردهای مهندسی استخراج، ترکیب شبکه‌های عصبی متفاوت در یک شبکه جامع است (هرندی و همکاران^۷، ۲۰۲۴). نتایج مدل‌های ترکیبی (تشکیل شده از شبکه‌هایی عصبی متفاوت) در تحقیقات مختلف، نشان‌دهنده افزایش دقت این روش به نسبت شبکه‌های عصبی ترکیب نشده است (وان، بای، لو و دینگ^۸، ۲۰۲۲).

در حوزه مالی نشان داده شده است که روش‌های نوین پیش‌بینی مبتنی بر یادگیری ماشینی و هوش مصنوعی به نسبت روش‌های سنتی، دقت بیشتری دارد (گوتز، گورتلر و ویتووسکی^۹، ۲۰۲۳). یک شبکه عصبی هیبرید متشکل از شبکه عصبی کانولوشن (CNN)^{۱۰} و حافظه طولانی کوتاه‌مدت (LSTM)^{۱۱} برای استخراج ویژگی‌های زمانی، در حوزه‌های مختلفی چون تشخیص اثر انگشت (شانموگاپریا^{۱۲}، ۲۰۲۳)، پیش‌بینی سرعت باد (جیانگ، ژانگ، گنگ و لیو^{۱۳}، ۲۰۲۳) و پیش‌بینی روندهای بازار سهام (هائو و ژائو^{۱۴}، ۲۰۲۰) استفاده شده است که این پژوهش‌ها، دقت بهتر شبکه

1. Umar, Awais & Muzammul
2. Artificial Intelligence Algorithms
3. Machine Learning
4. Yao & Zheng
5. Thorat, Pandit & Balote
6. Extraction engineering approaches
7. Harandi
8. Wan, Bai, Lu & Ding
9. Götze, Gürtler & Witowski
10. Convolutional Neural Network
11. Long Short-Term Memory
12. Shanmugapriya
13. Jiang, Zhang, Geng & Liu
14. Hao & Gao

عصبی هیبرید متشکل از کانولوشن و حافظه طولانی کوتاه‌مدت را نشان می‌دهد. همچنین شبکه عصبی چندفیلتره^۱، شبکه عصبی کانولوشن و شبکه عصبی بازگشتی (RNN)^۲ را ترکیب می‌کند تا باعث افزایش عملکرد آن شود (هوانگ، زنگ، ژانگ و زو^۳، ۲۰۲۰). با توجه به اینکه پژوهش‌ها نشان داده‌اند که این روش‌ها نسبت به شبکه عصبی کانولوشن، حافظه طولانی کوتاه‌مدت و شبکه عصبی بازگشتی به تنهایی عملکرد بهتری دارد، این پژوهش قصد دارد تا با استفاده از مدل‌های هیبریدی و چندفیلتره مبتنی بر استخراج ویژگی زمانی چندگانه به‌طور هم‌زمان، به پیش‌بینی روند شاخص بورس اوراق بهادار تهران بپردازد.

پیشینه نظری پژوهش

پیش‌بینی فرایند استفاده از داده‌های تاریخی و فعلی برای استنتاج روند آینده پدیده‌ها از طریق روش‌های علمی است (ساتو، ماتسوبایاشی و تودا^۴، ۲۰۲۲). یکی از موضوعات مهم در خصوص پیش‌بینی در حوزه مالی دقت آن است؛ به‌طوری که الگوریتم‌های پیش‌بینی بر اساس دقت پیش‌بینی و دقت ارزیابی مورد مقایسه قرار می‌گیرند (ائوسو ادجی، هایفرون آکوا، فریمپونگ و عبدالسلام^۵، ۲۰۲۳).

پیش‌بینی بازار سهام به‌دلیل عواملی چون نوسان‌ها (دهاکار و شیوانی^۶، ۲۰۲۳) و غیرخطی بودن رفتار بازار سهام (علی، خان، الشنبیری و الباغوری^۷، ۲۰۲۳) پیچیده‌اند. وجود نویز و نوسان‌های ناگهانی تشخیص الگوهای بازار را پیچیده‌تر می‌کند. این موارد تکنیک‌های پیش‌بینی سنتی را برای به‌کارگیری مؤثر، چالش‌برانگیز می‌کند. همان‌طور که استاموس^۸ (۲۰۲۲) اشاره می‌کند، به‌دلیل پیچیدگی پیش‌بینی بازار سهام، به بررسی حجم زیادی از داده نیاز داریم که این موضوع نشان‌دهنده اهمیت استفاده از تکنیک‌های پیشرفته مانند یادگیری ماشینی و یادگیری عمیق^۹ برای استخراج اطلاعات مرتبط برای پیش‌بینی دقیق قیمت سهام است.

روش‌های سنتی پیش‌بینی بازار سهام، مانند تحلیل بنیادی و تحلیل تکنیکال، در کنار رویکردهای جدیدتر مانند یادگیری ماشینی به‌طور گسترده مورد مطالعه قرار گرفته است (ژی^{۱۰}، ۲۰۲۴). تحلیل بنیادی شامل ارزیابی وضعیت مالی یک شرکت برای تعیین ارزش ذاتی آن با استفاده از صورت‌های مالی است، در حالی که تحلیل تکنیکال بر حرکات تاریخی قیمت و حجم معاملات برای پیش‌بینی روند قیمت‌ها در آینده تمرکز دارد. با این حال روش‌های سنتی به‌دلیل ماهیت پویا و پیش‌بینی‌ناپذیر بازارهای سهام با محدودیت‌هایی مواجهند (دسوال و کومار^{۱۱}، ۲۰۲۳).

1. Multifilter Neural Network
2. Recurrent Neural Networks
3. Huang, Zeng, Zhang & Xu
4. Sato, Matsubayashi & Toda
5. Owusu-Adjei, Hayfron-Acquah, Frimpong & Abdul-Salaam
6. Dhakar & Shiwani
7. Ali, Khan, Alshanbari & El-Bagoury
8. Stamos
9. Deep Learning
10. Xie
11. Deswal & Kumar

روش‌های نوین پیش‌بینی بازار سهام عبارت‌اند از: استفاده از تکنیک‌های یادگیری ماشینی مانند حافظه طولانی کوتاه‌مدت و یادگیری شبکه عصبی بازگشتی (سینق و مارکانده^۱، ۲۰۲۳). این مدل‌ها در شناسایی الگوهای زمانی و روابط در داده‌های بازار سهام و بهبود تصمیم‌گیری سرمایه‌گذاری با تجزیه و تحلیل داده‌های بازار سهام در دنیای واقعی در مقایسه با روش‌های سنتی برترند (صبور، حسین، عقلی، لی و کومار^۲، ۲۰۲۳).

یادگیری ماشینی زیر مجموعه‌ای از هوش مصنوعی است که رایانه‌ها را قادر می‌سازد تا از مثال‌ها یاد بگیرند و دانش کسب شده از مثال‌ها را در موقعیت‌های جدید به کار برند (هان^۳، ۲۰۱۹). یادگیری ماشینی با افزایش دقت پیش‌بینی و ترکیب منابع داده‌های مختلف، در پیش‌بینی بازارهای مالی نقش مهمی بازی می‌کند. در پژوهش‌های مربوط به این حوزه، تلاش شده است تا با استفاده از مدل‌های مبتنی بر مجموعه ترکیبی از تجزیه و تحلیل احساسات از رسانه‌های اجتماعی با استفاده از داده‌های مالی (الریدهاوی و العثمان^۴، ۲۰۲۳)، حافظه طولانی کوتاه‌مدت و شبکه عصبی بازگشتی (روت، بار، ساها و چاودهوری^۵، ۲۰۲۲) و مدل ARIMA^۶ برای تجزیه و تحلیل داده‌های سری زمانی (آداگالا و همکاران^۷، ۲۰۲۳)، دقت پیش‌بینی‌های این حوزه افزایش یابد.

شبکه عصبی یکی از الگوریتم‌های یادگیری ماشینی بسیار مشهور است. در حوزه پیش‌بینی اقتصادی، تکنیک‌های شبکه‌های عصبی برای پیش‌بینی و پیشگیری بحران‌های اقتصادی به کار گرفته می‌شود که پتانسیل شبکه‌های عصبی را به عنوان ابزاری قدرتمند در پیشگیری از بحران نشان می‌دهد (چیریتا^۸، ۲۰۱۲). این شبکه‌ها از گره‌های^۹ به هم پیوسته تشکیل شده‌اند که در لایه‌های مختلفی قرار می‌گیرند، جایی که اطلاعات پردازش می‌شوند و از طریق اتصالات وزن‌دهی شده برای پیش‌بینی یا طبقه‌بندی انتقال می‌یابند (کروئلی^{۱۰}، ۲۰۲۱). شبکه‌های عصبی با تقلید از توانایی پردازش مغز، در وظایفی مانند بینایی رایانه، تشخیص گفتار و پردازش زبان طبیعی، با استفاده از یادگیری با نظارت، برتری زیادی دارند (دوکار و شارما^{۱۱}، ۲۰۲۳). علاوه بر این، شبکه‌های عصبی می‌توانند الگوهای پیچیده و مشکلات پیش‌بینی را با پردازش داده‌ها به شیوه‌ای توزیع شده و موازی^{۱۲}، شبیه به قابلیت پردازش اطلاعات مغز انسان مدل‌سازی کنند (یوسینسکی، کلون، نگوین، فوکس و لیبسون^{۱۳}، ۲۰۱۵).

ورودی‌های شبکه عصبی را داده‌ها به صورت مجموعه‌ای از اعداد تشکیل می‌دهند. داده‌های وارد شده به شبکه عصبی بعد از انجام دسته‌ای از فرایندها به چندین (یا یک) خروجی تبدیل می‌شوند. نورون‌ها در یک شبکه عصبی،

1. Singh & Markande
2. Saboor, Hussain, Agbley, Li & Kumar
3. Hahn
4. Al Ridhawi & Al Osman
5. Rout, Bar, Saha & Chaudhuri
6. Autoregressive Integrated Moving Average
7. Addagalla
8. Chiriță
9. Nodes
10. Crowley
11. Devkar & Sharma
12. Distributed and parallel manner
13. Yosinski, Clune, Nguyen, Fuchs & Lipson

کوچک‌ترین واحد پردازش اطلاعات هستند. شبکه عصبی سه لایه دارد که نورون‌های ورودی، پنهان و خروجی در آن‌ها قرار می‌گیرند. در دو لایه نورون‌های پنهان و خروجی، پردازش اطلاعات انجام می‌شود و در لایه نورون‌های ورودی دریافت اطلاعات به‌وقوع می‌پیوندد (یونگ، اندروز، کوپر و لو^۱، ۲۰۲۰). شبکه‌های عصبی دولایه نیز وجود دارند که در آن‌ها لایه نورون‌های حذف شده است. در شبکه‌های عصبی، داده‌ها قبل از ورود به لایه‌های نورون‌های پنهان و خروجی، وزن‌دهی می‌شوند. وزن‌دهی به داده‌های ورودی در شبکه‌های عصبی، برای شناسایی سوگیری‌های داده بسیار مهم است (شو، یوان، منگ و زو^۲، ۲۰۲۳).

