Publisher: Shahid Beheshti University
Journal of Financial Management Perspective
2024, Vol. 14, No.45
PP: 123-147
DOI: 10.48308/jfmp.2024.104892
P-ISSN: 2645-4637
E-ISSN: 2645-4645
Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by /4.0/)

Presenting an Optimized CNN-LSTM Model for Stock Price Forecasting in the Tehran Stock Exchange

Nima Gholami *, Naser Shams Gharneh **

Research Paper

Abstract

Purpose: One of the primary objectives for investors and traders in financial markets is understanding stock price behavior. Accurate price predictions can assist individuals in making informed decisions about buying, selling, or holding stocks, thereby maximizing potential profits or minimizing losses through effective transaction timing. Due to the nonlinear behavior of prices and their dependence on various factors, predicting stock prices is challenging. To address this, the research employs deep learning models, a subset of machine learning models known for their capability to handle large datasets. Their structure, featuring multiple layers and interconnected nodes (neurons), allows for recognizing patterns and relationships, facilitating more accurate price predictions.

Method: This study utilizes two proposed models, LSTM-CNN and CNN-LSTM, to predict stock prices in the Tehran Stock Exchange. These models are optimized through hyperparameter tuning using the PSO algorithm, a population-based optimization technique, along with model enhancement techniques such as adversarial training, attention mechanism, and residual blocks. The proposed models are compared with CNN, LSTM, and CNN-

Received: 2024. March. 11 Accepted: 2024. September. 14.
* MSc. Student In Financial Engineering , Amirkabir University of Technology, Tehran, Iran. Email: nima.gh@aut.ac.ir
** Associate Professor, Department of Industrial Engineering & Management Systems, Amirkabir University of Technology, Tehran, Iran. (Corresponding Author). Email: nshams@aut.ac.ir

LSTM models. Data from 10 stocks, spanning September 11, 2013, to September 11, 2023, were analyzed. The input data include adjusted stock prices, indicators, oscillators, the free market US dollar price, and the inflation rate.

Findings: Comparative results of evaluation metrics, including RMSE, MAE, R-squared, and MAPE, indicate that the two proposed models outperform other models. The LSTM-CNN model, in particular, has demonstrated the best performance. Evaluation of the proposed models, with and without the PSO algorithm, reveals that the algorithm aids in optimizing the models. Furthermore, strategy analysis based on the proposed models, applied to five stocks from five different industries over four time periods, shows superior financial performance. In essence, the LSTM-CNN and CNN-LSTM models have achieved notable success in terms of financial returns and the Sharpe ratio compared to other strategies, with the LSTM-CNN model showing the most favorable performance.

Conclusion: Hyperparameters are critical as they impact model outcomes, and varying these parameters can produce different results, which may be either suitable or unsuitable. These parameters are not learned during training and must be set before training. The findings suggest that hyperparameter optimization can fine-tune models based on stock behavior, leading to more accurate price predictions. Additionally, the LSTM-CNN model excels in feature extraction and understanding dependencies within the data compared to the CNN-LSTM model, thereby enhancing prediction accuracy. In summary, leveraging deep learning models, specifically LSTM-CNN and CNN-LSTM, optimized with PSO and advanced techniques, significantly improves stock price prediction in the Tehran Stock Exchange. The results indicate that these models, particularly the LSTM-CNN, offer superior performance in financial return metrics and prediction accuracy. This approach provides a robust framework for investors and traders seeking to make informed decisions in financial markets, highlighting the importance of advanced machine learning techniques in enhancing predictive analytics.