شبکه عصبی بازگشتی، نوعی شبکه عصبی مصنوعی است که برای تشخیص الگوها در توالی داده‌ها، مانند سری‌های زمانی، گفتار، متن و ویدئو طراحی شده است. شبکه عصبی بازگشتی به این صورت عمل می‌کند که ورودی‌های جدید به همراه نورون‌های لایه‌های قبلی جهت یادگیری به کار گرفته می‌شوند (کانگ‌چیدامبارسن، روالی، بانرجی و پراکاش^۳، ۲۰۲۱). شبکه‌های عصبی بازگشتی با استفاده از روشی به نام BPTT^۴ آموزش داده می‌شوند که شبکه را در طول زمان باز می‌کند. اگرچه این فرایند می‌تواند با ناپدید شدن گرادیان^۵ به چالش کشیده شود (مارهون، کامرون و کرمر^۶، ۲۰۱۳). انواع شبکه‌های عصبی بازگشتی، مانند شبکه‌های حافظه طولانی کوتاه‌مدت برای حل این مسئله با ترکیب سازوکارهایی برای درک بهتر وابستگی‌های طولانی مدت توسعه یافته‌اند (کابسا و سیگلمن^۷، ۲۰۱۱).

حافظه طولانی کوتاه‌مدت نوعی از شبکه عصبی بازگشتی است که برای مدل‌سازی توالی‌های زمانی و وابستگی‌های دوربرد^۸ طراحی شده است و از نسبت به شبکه‌های عصبی بازگشتی معمولی دقت بیشتری دارد. هاجریتر و اشمیدهوربر^۹، حافظه طولانی کوتاه‌مدت را نخستین بار در سال ۱۹۹۷ معرفی کردند تا مشکل جریان بازگشتی خطا در شبکه‌های عصبی بازگشتی سنتی را رفع کنند (هاجریتر و اشمیدهوربر، ۱۹۹۷). شبکه‌های عصبی بازگشتی توانایی خود را در حوزه‌های مختلفی مانند پیش‌بینی سری زمانی در حوزه مالی نشان داده‌اند (اگوس^{۱۰}، ۲۰۱۹)، شبکه عصبی بازگشتی به دلیل توانایی در حفظ حافظه بلندمدت و روی هم قرار دادن لایه‌های پنهان مکرر برای یادگیری ویژگی‌های زمانی سطح بالاتر در یادگیری توالی با الگوهای بلندمدت مؤثر است (موجیکا، ویلاسنور، آلانیس و آرانا دانیل^{۱۱}، ۲۰۱۹).

پیشینه تجربی پژوهش

در زمینه پیش‌بینی بازارهای مالی با استفاده از شبکه‌های عصبی، حافظه طولانی کوتاه‌مدت و شبکه عصبی بازگشتی

1. Young, Andrews, Cooper & Lu
2. Shu, Yuan, Meng & Xu
3. Kanagachidambaresan, Ruwali, Banerjee & Prakash
4. Backpropagation through time
5. Gradient
6. Marhon, Cameron & Kremer
7. Cabessa & Siegelmann
8. Long-range dependencies
9. Hochreiter & Schmidhuber
10. Ogus
11. Mojica, Villaseñor, Alanis & Arana-Daniel

تحقیقات مختلفی صورت گرفته است. در این بخش به بررسی تحقیقات انجام شده در این خصوص پرداخته خواهد شد. انریکو، گوستیانا و کریشنا^۱ (۲۰۲۳) به پیش‌بینی قیمت سهام JPFA^۲ (یکی از شرکت‌های پیشرو در صنایع غذایی در اندونزی) با استفاده از الگوریتم حافظه طولانی کوتاه‌مدت می‌پردازند. این پژوهش نشان می‌دهد که مدل بهینه‌سازی آدام^۳ با ۱۰۰ دوره بالاترین دقت را به دست آورده (۹۸/۴۴ درصد)، در حالی که بهینه‌سازی گرادیان تصادفی (SGD)^۴ با ۲۵ دوره، کمترین دقت را داشته است (۹۲/۰۴ درصد).

فجلستروم^۵ (۲۰۲۲) به بررسی استفاده از شبکه‌های عصبی حافظه طولانی کوتاه‌مدت برای پیش‌بینی حرکات قیمت سهام می‌پردازد و از توانایی آن برای ترکیب اطلاعات گذشته برای داده‌های سری زمانی استفاده می‌کند. در این پژوهش از مجموعه‌ای از شبکه‌های حافظه طولانی کوتاه‌مدت مستقل و موازی برای بهبود دقت پیش‌بینی و کاهش انحراف در نتایج استفاده شده است. این پژوهش بر شاخص استکهلم تمرکز می‌کند. نتایج این پژوهش نشان می‌دهد که پرتفوی مبتنی بر حافظه طولانی کوتاه‌مدت، نوسان‌های کمتری از خود نشان می‌دهد که به نسبت ریسک به بازده بالاتر منجر می‌شود.

پاتیل، شانکار، دیوی، سینگ و آپادحیای^۶ (۲۰۲۳) به پیش‌بینی بازار سهام با استفاده از شبکه‌های عصبی مبتنی بر حافظه طولانی کوتاه‌مدت می‌پردازند و با استفاده از معیارهای کلیدی مانند خطای میانگین مربعات ریشه (RSME)^۷ مدل مورد استفاده را ارزیابی می‌کنند. نتایج این پژوهش نشان می‌دهد که دقت پیش‌بینی مدل مورد استفاده ۸۸/۲ درصد است که اثربخشی بالقوه آن را در پیش‌بینی بازار سهام نشان می‌دهد.

کوه‌ساسی، هایادی، جولیانندی و پریادی^۸ (۲۰۲۲) با استفاده از حافظه طولانی کوتاه‌مدت و شبکه عصبی بازگشتی به تجزیه و تحلیل احساسات در اخبار مربوط به حوزه مالی می‌پردازند. نتایج این مدل حاکی از دقت ۹۱/۵۴ درصدی آن است که اثربخشی این مدل در درک روندها و نظرهای مالی را نشان می‌دهد.

ساکو، مپیندا و رودریگس^۹ (۲۰۲۲) با استفاده از شبکه‌های عصبی بازگشتی، حافظه طولانی کوتاه‌مدت و واحدهای دروازه بازگشتی (GRU)^{۱۰}، به پیش‌بینی قیمت‌های پایانی هشت شاخص بازار سهام و شش نرخ مبادله ارز مرتبط با آمریکا می‌پردازند. نتایج این پژوهش نشان می‌دهد که مدل‌های واحدهای دروازه بازگشتی، به‌ویژه برای پیش‌بینی تک‌متغیره خارج از نمونه برای نرخ ارز و پیش‌بینی چندمتغیره خارج از نمونه برای شاخص بازار سهام، بهترین نتایج کلی را ارائه می‌کنند.

1. Enriko, Gustiyana & Krishna
2. Japfa Comfeed
3. Adam
4. Stochastic Gradient Descent
5. Fjellström
6. Patil, Shankar, Devi, Singh & Upadhyay
7. Root Mean Squared Error
8. Kohsasih, Hayadi, Juliandy & Pribadi
9. Sako, Mpinda & Rodrigues
10. Gate Recurrent Unit

حاجی‌زاده طحان، قاسم‌زاده و رضاییان (۱۳۹۹) برای غلبه بر مشکل کافی نبودن توانایی حافظه طولانی کوتاه‌مدت در اختصاص درجه‌های مختلف توجه به ویژگی‌های زیرپنجره در چند مرحله زمانی، از رویکرد حافظه طولانی کوتاه‌مدت عمیق مبتنی بر توجه تکاملی برای پیش‌بینی سری‌های زمانی چند متغیره استفاده کردند که به‌طور خودکار، یکی از بهترین ترکیب‌ها از مقادیر پارامترهای حافظه طولانی کوتاه‌مدت و وزن توجه به ویژگی‌ها را می‌یابد. نتایج این پژوهش نشان می‌دهد که الگوریتم پیشنهادی، در مقایسه با سایر مدل‌های پایه‌ای از عملکرد بهتری برخوردار است.

عبدی، مرادزاده فرد، احمدزاده و خدام (۱۴۰۰) به بررسی مدلی ترکیبی جهت بهینه‌سازی سبد سهام با استفاده از شبکه عصبی بازگشتی و حافظه طولانی کوتاه‌مدت، به کمک محدودیت‌های کاردینالیتهی تصمیم‌گیری چند معیاره طی دوره خرداد ۱۳۹۵ تا خرداد ۱۴۰۰ می‌پردازند. نتیجه پژوهش آن‌ها نشان‌دهنده بازدهی بهتر سبد انتخاب شده با این روش به نسبت سبد انتخاب شده با روش‌های سنتی است.

شریف‌فر، خلیلی عراقی، رئیسی وانانی و فلاح شمس (۱۴۰۱) توانایی معماری‌های مختلف الگوریتم CNN برای پیش‌بینی قیمت سهام را بررسی کردند. آن‌ها با ۵۴ بار تکرار الگوریتم با معماری و پارامترهای متفاوت و با استفاده از دو دسته اصلی داده‌های ورودی شامل اطلاعات قیمتی روزانه سهام و ده شاخص منتخب تکنیکال برای سهام شرکت ذوب آهن اصفهان نشان می‌دهند که استفاده از CNN همراه با لایه ادغام بیشینه، دقتی به اندازه ۷۹/۱ درصد دارد که گویای عملکرد بهتر آن نسبت به سایر معماری‌ها و الگوریتم‌های RNN است.

حیدری و امیری (۱۴۰۱) پس از جمع‌آوری داده‌های ۱۵۰ شرکت بزرگ پذیرفته‌شده در بورس اوراق بهادار تهران، از سال ۱۳۹۰ تا ۱۳۹۹، با تنظیم دقیق روش‌های یادگیری ماشین برای هر یک از سهام، به پیش‌بینی روند قیمت سهام و صحت‌سنجی هر یک از روش‌های مدل‌های خطی، مدل‌های خودهم‌بسته، جنگل تصادفی و شبکه‌های عصبی پرداختند. مدل‌های مبتنی بر یادگیری عمیق نسبت به سایر مدل‌ها عملکرد بهتری از خود نشان می‌دهند و در پیش‌بینی روند کوتاه‌مدت قیمت سهام، از دقتی حدود ۷۰ تا ۸۰ درصد برخوردارند. همچنین، مدل‌های یادگیری کم‌عمق دقت بالاتری داشتند. به‌طور کلی، بیشتر مدل‌ها در پیش‌بینی روندهای منفی سهام، عملکرد بهتری نشان می‌دهند.

بهشتی مسئله‌گو، افشار کاظمی، حقیقت منفرد و رضاییان (۱۴۰۲) به بررسی روشی جهت پیش‌بینی قیمت روز آینده سهام بر اساس ساختار شبکه عصبی عمیق با استفاده از داده‌های قیمت، مجموعه‌ای از شاخص‌های فنی و سر تیتراخبار به‌عنوان ورودی مدل می‌پردازند. آن‌ها در این پژوهش از داده‌های شاخص داوجونز و داده‌های خبری کانال ردیت استفاده می‌کنند. آن‌ها ورودی‌ها را به شبکه حافظه طولانی کوتاه‌مدت وارد می‌کنند. نتایج پژوهش آن‌ها نشان می‌دهد که مدل حافظه طولانی کوتاه‌مدت با دقت ۶۹/۱۹ درصد، بیشترین دقت پیش‌بینی را دارد و داده‌های خبری با ۶۵/۶۲ درصد و داده‌های عددی با ۵۱/۸۹ درصد دقت در رتبه‌های بعدی قرار می‌گیرند.

زیادی، صلواتی و لطفی هروی (۱۴۰۲) با استفاده از الگوریتم بر پایه شبکه‌های عصبی بازگشتی و الگوریتم LSTM با توجه به قابلیت حفظ اطلاعات گذشته، مدلی جهت پیش‌بینی سری‌های زمانی ارائه کردند. در هر دو دسته از سری‌های زمانی تک‌متغیره و چندمتغیره، از معماری stacked-LSTM استفاده شده است. این پژوهش نشان می‌دهد که

قیمت طلا، قیمت ارز، شاخص بهای کالا و خدمات و همچنین حجم نقدینگی، بیشترین همبستگی را با قیمت مسکن داشته‌اند. با استفاده از داده‌های این شاخص‌های اقتصادی، پیش‌بینی‌هایی با دقت‌های زیاد به‌دست آمده است.

طهرانی یزدی، واعظی، ستایشی و رئیسی وانانی (۱۴۰۳) با استفاده از هوش مصنوعی و یادگیری ماشین تعداد حوادث و رخداد‌های روزانه با خسارت‌های جانی را در بیمه نوع شخص ثالث برآورد کرده‌اند. خروجی این مدل، ترکیب تعداد و زمان رخداد حوادث، نقشی مهم در محاسبه تعهدات آتی و ذخیره خسارت در این نوع بیمه دارد. در این پژوهش، از الگوریتم شبکه عصبی حافظه کوتاه‌مدت بلند (LSTM) استفاده شده است. این الگوریتم از توانایی بالای پیش‌بینی در سری‌های زمانی برخوردار است. این مدل برتری جست‌وجوی شبکه‌ای را نشان می‌دهد.

باجلان، فلاح‌پور و رئیسی (۱۴۰۳) به بررسی ریسک اعتباری پرتفوی تسهیلات بانکی با رویکرد اکچوئری پرداخته و با استفاده از شبکه عصبی پرسپترون و محدودیت‌های بانک‌ها در ارائه تسهیلات، ترکیب بهینه پرتفوی اعتباری را تعیین کردند. نمونه مورد استفاده شامل تسهیلات اعطایی بانک به ۲۸۰ مشتری کلان در ۴ بخش صنعتی، خدماتی، کشاورزی و ساختمان در سال ۱۳۹۲ بوده است. نتیجه پژوهش نشان می‌دهد که استفاده از مدل اکچوئری برای تعیین ریسک اعتباری و بهینه‌سازی با استفاده از شبکه عصبی، به بهبود بهینه‌سازی پرتفوی اعتباری بانک‌ها منجر می‌شود.

با توجه به مباحث مطرح شده، سؤال‌های پژوهش به‌شکل زیر مطرح می‌شود:

۱. آیا شبکه‌های عصبی یادگیری عمیق هیبریدی مبتنی بر استخراج ویژگی مقیاس زمانی چندگانه، در بورس اوراق بهادار تهران قدرت پیش‌بینی دارند؟
۲. آیا استفاده از شبکه‌های عصبی یادگیری عمیق هیبریدی مبتنی بر استخراج ویژگی مقیاس زمانی چندگانه در مقایسه با سایر شبکه‌های عصبی (شبکه‌های عصبی هیبریدی ساده^۱، CNN و LSTM) باعث بهبود میزان دقت پیش‌بینی می‌شود؟

روش‌شناسی پژوهش

در این پژوهش با استفاده از شبکه عصبی هیبریدی شامل دو رویکرد شبکه عصبی بازگشتی و شبکه عصبی حافظه طولانی کوتاه‌مدت مبتنی بر استخراج ویژگی زمانی چندگانه، به پیش‌بینی روند شاخص کل بورس اوراق بهادار تهران پرداخته شده است. در ابتدا به‌واسطه لایه‌های اول و دوم شبکه عصبی کانولوشن، به استخراج دو ویژگی در مقیاس‌های زمانی متفاوت پرداخته می‌شود. در ادامه با استفاده از حافظه طولانی کوتاه‌مدت، به شناسایی وابستگی‌های زمانی دو ویژگی استخراج شده و یک ویژگی خام پرداخته خواهد شد. در نهایت با استفاده از لایه‌های کاملاً مرتبط به یادگیری ویژگی‌های مشترک جهت پیش‌بینی روند شاخص بورس اوراق بهادار تهران پرداخته خواهد شد. برای انجام این پژوهش، از داده‌های آماری مربوط به شاخص قیمت و بازده نقدی مربوط به شاخص بورس اوراق بهادار تهران، طی سال‌های ۱۳۷۷ تا ۱۴۰۱ و برای تحلیل داده‌ها، نرم‌افزار پایتون استفاده شده است.

مراحل پژوهش با توجه به پژوهش هائو و گائو (۲۰۲۰) در چهارده مرحله شرح داده خواهد شد. در مرحله اول به استخراج داده‌ها پرداخته می‌شود (در این پژوهش داده‌های روزانه ۲۵ ساله شاخص کل بورس اوراق بهادار تهران استفاده شده است). در مرحله دوم که مرحله پنجره غلتان^۱ نامیده می‌شود، برای ایجاد نمونه براساس مقدار پایانی شاخص کل در هر نمونه، ۴۰ داده متوالی شاخص کل قرار داده شده است:

$$X_i = \{x_{i+1}, x_{i+2}, \dots, x_{i+t}, \dots, x_{i+40}\} \quad \text{رابطه ۱}$$

که در آن x_{i+t} مقدار پایانی شاخص کل در روز $i + t$ ام است. مرحله سوم، مرحله نرمال‌سازی^۲ داده‌هاست. نرمال‌سازی داده‌ها با استفاده از رابطه ۲ انجام می‌پذیرد تا مقادیر روزانه شاخص کل در بازه $[0, 1]$ قرار گیرند.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \quad \text{رابطه ۲}$$

در این رابطه، X_{max} بالاترین مقدار روزانه شاخص؛ X_{min} کمترین مقدار روزانه شاخص و X مقدار پایانی شاخص کل است.

در مرحله چهارم، داده‌ها لیبیل‌گذاری^۳ می‌شوند. در این مرحله، روند شاخص کل در ۵ روز معاملاتی در هفته یا ۲۰ روز معاملاتی در ماه بررسی می‌شود. در صورت صعودی بودن بازار، این روند با عدد ۱ و در صورت نزولی بودن آن با عدد ۰ نشان داده می‌شود. برای انجام این کار از رابطه ۳ استفاده می‌شود:

$$\gamma_i = \begin{cases} 1 & x_{i+40} \leq x_{i+40+n} \\ 0 & x_{i+40} > x_{i+40+n} \end{cases} \quad \text{رابطه ۳}$$

که در این رابطه، x_{i+40} مقدار پایانی شاخص در روز $i + 40$ و x_{i+40+n} مقدار پایانی شاخص در روز $i + 40 + n$ است. n همان ۵ روز معاملاتی در هفته یا ۲۰ روز معاملاتی در ماه است.

با توجه به پیشینه پژوهش، شبکه‌های عصبی بازگشتی در خصوص سری‌های زمانی، برای استخراج ویژگی‌های بلندمدت دچار ناپدید شدن یا انفجار گرادیان می‌شوند، به همین دلیل از LSTM برای رفع این مشکل استفاده می‌شود. در مرحله پنجم، برای بهینه‌سازی شبکه عصبی، از تابع بهینه‌سازی کلی گرادیان نزولی تصادفی استفاده می‌شود. در پیاده‌سازی کلی گرادیان نزولی تصادفی، ابتدا بردار پارامترها را که برداری شامل تمام پارامترهای تابع هزینه است، θ می‌نامیم. θ برابر برداری دل‌خواه قرار می‌گیرد؛ سپس برای هر بار به‌روزرسانی این بردار یک عضو از مجموعه داده‌های آموزشی، به‌صورت تصادفی انتخاب شده و با نرخ α ضرب می‌شود و بردار حاصل از گرادیان تابع هزینه در نقطه θ از θ کم می‌شود:

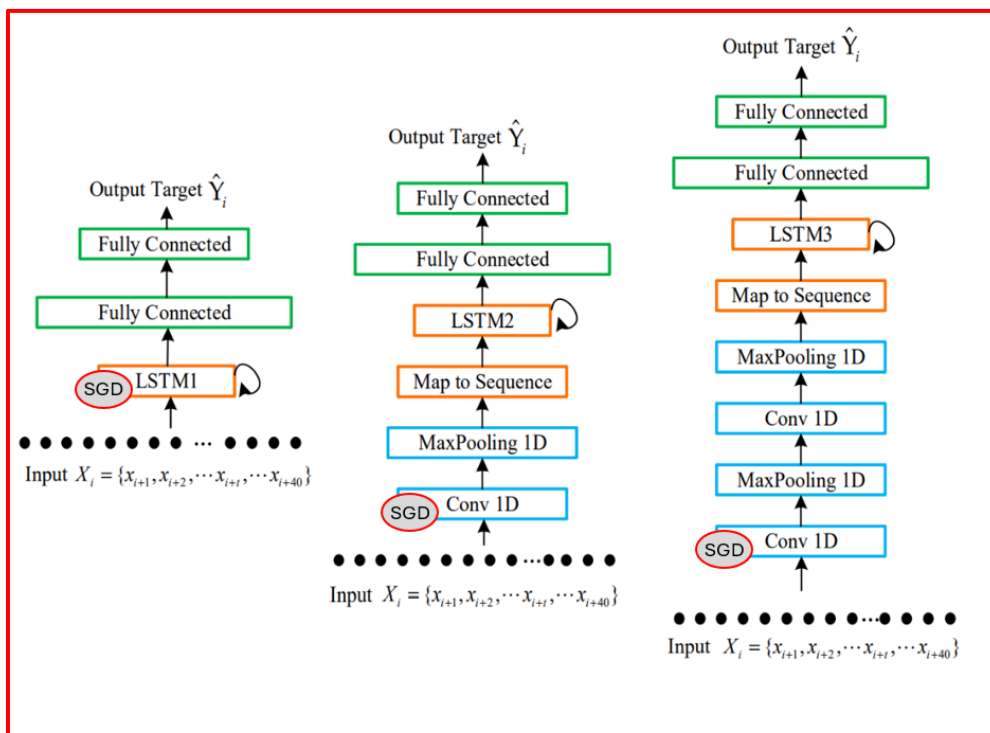
1. Rolling Window
2. Normalization
3. Labling

$$\theta = \theta - \alpha \nabla_{\theta} J_i(\theta; x^{(i)}, y^{(i)}) \tag{رابطه ۴}$$

J تابع هزینه و $x^{(i)}, y^{(i)}$ یک عضو از داده‌های آموزشی است که به صورت تصادفی انتخاب شده است. $J_i(\theta; x^{(i)}, y^{(i)})$ نشان‌دهنده جمله i ام از جمله‌های تابع هدف است. α نرخ است که با آن θ به‌روزرسانی می‌شود و مقداری تجربی دارد که اگر خیلی کوچک باشد، زمان رسیدن به هم‌گرایی را طولانی می‌کند و اگر خیلی بزرگ باشد، ممکن است هم‌گرایی رخ ندهد.