Keywords: Forecasting price, Deep Learning, LSTM, CNN, Tehran Stock Exchange

ناشر: دانشگاه شهیدبهشتی
نشریه چشم‌انداز مدیریت مالی
۱۴۰۳، دوره ۱۴، شماره ۴۵
ص‌ص. ۱۲۳ – ۱۴۷

شاپای چاپی: ۴۶۳۷- ۲۶۴۵
شاپای الکترونیکی: ۴۶۴۵- ۲۶۴۵

Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

مقاله پژوهشی

ارائه مدلی برای پیش بینی قیمت سهام مبتنی بر CNN-LSTM بهینه شده در بورس اوراق بهادار تهران

نیما غلامی [*]، ناصر شمس قارنه [**]

چکیده

هدف: یکی از مهم‌ترین اهداف سرمایه‌گذاران و معامله‌گران بازارهای مالی شناخت رفتار قیمت سهام است. پیش‌بینی قیمت، می‌تواند به افراد در مورد خرید، فروش یا نگهداری سهام کمک کند تا با زمان‌بندی موثر معاملات، سود بالقوه را به حداکثر یا زیان را به حداقل برسانند. به دلیل رفتار غیرخطی قیمت و وابستگی آن به عوامل گوناگون پیش‌بینی قیمت کار دشواری است که برای مقابله با این چالش، در این پژوهش از مدل‌های یادگیری عمیق کمک گرفته می‌شود که زیر شاخه ای از مدل‌های یادگیری ماشین هستند. ساختار مدل‌های یادگیری عمیق به دلیل تمرکز بر روی تعداد لایه‌های بیشتر و استفاده از گره‌های به هم پیوسته (نورون‌ها) در مواجه با مقادیر زیادی داده عملکرد مناسبی دارند و به دلیل توانایی آن‌ها در شناخت الگوها و روابط، امکان ارائه پیش‌بینی دقیق‌تری از قیمت‌ها را فراهم می‌کنند.

روش: در این پژوهش، برای پیش‌بینی قیمت سهام در بازار بورس اوراق بهادار تهران از دو مدل پیشنهادی LSTM-CNN و CNN-LSTM به وسیله بهینه‌سازی هایپرپارامترها توسط الگوریتم PSO که یک الگوریتم بهینه‌سازی مبتنی بر جمعیت است و تکنیک‌های تقویت مدل شامل: آموزش خصمانه، مکانسیم توجه و بلوک باقی‌مانده استفاده شده است. این مدل‌ها با مدل‌های CNN،LSTM وCNN-LSTM مورد مقایسه قرار گرفته اند. جهت دستیابی به این هدف، داده‌های۱۰سهم از تاریخ۲۰ شهریور ۱۳۹۲ تا تاریخ۲۰ شهریور ۱۴۰۲ مورد بررسی

تاریخ دریافت مقاله: ۱۴۰۲/۱۲/۲۱، تاریخ پذیرش مقاله: ۱۴۰۳/۰۶/۲۴

[*] دانشجوی کارشناسی ارشد مهندسی مالی، دانشگاه صنعتی امیرکبیر، تهران، ایران. E-Mail: nima.gh@aut.ac.ir
[**] دانشیار گروه مهندسی صنایع و سیستم های مدیریت ، دانشگاه صنعتی امیرکبیر، تهران، ایران (نویسنده مسئول). E-Mail: nshams@aut.ac.ir

قرارگرفته‌است.داده‌های ورودی شامل: قیمت‌های تعدیل‌شده سهم، اندیکاتورها، اسیلاتورها، قیمت دلارآزاد آمریکا و نرخ تورم می‌باشد.

یافته‌ها: مقایسه نتایج معیارهای ارزیابی شامل چهار معیار RMSE، MAE، R-squared و MAPE نشان از عملکرد مطلوب دو مدل پیشنهادی با سایر مدل‌ها می‌دهد. علاوه بر این، مدل پیشنهادی LSTM-CNN توانسته بهترین عملکرد را به ثبت رساند. همچنین، بررسی نتایج دو مدل پیشنهادی، با و بدون الگوریتم PSO مشخص کرد که این الگوریتم به بهینه‌سازی مدل‌ها کمک می‌کند. در نهایت، بررسی استراتژی‌های مبتنی بر مدل‌های پیشنهادی بر روی پنج سهم از پنج صنعت مختلف در چهار دوره زمانی نشان از عملکرد مالی برتر آن‌ها می‌دهد. به عبارت دیگر، مدل‌های پیشنهادی LSTM-CNN و CNN-LSTM توانسته‌اند در معیار بازده مالی و نسبت شارپ در مقایسه با سایر استراتژی‌ها موفقیت مطلوبی به‌دست آورند که مدل پیشنهادی LSTM-CNN مطلوب‌ترین عملکرد را به ثبت رسانده است.