در پیاده‌سازی دیگر در هر حلقه عضوی تصادفی از مجموعه داده‌ها انتخاب نمی‌شود؛ بلکه در هر حلقه کل مجموعه داده‌ها یک بار به صورت تصادفی بازچینی می‌شود؛ سپس عملیات به‌روزرسانی به ترتیب به ازای J_1, J_2, \dots, J_n انجام می‌شود که n نشان‌دهنده اندازه مجموعه داده‌های آموزشی است.

همان طور که پیش‌تر اشاره شد، معمولاً عملیات به‌روزرسانی برای J حاصل از یک تک عضو مجموعه، داده‌های آموزشی انجام نمی‌شود و برای زیرمجموعه‌ای از این داده‌ها انجام می‌شود که به آن دسته کوچک می‌گویند. این تابع از اجزای اصلی هریک از مدل‌های شبکه عصبی بوده و هریک از مدل‌ها، چه به صورت تکی و چه به صورت هیبریدی شامل این تابع است. شبکه عصبی حافظه طولانی کوتاه‌مدت، کانولوشن و ترکیبی از این شبکه‌های عصبی دارای بهینه‌سازی کلی گرادینان نزولی تصادفی است که در خود ساختار شبکه قبل از ورود به مراحل بعدی از آن استفاده می‌شود.



شکل ۱. بهینه‌سازی هایپرپارامترهای مدل

در هر LSTM وابستگی‌های زمانی از سری ویژگی متناظر خود استخراج می‌شود. هر سلول سه دروازه اصلی دارد: دروازه ورودی، دروازه فراموشی و دروازه خروجی. فرض کنید بردار ویژگی ورودی در زمان t ، $f v_t$ بوده و گزاره فراموشی در گام زمانی قبلی h_{t-1} باشد. i_t به‌عنوان دروازه ورودی شناخته می‌شود و برای محاسبه آن از رابطه ۵ استفاده می‌شود.

$$i_t = \text{sigmoid}(W_{ifv} f v_t + W_{ih} h_{t-1} + b_i) \quad (\text{رابطه ۵})$$

f_t دروازه فراموشی است و برای محاسبه آن از رابطه ۶ استفاده می‌شود.

$$f_t = \text{sigmoid}(W_{ffv} f v_t + W_{fh} h_{t-1} + b_f) \quad (\text{رابطه ۶})$$

o_t دروازه خروجی است و برای محاسبه آن از رابطه ۷ استفاده می‌شود.

$$o_t = \text{sigmoid}(W_{ofv} f v_t + W_{oh} h_{t-1} + b_o) \quad (\text{رابطه ۷})$$

کارکرد حافظه بدین صورت است که اطلاعات جدید از طریق دروازه ورودی وارد و اطلاعات گذشته از طریق دروازه فراموشی فراموش می‌شود. اطلاعات گذشته با c_{t-1} نشان داده شده است. آخرین اطلاعات به‌کمک رابطه ۸ محاسبه می‌شود.

$$\tilde{c}_t = \text{tanh}(W_{cfv} f v_t + W_{ch} h_{t-1} + b_c) \quad (\text{رابطه ۸})$$

اطلاعات ذخیره شده در واحد حافظه، با استفاده از رابطه ۹ به‌روزرسانی می‌شود.

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (\text{رابطه ۹})$$

در نهایت، خروجی LSTM به‌صورت رابطه ۱۰ نشان داده می‌شود.

$$h_t = o_t \odot \text{tanh}(c_t) \quad (\text{رابطه ۱۰})$$

از h_t به‌عنوان نماینده ویژگی‌های زمانی استخراج شده توسط سلول LSTM استفاده می‌شود.

در مرحله ششم، شبکه عصبی کانولوشن نیز همانند LSTM، به‌عنوان بخشی از شبکه عصبی هیبریدی مورد استفاده قرار می‌گیرد و از تابع بهینه‌سازی گرادیان نزولی تصادفی^۱ جهت بهینه‌سازی شبکه عصبی هیبریدی استفاده می‌شود. ویژگی‌های استخراج شده از لایه پیچش، به‌عنوان ورودی لایه گردآوری قرار دارد. در این لایه گردآوری داده‌های ورودی با استفاده از یک نورون، سبب کاهش ابعاد و افزایش سرعت شبکه می‌شود.

پس از وارد کردن داده‌ها به لایه CNN، از لایه نرمال‌سازی دسته‌ای^۲ استفاده شده و خروجی این لایه، وارد دو لایه تمام متصل^۳ شده و به خروجی متصل می‌شود. تابع فعال‌سازی هر کدام از لایه‌های تماماً متصل، تابع واحد خطی یکسوساز نشت‌دار^۴ است.

1. Stochastic Gradient Descent
2. Batch Normalization
3. Fully Connected Layers
4. Leaky ReLU

در مرحله هفتم شبکه عصبی هیبریدی ساده CNN-LSTM مورد استفاده قرار می‌گیرد و مشابه CNN و LSTM در این مرحله نیز جهت بهینه سازی از تابع گرادیان نزولی تصادفی استفاده می‌شود. این مرحله بدین صورت است که پس از وارد کردن داده‌ها به لایه CNN، خروجی این لایه وارد لایه گردآوری بیشترین‌ها و خروجی این لایه به بردارهای متوالی نگاشت می‌شود^۱. ویژگی‌های نگاشت شده وارد لایه LSTM شده و در نهایت، خروجی لایه LSTM، وارد دو لایه تماماً متصل شده و به خروجی وصل می‌شود. تابع ضرر مورد استفاده در تمام مدل‌ها تابع آنتروپی متقاطع باینری است.

در مرحله هشتم، شبکه عصبی هیبریدی ساده CNN-CNN-LSTM استفاده می‌شود. تمام اجزا، توابع و کارکرد این شبکه مشابه شبکه عصبی هیبریدی ساده CNN-LSTM است، با این تفاوت که به ابتدای شبکه، یک لایه CNN و یک لایه گردآوری بیشترین‌ها اضافه شده است. پس شبکه عصبی هیبریدی ساده CNN-CNN-LSTM ترکیبی از دو شبکه CNN و یک شبکه LSTM است. تابع ضرر مورد استفاده در تمام مدل‌ها تابع آنتروپی متقاطع باینری است.

در مرحله نهم، سه شبکه عصبی LSTM، شبکه عصبی هیبریدی ساده CNN-LSTM و شبکه عصبی هیبریدی ساده CNN-CNN-LSTM با یکدیگر ترکیب می‌شوند. برای استخراج ویژگی زمانی چندگانه در این پژوهش، سه مقیاس زمانی مورد بررسی قرار می‌گیرد. مقیاس زمانی کوتاه‌مدت (F_1) که مقادیر روزانه شاخص کل بوده و همان شبکه LSTM است. دو مقیاس زمانی میان‌مدت (F_2) و بلندمدت (F_3) نیز در این پژوهش استفاده شده است. خروجی لایه‌های مختلف CNN نشان‌دهنده سری مقادیر شاخص کل در مقیاس‌های زمانی متفاوت (هفتگی، ماهانه و...) است که ویژگی‌های مقیاس‌های زمانی چندگانه را نشان می‌دهد. از آنجا که در شبکه عصبی هیبریدی ساده CNN-LSTM، از یک لایه CNN استفاده شده است، می‌توان ویژگی مقیاس زمانی میان‌مدت را استخراج کرد. از سوی دیگر، از آنجا که در شبکه عصبی هیبریدی ساده CNN-CNN-LSTM، از دو لایه CNN استفاده شده است، می‌توان ویژگی مقیاس زمانی بلندمدت را استخراج کرد.

برای پیش‌بینی روند شاخص با توجه به ویژگی‌هایی که در مقیاس‌های مختلف (F_1 ، F_2 و F_3) وجود دارد، به تعیین مقیاس‌های زمانی مناسب برای این ویژگی‌ها نیاز داریم. از آنجا که F_1 سری ویژگی‌های استخراج شده از مقادیر روزانه خام شاخص کل است، فقط باید مقیاس‌های زمانی برای ویژگی‌های F_2 و F_3 که با شبکه عصبی کانولوشن (CNN) آموزش داده شده‌اند، تعیین شود. مقیاس‌های زمانی F_2 و F_3 براساس فیلدهای گیرنده که توسط سائز کرنل^۲ و گام‌های maxpooling و عملیات CNN تعیین می‌شوند، عمل می‌کنند.

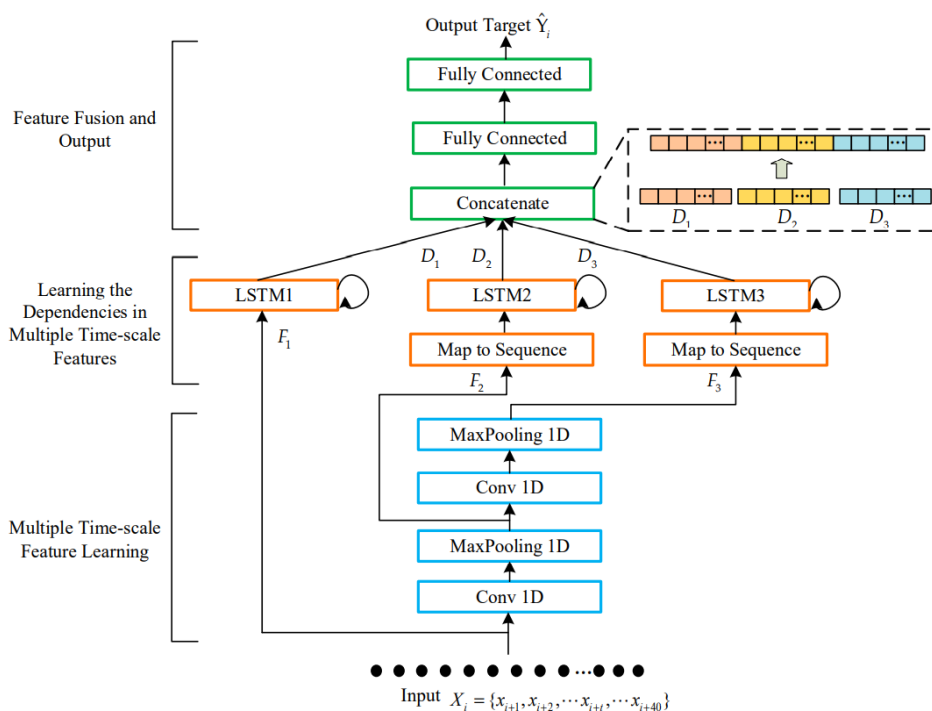
تنظیم سائز کرنل دو لایه CNN برای یافتن مقیاس‌های زمانی بهینه مورد نیاز است. برای پیش‌بینی روند شاخص کل در هفته آینده (پنج روز معاملاتی)، سائز کرنل لایه‌های اول و دوم CNN به ترتیب ۷ و ۵ است. بنابراین، مقیاس‌های زمانی F_2 و F_3 به ترتیب ۸ و ۱۸ روز معاملاتی تنظیم می‌شوند؛ به این معنا که هر کرنل در مقیاس‌های زمانی F_2 و F_3 به ترتیب شامل ۸ و ۱۸ مقدار پایانی شاخص کل در روزهای معاملاتی است.

1. Map to Sequence
2. Kernel

برای پیش‌بینی روند شاخص کل در ماه آینده (۲۰ روز معاملاتی)، سائز کرنل لایه‌های اول و دوم CNN به ترتیب ۹ و ۷ است. در این حالت، مقیاس‌های زمانی F_2 و F_3 به ترتیب ۱۰ و ۲۴ روز معاملاتی تنظیم می‌شوند؛ به این معنا که هر کرنل در مقیاس‌های زمانی F_2 و F_3 به ترتیب شامل ۱۰ و ۲۴ مقدار پایانی شاخص کل در روزهای معاملاتی است. تابع ضروری که در تمامی مدل‌ها استفاده می‌شود، تابع آنتروپی متقاطع باینری است.

در مرحله دهم، برای استخراج وابستگی زمانی از ویژگی‌ها در مقیاس‌های زمانی مختلف، از سه شبکه عصبی LSTM استفاده می‌شود. ماتریس ویژگی دوبعدی که از CNN استخراج شده است، باید از طریق لایه Flattening به یک سری (ماتریس تک‌بعدی) تبدیل شود تا برای LSTM آماده شود. توالی ویژگی‌ها به‌عنوان ورودی به LSTM ارائه می‌شوند. f, v_t بردار ویژگی در سری ویژگی‌ها و t نمایانگر جایگاه در سری است. هر بردار ویژگی از ماتریس دوبعدی ویژگی از چپ به راست به صورت ستونی استخراج شده است؛ یعنی بردار ویژگی، از الحاق ستون t ام از تمام ماتریس‌های ویژگی حاصل شده است.

هر شبکه LSTM وابستگی‌های زمانی را از سری ویژگی متناظر خود استخراج می‌کند که توسط رابطه‌های ۵ تا ۱۰ توضیح داده شد. پس از تکمیل فرایند رابطه‌های ۵ تا ۱۰ توسط بردارهای ویژگی، از خروجی نهایی h_t به‌عنوان نماینده ویژگی‌های زمانی استخراج شده توسط سلول LSTM استفاده می‌شود. این وابستگی‌های زمانی استخراج شده توسط سه شبکه LSTM را با D_1 ، D_2 و D_3 که همگی بردارهای تک‌بعدی هستند، نشان می‌دهیم. با توجه به شکل ۲ می‌توان مدل شبکه عصبی هیبریدی پیشنهاد شده را مشاهده کرد.



شکل ۲. مدل شبکه عصبی هیبریدی پیشنهادی

در مرحله یازدهم، برای ادغام خروجی‌های استخراج شده از سه شبکه عصبی LSTM، از لایه تجمیع استفاده می‌شود. همان طور که در شکل ۲ نشان داده شد، D_1 ، D_2 و D_3 در کنار هم قرار می‌گیرند تا یک ویژگی پیوسته تشکیل شود. این ویژگی پیوسته وارد لایه‌های کاملاً متصل می‌شود تا پیش‌بینی روند انجام گیرد. به صورت ریاضی، پیش‌بینی شبکه عصبی هیبریدی به صورت رابطه ۱۱ بیان می‌شود.

$$\gamma_i = f(D_1, D_2, D_3) = \varphi(W(W_1D_1 + W_2D_2 + W_3D_3) + b) \quad \text{رابطه ۱۱}$$

که φ نشان‌دهنده تابع فعال‌سازی sigmoid، W_1 ، W_2 و W_3 وزن‌های اولین لایه تمام متصل و W و b به ترتیب وزن و خطای دومین لایه تمام متصل هستند.

در گام دوازدهم، با توجه به اینکه در این پژوهش از مقادیر روزانه شاخص مربوط به ۲۵ سال اخیر شاخص استفاده شده است، برای آموزش از نیمه دوم داده‌ها برای اعتبارسنجی و از نیمه اول داده‌ها برای آموزش استفاده شده است. برای افزایش دقت، این بررسی در ۵ حالت متفاوت انجام پذیرفته است.

- حالت اول: نیمه اول آموزش و نیمه دوم اعتبارسنجی؛
- حالت دوم: ۶۰ درصد اول آموزش و ۴۰ درصد آخر اعتبارسنجی؛
- حالت سوم: ۷۰ درصد اول آموزش و ۳۰ درصد آخر اعتبارسنجی؛
- حالت چهارم: ۸۰ درصد اول آموزش و ۲۰ درصد آخر اعتبارسنجی؛
- حالت پنجم: ۹۰ درصد اول آموزش و ۱۰ درصد آخر اعتبارسنجی.

در مرحله سیزدهم، شبکه عصبی هیبریدی پیشنهادی شامل یک شبکه عصبی CNN برای استخراج ویژگی‌ها در مقیاس‌های زمانی مختلف، سه شبکه عصبی LSTM برای یادگیری وابستگی‌های زمانی و لایه‌های کاملاً متصل برای استخراج ویژگی‌های سطح بالاتر است. شبکه CNN دارای دو لایه است که هر کدام شامل فرایندهای CNN تک‌بعدی، فعال‌سازی و pooling هستند. تعداد فیلترهای لایه‌های CNN به ترتیب ۱۰ و ۲۰ است. با توجه به اینکه داده‌ها به صورت سری تک‌بعدی هستند، استفاده از CNN تک‌بعدی مناسب است و تابع فعال‌سازی مورد استفاده LeakyReLU است. عملیات pooling از نوع maxpooling با سایز دو و استراید^۱ دو انجام می‌شود. تعداد واحدهای LSTM، ۱۰ و تعداد واحدهای لایه‌های کاملاً متصل نیز ۱۰ است. تابع ضرر آنتروپی متقابل باینری است و برای آموزش شبکه از بهینه‌ساز آدام استفاده می‌شود. نرخ یادگیری اولیه ۰/۰۰۱ و ضریب تضعیف^۲ ۱۰^{-۶} است. لایه Batch Normalization نیز برای بهبود یادگیری استفاده می‌شود. سایز batch برابر با ۵۰ است. برای جلوگیری از بیش‌برازش^۳، از توقف زود هنگام^۳ استفاده می‌شود؛ یعنی اگر دقت سری اعتبار در ۲۰۰ گام بهبود نیافت، فرایند یادگیری متوقف می‌شود. همچنین، به منظور جلوگیری از بیش‌برازش و بهبود عملکرد شبکه، واحدها به طور تصادفی غیرفعال می‌شود.

1. Strides
2. Overfitting
3. Early Stopping

می‌شوند و پس از عملیات کانولوشن، برای کاهش تفاوت بین توزیع داده ورودی و داده‌های داخلی، از نرمال‌سازی دسته‌ای استفاده می‌شود.

با وجود تفاوت‌های مدل‌های مختلف شبکه‌های عصبی در پردازش داده‌ها، روش یادگیری و هدف یادگیری، مقایسه آن‌ها کار ساده‌ای نیست. برای این منظور در مرحله چهاردهم، برخی از مدل‌های رایج که در پیش‌بینی‌های مالی استفاده می‌شوند، انتخاب و تعدیلات لازم روی آن‌ها انجام می‌شود تا با مدل شبکه عصبی هیبریدی قابل مقایسه باشند. مدل‌های رایج مذکور شامل شبکه‌های عصبی هیبریدی ساده، CNN و LSTM هستند.

پیش‌بینی روند شاخص کل بورس اوراق بهادار، به‌طور کلی به‌عنوان یک مسئله طبقه‌بندی در نظر گرفته می‌شود. برای ارزیابی کیفیت پیش‌بینی، از معیار دقت استفاده می‌شود. دقت نشان‌دهنده نسبت تعداد نمونه‌های درست پیش‌بینی شده به تعداد کل نمونه‌هاست. هرچه دقت بیشتر باشد، عملکرد مدل پیش‌بینی بهتر خواهد بود. دقت با استفاده از رابطه ۱۲ محاسبه می‌شود.

$$\text{دقت} = \frac{N_{correct}}{N_{all}} * 100\% \quad \text{رابطه ۱۲}$$

$N_{correct}$ تعداد نمونه‌هایی است که پیش‌بینی مدل با روند واقعی یکسان بوده و N_{all} نشان‌دهنده تعداد کل نمونه‌هاست.

برای پاسخ به سؤال اول پژوهش، داده‌ها پس از استخراج، نرمال‌سازی می‌شوند و سپس برای آموزش به مدل ارائه می‌شوند. در ابتدا، دقت برای هر ۵ فرم آموزشی محاسبه و بهترین فرم یا بخش‌بندی برای آموزش و اعتبارسنجی انتخاب می‌شود. پس از انتخاب بهترین بخش‌بندی برای آموزش و اعتبارسنجی، داده‌های اعتبارسنجی در بهترین فرم، به ۱۰ زیربخش بر اساس ترتیب سری مقادیر شاخص کل تقسیم می‌شوند. معیار دقت بر اساس اعتبارسنجی، روی هر یک از این زیربخش‌ها، در بهترین فرم محاسبه می‌شود. در نهایت دقت مدل با استفاده از معیار دقت بررسی و نتیجه عملکرد شبکه عصبی هیبریدی پیشنهادی ارائه می‌شود.

به‌منظور پاسخ به سؤال دوم پژوهش، معیار دقت شبکه عصبی هیبریدی پیشنهادی با معیار دقت شبکه‌های عصبی هیبریدی ساده، CNN و LSTM مقایسه می‌شود. در این پژوهش، زمانی که یک شاخص (شبکه عصبی هیبریدی پیشنهادی) و مدل‌های رقیب وجود دارند و پژوهشگر قصد دارد تا شاخص را با مدل‌های رقیب به‌صورت کیفی و جفت‌جفت^۱ مورد مقایسه قرار دهد، از آزمون ویلکاکسون رتبه علامت‌دار^۲ استفاده می‌شود که از روش‌های آماری ناپارامتریک است. آزمون فرض آماری ناپارامتریک ویلکاکسون رتبه علامت‌دار، روی نمونه شاخص کل انجام می‌شود تا بررسی شود که آیا شبکه پیشنهادی نسبت به شبکه‌های رقیب تفاوت آماری دارد یا خیر. آزمون فرض آماری، یک آزمون یک طرفه

1. Pairwise

2. Wilcoxon Signed Rank

است که در زیر توضیح داده شده است (لانگ، لو و کوی^۱، ۲۰۱۹). در رابطه^{۱۳}، μ_1 نشان‌دهنده دقت شبکه عصبی هیبریدی پیشنهادی است و μ_2 دقت مدل‌های رقیب را نشان می‌دهد.

$$H_0 = \mu_1 \leq \mu_2 \quad \text{رابطه ۱۳}$$

$$H_1 = \mu_1 > \mu_2$$

در آزمون فرض طبق روش ویلکاکسون رتبه علامت‌دار، ابتدا دقت‌های استخراج شده برای هر دو نوع شبکه عصبی (شبکه‌های رقیب و شبکه عصبی هیبریدی پیشنهادی) را در جدولی ثبت می‌کنیم. سپس تفاوت‌های میانگین بین دقت شبکه عصبی هیبریدی پیشنهادی و دقت مدل‌های رقیب را محاسبه می‌کنیم. این تفاوت‌ها را به ترتیب مرتب می‌کنیم و به آن‌ها رتبه می‌دهیم. سپس از تفاوت‌هایی که نشان‌دهنده اختلاف معنادار بین دقت شبکه عصبی هیبریدی پیشنهادی و دقت مدل‌های رقیب است، استفاده می‌کنیم تا نتیجه‌ای آماری بگیریم.

برای نمونه‌های کم، جداول خاصی برای آزمون فرض صفر تهیه می‌شوند؛ اما برای نمونه‌های بیشتر از حجم ۱۵، توزیع T^+ تقریباً نرمال است و برای انجام آزمون، امید ریاضی و واریانس محاسبه می‌شود. سپس با استفاده از تفاوت‌های محاسبه شده، مقادیر p-value و واریانس محاسبه می‌شود.