نتیجه‌گیری: هایپرپارامترها، پارامترهایی هستند بر روی خروجی مدل‌ها تاثیر دارند و با تغییر هرکدام از این هایپارامترها نتایج مختلفی به‌دست می‌آید که می‌تواند مناسب یا غیر مناسب باشند این پارامترها در طول آموزش یاد نمی‌گیرند و باید قبل از آموزش تنظیم شوند. طبق نتایج به‌دست آمده، بهینه‌سازی هایپرپارامترها، می‌تواند سهم را بر اساس رفتار آن سهم بهینه کند که این امر منجر به پیش‌بینی دقیق‌تر قیمت می‌شود. همچنین، با استفاده از مدلLSTM-CNN استخراج ویژگی‌ها از داده‌ها و شناسایی و درک وابستگی‌ها در ویژگی‌های استخراج شده به طور مناسب تری از CNN-LSTM صورت می‌گیرد که می‌تواند در بهبود دقت پیش‌بینی کمک کند.

کلیدواژه‌ها: پیش‌بینی قیمت، یادگیری عمیق، حافظه طولانی کوتاه مدت، شبکه عصبی کانولوشن، بورس اوراق بهادار تهران.

۱. مقدمه

شناخت و بررسی رفتار قیمت سهام همواره مورد توجه سرمایه‌گذاران و پژوهشگران حوزه مالی بوده است. هدف اصلی سیستم‌های مالی در دنیا ایجاد فرصت برای سرمایه‌گذاری با بهره‌وری بالاتر توسط افراد و بنگاه‌ها است(قاضی عسگری نایینی و همکاران،۱۴۰۱). فعالان این بازار درصدد دستیابی و به کارگیری روش‌هایی هستند تا بتوانند با پیش‌بینی آتی قیمت سهام، سود سرمایه خود را افزایش دهند. بنابراین، ضروری است که روش‌های مناسب، صحیح و متکی به اصول علمی در تعیین قیمت آینده سهام توسط افراد سرمایه‌گذار صورت گیرد. پیش‌بینی بازار سهام یک مشکل چالش برانگیز است زیرا بازار سهام اساساً یک سیستم غیرخطی و پویا است (اقتصادی و محمدی،۱۴۰۲). قیمت سهام، تحت تأثیر عوامل مختلف داخلی و خارجی، وضعیت بین‌المللی، چشم‌انداز صنعت، داده‌های مالی شرکت‌های پذیرفته شده در بورس و عملکرد بازار سهام قرار دارد. داده‌های تاریخی نشان‌دهنده ویژگی‌های پیچیده قیمت سهام، مانند: عدم قطعیت و نوسان، پیش‌بینی آن را دشوار می‌کند و نتایج پیش‌بینی را با عدم قطعیت زیادی مواجه می‌سازد (خایدم[1]و همکاران، ۲۰۱۶).