امید ریاضی و واریانس T^+ طبق روابط ۱۴ و ۱۵ به دست می‌آید:

$$E(T^+) = n(n+1)/4 \quad \text{رابطه ۱۴}$$

$$V(T^+) = n(n+1)(2n+1)/24 \quad \text{رابطه ۱۵}$$

در نهایت، آماره طبق رابطه ۱۶ به دست می‌آید.

$$Z = \frac{\sum T^+ - E(T^+)}{\sqrt{V(T^+)}} \quad \text{رابطه ۱۴}$$

که $\sum T^+$ بیانگر مجموع رتبه مشاهده‌ها «+» است.

یافته‌های پژوهش

برای انجام آمار توصیفی داده‌ها، داده‌های ورودی از مقادیر روزانه شاخص کل بورس اوراق بهادار تهران از سال ۱۳۷۷ تا ۱۴۰۱ (به مدت ۲۵ سال) گردآوری شده‌اند. این داده‌ها به صورت بازدهی ساده محاسبه شده‌اند. تعداد داده‌ها برابر با تعداد بازدهی‌ها است و شاخص‌های آماری معمولاً بر روی میانگین داده‌ها اعمال می‌شود.

در این بخش، می‌توانیم از مقادیر آماری مختلفی مانند میانگین، انحراف معیار و غیره برای توصیف داده‌ها استفاده کنیم. این اطلاعات می‌توانند به ما کمک کنند تا درک بهتری از ویژگی‌های داده‌ها و توزیع آن‌ها داشته باشیم. برای

مثال، میانگین بازدهی برای تمام دوره ۲۵ ساله، انحراف معیار بازدهی و میانه داده‌ها از جمله شاخص‌های آماری مهمی هستند که می‌توانند ما را در درک بهتر داده‌ها یاری کنند.

جدول ۱. دقت شبکه‌های عصبی در هر فرم در مقیاس زمانی هفتگی

آماره	تعداد داده	میانگین	انحراف معیار	مینیمم	میانگین ابتدایی داده‌ها		
					%۷۵	%۵۰	%۲۵
مقدار	۶۰۴۳	%۰/۱۳	%۰/۷۹	%-۵/۵۳	%۰/۳۴	%۰/۰۷	%-۰/۱۶

با توجه به اطلاعات ارائه شده، میانگین داده‌ها ۰/۱۳ درصد و انحراف معیار آن ۰/۷۹ درصد است که نشان می‌دهد داده‌ها به نسبت زیادی از میانگین پراکنده‌اند. همچنین، وجود چولگی به چپ و کشیدگی بیشتر از توزیع نرمال نشان می‌دهد که داده‌ها به راست متمایل هستند. این اطلاعات مهمی هستند که برای تحلیل و تفسیر داده‌ها باید در نظر گرفته شوند.

با توجه به اینکه سری داده‌های ورودی برای پیش‌بینی به دو بخش، یعنی داده‌های آموزش و داده‌های اعتبارسنجی تقسیم می‌شود، باید نسبت مناسبی برای این تقسیم را تعیین کنیم. برای این منظور، ۵ حالت مختلف از نسبت تقسیم داده‌ها به دو بخش مختلف (آموزش و اعتبارسنجی) در نظر گرفته می‌شود. سپس، معیار دقت برای هر یک از این ۵ فرم برای مدل‌ها و در مقیاس زمانی هفتگی و ماهانه (۵ روزه و ۲۰ روزه) مورد بررسی قرار می‌گیرد تا بهترین فرم برای بخش‌بندی داده‌ها تعیین شود. دقت نیز به کمک رابطه ۱۲ محاسبه می‌شود.

با توجه به نتایج جدول ۲، دقت پیش‌بینی هفتگی براساس بخش‌بندی‌های مختلف آموزش و اعتبارسنجی برای همه مدل‌های شبکه عصبی بررسی شده است. این نتایج نشان می‌دهد که در پیش‌بینی هفتگی، فرم اعتبارسنجی ۳۰ درصد، دارای بیشترین دقت یادگیری به صورت میانگین است. بنابراین، بهترین فرم (بخش‌بندی جهت آموزش و اعتبارسنجی) برای پیش‌بینی هفتگی، فرم ۳۰ درصد (۳۰ درصد اعتبارسنجی و ۷۰ درصد آموزش) است.

جدول ۲. دقت شبکه‌های عصبی در هر فرم در مقیاس زمانی هفتگی

فرم‌های بخش‌بندی					شبکه عصبی
%۱۰	%۲۰	%۳۰	%۴۰	%۵۰	
۰/۳۵۷۴۹	۰/۳۷۸۴۷	۰/۵۶۳۲۳	۰/۴۱۷۶۵	۰/۴۸۲۶۳	مدل پیشنهادی
۰/۶۶۲۵۴	۰/۶۲۴۷۵	۰/۵۶۱۵۴	۰/۵۶۹۳۵	۰/۴۹۶۸۹	CNN-CNN-LSTM
۰/۴۱۸۳۴	۰/۴۸۹۵۱	۰/۴۳۶۱۲	۰/۴۸۱۴۳	۰/۴۵۷۵۶	CNN-LSTM
۰/۳۳۹۷۱	۰/۳۶۹۴۱	۰/۵۸۱۴۵	۰/۴۸۸۴۲	۰/۵۰۹۳۱	CNN
۰/۳۹۹۸۱	۰/۳۶۸۳۴	۰/۴۵۱۲۳	۰/۴۸۹۲۱	۰/۴۰۸۹۳	LSTM
۰/۴۳۵۵۷	۰/۴۴۶۰۹	۰/۵۱۸۷۱	۰/۴۸۹۲۱	۰/۴۷۱۰۶	میانگین در هر فرم

با توجه به نتایج مشاهده شده در جدول ۳، دقت پیش‌بینی ماهانه براساس بخش‌بندی‌های مختلف آموزش و اعتبارسنجی برای همه مدل‌های شبکه‌ی عصبی بررسی شده است. این نتایج نشان می‌دهد که در پیش‌بینی ماهانه، فرم اعتبارسنجی ۱۰ درصد دارای بیشترین دقت یادگیری به‌صورت میانگین است. بنابراین، بهترین فرم (بخش‌بندی جهت آموزش و اعتبارسنجی) برای پیش‌بینی ماهانه، فرم ۱۰ درصد (۱۰ درصد اعتبارسنجی و ۹۰ درصد آموزش) است.

جدول ۳. دقت شبکه‌های عصبی پیشنهادی در مقیاس زمانی ماهانه

فرم‌های بخش‌بندی					شبکه‌ی عصبی
٪۱۰	٪۲۰	٪۳۰	٪۴۰	٪۵۰	
۰/۳۰۰۱۲	۰/۳۵۸۷۵	۰/۳۹۹۷۸	۰/۴۰۱۲۹	۰/۵۱۷۴۱	مدل پیشنهادی
۰/۶۹۹۸۵	۰/۶۸۴۱۲	۰/۶۰۱۲۱	۰/۵۹۹۸۴	۰/۴۰۱۲۳	CNN-CNN-LSTM
۰/۴۲۸۵۳	۰/۳۱۵۷۶	۰/۴۲۹۷۶	۰/۴۸۹۵۳	۰/۴۶۹۸۴	CNN-LSTM
۰/۷۴۳۵۲	۰/۳۲۲۱۳	۰/۴۳۱۰۲	۰/۳۸۹۷۶	۰/۵۳۰۱۳	CNN
۰/۳۰۱۸۵	۰/۳۱۸۷۱	۰/۳۹۸۹۰	۰/۳۹۹۸۷	۰/۳۸۰۴۳	LSTM
۰/۴۹۴۷۷	۰/۳۹۹۸۹	۰/۴۵۲۱۳	۰/۴۵۶۰۵	۰/۴۵۹۸۰	میانگین در هر فرم

برای پاسخ به سؤال اول پژوهش به‌صورت زیر عمل می‌کنیم:

در مقیاس هفتگی، با انتخاب شدن فرم ۳۰ درصد (آموزش ۷۰ درصد ابتدایی داده‌ها و اعتبارسنجی ۳۰ درصد انتهایی داده‌ها)، بخش مورد اعتبارسنجی (۳۰ درصد انتهایی داده‌ها) به ۳۰ زیربخش تقسیم می‌شود (لانگ و همکاران، ۲۰۱۹). برای ارزیابی دقت هریک از این زیربخش‌ها به این صورت عمل می‌شود که ۷۰ درصد ابتدایی داده‌ها به‌عنوان آموزش و ۳۰ زیربخش هر یک به‌عنوان بخش اعتبارسنجی مورد ارزیابی قرار گرفته و براساس معیار دقت بیشتر بیان شده، ۳۰ دقت برای هریک از مدل‌های شبکه‌ی عصبی در مقیاس زمانی هفتگی به‌دست می‌آید.

در مقیاس ماهانه، با انتخاب شدن فرم ۱۰ درصد (آموزش ۹۰ درصد ابتدایی داده‌ها و اعتبارسنجی ۱۰ درصد انتهایی داده‌ها)، بخش مورد اعتبارسنجی (۱۰ درصد انتهایی داده‌ها) به ۳۰ زیربخش تقسیم می‌شود (لانگ و همکاران، ۲۰۱۹). برای ارزیابی دقت هریک از این زیربخش‌ها به این صورت عمل می‌شود که ۹۰ درصد ابتدایی داده‌ها به‌عنوان آموزش و ۳۰ زیربخش هر یک به‌عنوان بخش اعتبارسنجی مورد ارزیابی قرار گرفته و براساس معیار دقت بیشتر بیان شده، ۳۰ دقت برای هریک از مدل‌های شبکه‌ی عصبی در مقیاس زمانی ماهانه به‌دست می‌آید.

ادعای این پژوهش این است که عملکرد شبکه پیشنهادی بهتر از شبکه‌های رقیب است ($\mu_1 > \mu_2$).

مقدار بحرانی^۱ در سطح اطمینان ۹۵ درصد در نظر گرفته شده است. اگر مقدار آماره آزمون از مقدار بحرانی بیشتر

باشد، H_0 رد می‌شود و ادعای ما تأیید می‌شود.

1. Critical Value

جدول ۴. آزمون فرض آماری هفتگی

آماره‌ها				
LSTM	CNN-LSTM	CNN-CNN-LSTM	CNN	
۱/۲۹۷	۱/۲۹۷	-۱	-۱/۵۹۸۷	مدل پیشنهادی
۱/۷۵۶	۱/۷۵۶	۱/۲۹۷		CNN
۱/۲۹۷	۱/۲۹۷			CNN-CNN-LSTM
-۱				CNN-LSTM
تفاضل مقدار بحرانی (۱/۶۵) از آماره‌ها				
LSTM	CNN-LSTM	CNN-CNN-LSTM	CNN	
-۰/۳۵۳	-۰/۳۵۳	-۲/۶۵	-۳/۲۴۸۷	مدل پیشنهادی
۰/۱۰۶	۰/۱۰۶	-۰/۳۵۳		CNN
-۰/۳۵۳	-۰/۳۵۳			CNN-CNN-LSTM
-۲/۶۵				CNN-LSTM

جدول ۴ شامل دو بخش است: جدول آماره‌ها و جدول تفاضل مقدار بحرانی از آماره‌ها. در جدول آماره‌ها، آماره‌های به‌دست آمده از آزمون فرض ویلکاکسون رتبه علامت‌دار مشاهده می‌شود. مدل‌های موجود در ردیف‌ها به‌عنوان μ_1 و مدل‌های موجود در ستون‌ها به‌عنوان μ_2 معرفی شده‌اند. هدف این جدول، مقایسه برتری مدل‌های موجود در ردیف با مدل‌های موجود در ستون است.

زمانی که آماره آزمون در هر سلول بیشتر از مقدار بحرانی باشد (که در اینجا ۱.۶۵ در نظر گرفته شده است)، فرض H_0 رد می‌شود. بنابراین، براساس جدول تفاضل مقدار بحرانی از آماره‌ها، در صورتی که عدد موجود در هر سلول بزرگتر از صفر باشد، فرض H_0 رد می‌شود. به عبارت دیگر، مدل‌هایی که در ردیف قرار دارند، بهتر از مدل‌های موجود در ستون هستند. در صورتی که عدد موجود در هر سلول کوچک‌تر یا مساوی با مقدار بحرانی باشد، مدل‌های ردیف هیچ برتری نسبت به مدل‌های ستون ندارند.

این تحلیل آماری به ما کمک می‌کند تا به‌صورت دقیق‌تر بتوانیم برتری مدل‌های مختلف شبکه عصبی را نسبت به یکدیگر ارزیابی کنیم و اطمینان حاصل کنیم که کدام مدل‌ها دارای عملکرد بهتری در پیش‌بینی شاخص کل بورس اوراق بهادار تهران هستند.

جدول ۵، مشابه جدول ۴، شامل دو بخش است: جدول آماره‌ها و جدول تفاضل مقدار بحرانی از آماره‌ها، اما در مقیاس ماهانه. همانند توضیحات مربوط به جدول ۴، هدف این جدول نیز مقایسه برتری مدل‌های موجود در ردیف‌ها نسبت به مدل‌های موجود در ستون‌هاست.

جدول ۵. آزمون فرض آماری ماهانه

آماره‌ها				
LSTM	CNN-LSTM	CNN-CNN-LSTM	CNN	
-۱	-۱/۷۹۴	-۳/۱۰۲	-۳/۴۶۸	مدل پیشنهادی
۳/۴۸۱	۲/۷۸۹	۰/۵۹۶		CNN
۲/۷۸۹	۲/۱۹۷			CNN-CNN-LSTM
۱/۹۹۰				CNN-LSTM
تفاضل مقدار بحرانی (۱/۶۵) از آماره‌ها				
LSTM	CNN-LSTM	CNN-CNN-LSTM	CNN	
-۲/۶۵	-۳/۴۴۴	-۴/۷۵۲	-۵/۱۱۸	مدل پیشنهادی
۱/۸۳۱	۱/۱۳۹	-۱/۰۵۴		CNN
۱/۱۳۹	۰/۵۴۷			CNN-CNN-LSTM
۰/۳۴۰				CNN-LSTM

در این بخش برای پاسخ به سؤال ۱ پژوهش، دقت شبکه عصبی هیبریدی پیشنهادی در حالت‌های مختلف آموزش و اعتبارسنجی محاسبه شد. سپس، دقت این مدل در ۳۰ زیربخش اعتبارسنجی در فرم ۳۰ درصد برای مقیاس هفتگی و فرم ۱۰ درصد برای مقیاس ماهانه به دست آمد.

برای پاسخ به سؤال ۲ پژوهش، دقت شبکه عصبی هیبریدی پیشنهادی در فرم‌های مختلف آموزش و اعتبارسنجی محاسبه شد و در ۳۰ زیربخش اعتبارسنجی در فرم ۳۰ درصد برای مقیاس هفتگی و فرم ۱۰ درصد برای مقیاس ماهانه با دقت مدل‌های رقیب (شبکه‌های عصبی ترکیبی ساده، CNN و LSTM) مقایسه شد. به منظور انجام این مقایسه آماری، از روش ناپارامتریک ویلکاکسون رتبه علامت‌دار استفاده شد.

نتیجه‌گیری و پیشنهادها

در این پژوهش، یک شبکه عصبی هیبریدی مبتنی بر استخراج ویژگی مقیاس زمانی چندگانه برای پیش‌بینی روند شاخص کل بورس اوراق بهادار تهران پیشنهاد شد. شبکه عصبی هیبریدی شامل دو رویکرد نوین، یعنی شبکه عصبی CNN و شبکه عصبی LSTM بود که به صورت هیبریدی^۱ از آن‌ها استفاده شد. همچنین به این مسئله پرداخته شد که آیا شبکه عصبی هیبریدی پیشنهادی، توانایی پیش‌بینی شاخص کل بورس را دارد یا خیر. این پژوهش روی داده‌های جمع‌آوری شده شاخص کل بورس اوراق بهادار تهران، در بازه زمانی ۱۳۷۷ تا ۱۴۰۱ انجام شد.

نتایج پژوهش نشان داد که مدل شبکه عصبی هیبریدی پیشنهادی، قادر نیست شاخص کل بورس اوراق بهادار تهران را پیش‌بینی کند. این نتیجه در تضاد با مطالعات لانگ و همکاران (۲۰۱۹) و سزر و اوزبایوگلو^۲ (۲۰۱۸) است. آن‌ها

1. Hybrid

2. Sezer & Ozbayoglu

مدل‌های موفق‌تری را برای پیش‌بینی حرکت قیمت سهام CSI300 و شاخص داوجونز ارائه داده‌اند.

همچنین نتیجه پژوهش حاضر برخلاف نتایج پژوهش هائو و گائو (۲۰۲۰) است. آن‌ها یک شبکه عصبی هیبریدی مبتنی بر استخراج ویژگی مقیاس زمانی چندگانه، برای پیش‌بینی روند شاخص S&P500 ارائه داده بودند. در این مطالعه مشخص شد که دقت شبکه عصبی هیبریدی پیشنهادی، در هر دو مقیاس هفتگی و ماهانه پایین‌تر از مدل‌های رقیب است.

مدل شبکه عصبی CNN که به‌عنوان یکی از مدل‌های رقیب بررسی شد، از نظر آماری و بر اساس معیار دقت، نسبت به شبکه عصبی هیبریدی پیشنهادی عملکرد بهتری نشان داد. این یافته‌ها نشان می‌دهد که شبکه عصبی CNN در پیش‌بینی شاخص کل بورس اوراق بهادار تهران، به‌ویژه در مقیاس‌های زمانی هفتگی و ماهانه، نتایج بهتری ارائه کرده است.

با توجه به دسترسی عمومی به اطلاعات شاخص‌های بورس اوراق بهادار تهران، محدودیتی برای داده‌های این پژوهش وجود ندارد. در پژوهش‌های بعدی، می‌توان با تغییر داده‌های روزانه به داده‌های بین‌روزی (برای مثال داده‌های دقیقه‌ای)، مدل شبکه عصبی هیبریدی مبتنی بر استخراج ویژگی زمانی چندگانه را بار دیگر بررسی کرد و همچنین، آن را مبنایی برای معاملات الگوریتمی قرار داد.

در این پژوهش از مقدار پایانی شاخص استفاده شده است. در پژوهش‌های آتی می‌توان از مدل‌های پیچیده با داده‌های ورودی مقدار آخرین، آغازین، ارزش معاملات، حجم، حداقل، حداکثر، انواع میانگین متحرک، شاخص جریان نقدینگی^۱، شاخص قدرت نسبی^۲ و ترکیبی از شاخص‌های نام‌برده استفاده کرد. علاوه بر این، در پژوهش‌های آتی می‌توان از مدل‌های ترکیبی ساده و مدل‌های پیچیده شبکه عصبی مبتنی بر استخراج ویژگی زمانی چند متغیره با ورودی خام یا پردازش شده مذکور استفاده کرد.

با توجه به دقت بیشتر سایر مدل‌های عصبی هیبریدی در ادبیات موضوع، در پژوهش‌های بعدی استفاده از سایر روش‌های هیبریدی در پیشینه پژوهش، برای پیش‌بینی شاخص کل و همچنین، استفاده از مدل پیشنهادی برای پیش‌بینی سایر شاخص‌ها و بازارهای مالی (شاخص هم‌وزن، شاخص فرابورس، قیمت انرژی و کالا) پیشنهاد می‌شود. استفاده از مدل رگرسیونی مشابه با داده‌های ورودی مدل پیشنهادی و در نتیجه آن، به‌کارگیری معیارهای ارزیابی پارامتریک مانند انواع خطا، Precision، Recall، F1 و مساحت زیرمنحنی، برای انجام پژوهش‌های آتی پیشنهاد می‌شود.

منابع

باجلان، سعید؛ فلاح‌پور، سعید و رئیسی، سارا (۱۴۰۳). بهینه‌سازی پرتفوی اعتباری بانک‌ها با استفاده از رویکرد اکچوئری و شبکه عصبی مصنوعی. *تحقیقات مالی*، ۲۶(۳)، ۷۱۰-۷۳۳.

- بهشتی مسئله گو، سیده مژگان؛ افشار کاظمی، محمدعلی؛ حقیقت منفرد، جلال و رضاییان، علی (۱۴۰۲). یادگیری عمیق برای پیش‌بینی بازار سهام با استفاده از اطلاعات عددی و متنی (رویکرد الگوریتم حافظه کوتاه‌مدت ماندگار LSTM). مهندسی مالی و مدیریت اوراق بهادار، ۱۴(۵۵)، ۶۵-۸۷.
- حاجی‌زاده طحان، مرضیه؛ قاسم‌زاده، محمد و رضاییان، مهدی (۱۳۹۹). یک شبکه حافظه طولانی کوتاه‌مدت عمیق مبتنی بر توجه تکاملی برای پیش‌بینی سری‌های زمانی. هوش محاسباتی در مهندسی برق، ۱۱(۴)، ۱۵-۲۸.
- حیدری، مهدی و امیری، حمیدرضا (۱۴۰۱). بررسی قدرت مدل‌های مبتنی بر هوش مصنوعی در پیش‌بینی روند قیمت سهام بورس اوراق بهادار تهران. تحقیقات مالی، ۲۴(۴)، ۶۰۲-۶۲۳.
- زیادی، حسین؛ صلواتی، عرفان و لطفی هروی، محمد مهدی (۱۴۰۲). پیش‌بینی قیمت مسکن با استفاده از الگوریتم هوش مصنوعی LSTM. تحقیقات مالی، ۲۵(۴)، ۵۵۷-۵۷۶.
- شریف‌فر، امیر؛ خلیلی عراقی، مریم؛ رئیسی وانانی، ایمان و فلاح شمس، میرفیض (۱۴۰۱). کاربرد معماری‌های یادگیری عمیق در پیش‌بینی قیمت سهام (رویکرد شبکه عصبی پیچشی CNN). مدیریت دارایی و تأمین مالی، ۱۰(۳)، ۱-۲۰.
- طهرانی یزدی، نگار؛ واعظی، رضا؛ ستایشی، سعید و رئیسی وانانی، ایمان (۱۴۰۳). پیش‌بینی تعهدات آتی شرکت‌های بیمه با استفاده از مدل حافظه بلند مدت - کوتاه‌مدت. تحقیقات مالی، ۲۶(۴)، ۸۵۴-۸۷۹.
- عبدی، نسیمه؛ مرادزاده فرد، مهدی؛ احمدزاده، حمید و خدام، محمود (۱۴۰۰). ارائه مدل ترکیبی بهینه‌سازی سید سهام براساس پیش‌بینی قیمت با شبکه عصبی بازگشتی LSTM به کمک محدودیت‌های کاردینالیتی و روش‌های تصمیم‌گیری چندمعیاره (مطالعه موردی بورس اوراق بهادار تهران). چشم‌انداز مدیریت مالی، ۱۱(۳۶)، ۱۱۹-۱۴۳.

References

- Abdi, N., Moradzadeh Fard, N., Ahmadzadeh, K. & Mahmoud, M. (2021). Presenting an optimal portfolio optimization model based on price prediction with LSTM recurrent neural network using cardinality constraints and multi-criteria decision-making methods (case study: Tehran Stock Exchange). *Financial Management Perspective*, 11(36), 119-143. doi: 10.52547/JFMP.11.36.119 (in Persian)
- Addagalla, S., Koppuravuri, S., Krosuri, R., Kunapareddy, M. S., Mallu, S. R. & Rashmi, M. (2023). Stock Market Price Prediction Using Machine Learning Techniques. *Paper presented at the 2023 4th International Conference for Emerging Technology (INCET)*, 2023.
- Al Ridhawi, M. & Al Osman, H. (2023). Stock Market Prediction from Sentiment and Financial Stock Data Using Machine Learning. *In Canadian AI*.
- Ali, M., Khan, D. M., Alshanbari, H. M. & El-Bagoury, A. A. H. (2023). Prediction of Complex Stock Market Data Using an Improved Hybrid Emd-Lstm Model. *Applied Sciences*, 13(3), 1429.

- Bajalan, S., Fallahpour, S. & Raeesi, S. (2024). Bank's Credit Portfolio Optimization Using Actuarial Approach and Artificial Neural Networks. *Financial Research Journal*, 26(3), 710-733. doi: 10.22059/frj.2021.311064.1007074 (in Persian)
- Beheshti-e-Masale-goo, S.M., Afshar-Kazemi, M.A., Haghghat Monfared, J. & Rezaian, A. (2013). Deep learning for stock market forecasting using numerical and textual information (LSTM long-term memory algorithm approach). *Financial Engineering and Securities Management*, 14(55), 65-87. (in Persian)
- Cabessa, J. & Siegelmann, H. T. (2011). Evolving Recurrent Neural Networks Are Super-Turing. *Paper presented at the 2011 International Joint Conference on Neural Networks*.
- Chiriță, M. (2012). Usefulness of Artificial Neural Networks for Predicting Financial and Economic Crisis. *Econ. Appl. Inform*, 2, 61-66.
- Crowley, J. L. (2021). *Machine Learning with Neural Networks*. In *Eccai Advanced Course on Artificial Intelligence* (pp. 39-54). Springer.
- Deswal, V. & Kumar, D. (2023). Stock Market Price Prediction Using Machine Learning Techniques: A Review. *Paper presented at the 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*.
- Devkar, C. & Sharma, S. (2023). A Study on Aspect of Artificial Neural Networks for Machine Learning. *International Journal of Advanced Research in Science, Communication and Technology*, 3(1).
- Dhakar, D. S. & Shiwani, S. (2023). Stock Market Prediction Employing Discrete Wavelet Transform and Moving Average Gradient Descent. *Paper presented at the Proceedings of the International Conference on Cognitive and Intelligent Computing: ICCIC 2021, Volume 2*.
- Enriko, I. K., Gustiyana, F. & Krishna, H. (2023). Forecasting Jpfa Share Price Using Long Short Term Memory Neural Network. *JAICT*, 8, 157.
- Fjellström, C. (2022). Long Short-Term Memory Neural Network for Financial Time Series. *Paper presented at the 2022 IEEE International Conference on Big Data (Big Data)*.
- Götze, T., Gürtler, M. & Witowski, E. (2023). Forecasting Accuracy of Machine Learning and Linear Regression: Evidence from the Secondary Cat Bond Market. *Journal of Business Economics*, 93(9), 1629-1660.
- Hahn, P. (2019). [Artificial Intelligence and Machine Learning]. *Handchir Mikrochir Plast Chir*, 51(1), 62-67.
- Hao, Y. & Gao, Q. (2020). Predicting the Trend of Stock Market Index Using the Hybrid Neural Network Based on Multiple Time Scale Feature Learning. *Applied Sciences*, 10(11), 3961.
- Hajizadeh Tahan, M., Ghasemzadeh, M. & Rezaeian, M. (2020). An evolutionary attention-based deep long short-term memory for time series prediction. *Computational Intelligence in Electrical Engineering*, 11(4), 15-28. doi: 10.22108/isee.2020.121597.1334 (in Persian)

- Harandi, A., Moeineddin, A., Kaliske, M., Reese, S. & Rezaei, S. (2024). Mixed Formulation of Physics-Informed Neural Networks for Thermo-Mechanically Coupled Systems and Heterogeneous Domains. *International Journal for Numerical Methods in Engineering*, 125(4), e7388.
- Heidari, M. & Amiri, H. (2022). Inspecting the Predictive Power of Artificial Intelligence Models in Predicting the Stock Price Trend in Tehran Stock Exchange. *Financial Research Journal*, 24(4), 602-623. doi: 10.22059/frj.2022.320064.1007149 (in Persian)
- Hochreiter, S. & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.
- Huang, F., Zeng, J., Zhang, Y. & Xu, W. (2020). Convolutional Recurrent Neural Networks with Multi-Sized Convolution Filters for Sound-Event Recognition. *Modern Physics Letters B*, 34(23), 2050235.
- Jiang, X., Zhang, L., Geng, H. & Liu, X. (2023). A Temporal Convolution Memory Network Model for Short-Term Wind Speed Forecasting. *Paper presented at the 2023 8th Asia Conference on Power and Electrical Engineering (ACPEE)*.
- Kanagachidambaresan, G. R., Ruwali, A., Banerjee, D. & Prakash, K. B. (2021). Recurrent Neural Network. In *Programming with TensorFlow: Solution for Edge Computing Applications* (pp. 53-61).
- Kohsasih, K. L., Hayadi, B. H., Juliandy, C. & Pribadi, O. (2022). Sentiment Analysis for Financial News Using RNN-LSTM Network. *Paper presented at the 2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*.
- Long, W., Lu, Z. & Cui, L. (2019). Deep Learning-Based Feature Engineering for Stock Price Movement Prediction. *Knowledge-Based Systems*, 164, 163-173.
- Marhon, S. A., Cameron, C. J. F. & Kremer, S. C. (2013). Recurrent Neural Networks. In *Handbook on Neural Information Processing* (pp. 29-65).
- Mojica, F., Villaseñor, C., Alanis, A. Y. & Arana-Daniel, N. (2019). Long Short-Term Memory with Smooth Adaptation. *Paper presented at the 2019 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*.
- Ogus, D. (2019, August). *Long Short-Term Memory: Can Artificial Neural Networks beat Econometric Models?* *Business Economics*. Retrieved from: <http://hdl.handle.net/2105/47961>
- Owusu-Adjei, M., Hayfron-Acquah, J. B., Frimpong, T. & Abdul-Salaam, G. (2023). A Systematic Review of Prediction Accuracy as an Evaluation Measure for Determining Machine Learning Model Performance in Healthcare Systems. *medRxiv*, 2023.06.01.23290837.
- Patil, S., Shankar, V. G., Devi, B., Singh, A. & Upadhyay, N. (2023). Long Short-Term Memory-Driven Recurrent Neural Network for Real-Time Stock Monitoring and Prediction. In *Machine Intelligence Techniques for Data Analysis and Signal Processing: Proceedings of the 4th International Conference MISIP 2022, Volume 1* (pp. 813-823). Singapore: Springer Nature Singapore.

- Rout, A., Bar, A. K., Saha, S. P. & Chaudhuri, A. (2022). Stock Market Prediction Using Machine Learning Algorithm. *IJARCCCE*, 11, 97-116.
- Saboor, A., Hussain, A., Agbley, B. L. Y., Li, J. P. & Kumar, R. (2023). Stock Market Index Prediction Using Machine Learning and Deep Learning Techniques. *Intelligent Automation & Soft Computing*, 37(2).
- Sako, K., Mpinda, B. N. & Rodrigues, P. C. (2022). Neural Networks for Financial Time Series Forecasting. *Entropy*, 24(5), 657.
- Sato, D., Matsubayashi, T. & Toda, H. (2022). Prediction Device, Prediction Method, and Prediction Program. *Google Patents*.
- Sezer, O. B. & Ozbayoglu, A. M. (2018). Algorithmic Financial Trading with Deep Convolutional Neural Networks: Time Series to Image Conversion Approach. *Applied Soft Computing*, 70, 525-538.
- Shanmugapriya, D. (2023). A Hybrid Model for Fingerprint Recognition Via LSTM and CNN. *Paper presented at the 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*.
- Sharif far, 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. doi: 10.22108/amf.2022.129205.1673 (in Persian)
- Shu, J., Yuan, X., Meng, D. & Xu, Z. (2023). CMW-Net: Learning a Class-Aware Sample Weighting Mapping for Robust Deep Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Singh, A. & Markande, L. (2023). Stock Market Forecasting Using LSTM Neural Network. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 544-554.
- Stamos, M. (2022). Forecasting Stock Market Volatility. *The Journal of Portfolio Management*, 49.
- Tehraniyazdi, N., Vaezi, R., Setayeshi, S. & Raeesi Vanani, I. (2024). Forecasting Insurance Company Commitments with Long Short-Term Memory Models. *Financial Research Journal*, 26(4), 854-879. doi: 10.22059/frj.2024.367421.1007532 (in Persian)
- Thorat, M., Pandit, S. & Balote, S. (2022). Artificial Neural Network: A Brief Study. *Asian Journal For Convergence In Technology (AJCT)*, 8(3), 12-16.
- Umer, M., Awais, M. & Muzammul, M. (2019). Stock Market Prediction Using Machine Learning (ML) Algorithms. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 8, 97-116.
- Wan, W. D., Bai, Y. L., Lu, Y. N. & Ding, L. (2022). A Hybrid Model Combining a Gated Recurrent Unit Network Based on Variational Mode Decomposition with Error Correction for Stock Price Prediction. *Cybernetics and Systems*, 1-25.

- Xie, Y. (2023). Stock Price Forecasting: Traditional Statistical Methods and Deep Learning Methods. *Highlights in Business, Economics and Management*, 21, 740-745.
- Yao, K. & Zheng, Y. (2023). Fundamentals of Machine Learning. *In Nanophotonics and Machine Learning: Concepts, Fundamentals, and Applications (pp. 77-112)*. Springer.
- Yosinski, J., Clune, J., Nguyen, A., Fuchs, T. & Lipson, H. (2015). *Understanding Neural Networks through Deep Visualization*. arXiv preprint arXiv:1506.06579.
- Young, J. D., Andrews, B., Cooper, G. F. & Lu, X. (2020). Learning Latent Causal Structures with a Redundant Input Neural Network. *Paper presented at the Proceedings of the 2020 KDD Workshop on Causal Discovery*.
- Ziyadi, H., Salavati, E. & Lotfi Heravi, M. M. (2023). Housing Price Forecasting Using AI (LSTM). *Financial Research Journal*, 25(4), 557-576. doi: 10.22059/frj.2023.349924.1007398 (in Persian)