برای اینکه بتوان دید مناسبی نسبت به آینده قیمت سهم داشت، از روش‌های تحلیلی سنتی که شامل: روش تحلیل بنیادی و تکنیکال است، استفاده می‌شود که جز رایج‌ترین روش‌های مورد استفاده برای بسیاری از سازمان‌ها و سرمایه‌گذاران فردی می‌باشد. اما، در سال‌های اخیر، با پیشرفت علم، روش‌های جدیدی ظهور پیدا کردند، از جمله: هوش مصنوعی(AI) [2]و مبحث یادگیری ماشین (ML) [3]که یادگیری عمیق (DL) [4]شاخه ای از آن می‌باشد که با استفاده از الگوریتم‌های مختلف سعی در پیدا کردن بهترین جواب از ورودی‌ها می‌باشند. تفاوت مدل‌های یادگیری ماشین (ML) و یادگیری عمیق (DL) در معماری آن‌ها است به طوری که، مدل‌های یادگیری ماشین (ML) از الگوریتم‌های مختلفی برای یادگیری الگوها و روابط درون داده‌ها استفاده می‌نمایند که اغلب شامل، توابع ریاضی نسبتاً ساده و تعداد محدودی لایه هستند. اما، مدل‌های یادگیری عمیق (DL) برلایه‌های متعدد تمرکز دارند و از گره‌های به هم پیوسته (نورون‌ها) تشکیل شده‌اند. همچنین، مدل‌های یادگیری ماشین (ML) در مجموعه داده‌های کوچکتر و کارهای نسبتاً ساده تر عملکرد خوبی داشته اند و در مواجهه با مجموعه داده‌های بزرگ و با ابعاد بالا دچار مشکل می‌شوند. ولی، مدل‌های یادگیری عمیق (DL) در مواجه با مقادیر زیادی داده عملکرد مناسبی دارند و به دلیل

---

[1].Khaidem
[2].Artificial Intelligence
[3]. Machine Learning
[4].Deep Learning

توانایی در یادگیری خودکار الگوها و روابط پیچیده در داده‌ها، اغلب می‌توانند به دقت بالاتری دست یابند و پیش‌بینی‌های قوی‌تری ارائه دهند (تای[۱]، ۲۰۲۳).

همچنین، پیش‌بینی قیمت سهام می‌تواند بینشی در مورد ریسک‌های بالقوه مرتبط با سرمایه‌گذاری‌های خاص (پر ریسک) ارائه دهد. مدل استفاده شده در این پژوهش، ترکیبی از دو مدل یادگیری عمیق LSTM [۲](حافظه طولانی کوتاه مدت) و CNN [۳] (شبکه عصبی کانولوشن) می‌باشد که با استفاده از تکنیک‌های تقویت مدل و بهینه‌سازی هایپرپارامترها بهبود یافته‌اند. هایپرپارامترها، پارامترهایی هستند که در طول آموزش نمی‌آموزند و باید قبل از آموزش تنظیم شوند. این پارامترها جهت شناخت روابط و پیش‌بینی قیمت‌ها استفاده می‌شوند و بر روی خروجی تاثیر دارند و با تغییر هر کدام از آن‌ها نتایج مختلفی به‌دست می‌آید که می‌تواند مناسب یا غیر مناسب باشند، بهینه‌سازی هایپرپارامترها در پیش‌بینی قیمت سهام جهت افزایش دقت و در حین حال دستیابی به نتایج مناسب‌تر استفاده می‌شود.

۲. مبانی نظری و پیشینه پژوهش

در این بخش به بررسی، پیشنیه پژوهش‌های انجام شده در ارتباط با پیش‌بینی قیمت پرداخته خواهد شد. این موارد شامل: سه مدل، LSTM ، CNN و CNN-LSTM می‌باشند.

LSTM

مدل LSTM نوعی خاصی ازRNN [۴]هستند که قادر به یادگیری وابستگی‌های طولانی مدت می‌باشند و کاربرد آن بیشتر در داده‌های دنباله دار و چندتایی است. مدل LSTM در سال ۱۹۹۷ توسط اشمیدوبرو[۵] هوکرایتر[۶] ساخته شده است (اشمیدوبرو و هوکرایتر، ۱۹۹۸). مدل LSTM در تجزیه و تحلیل داده‌های سری زمانی عملکرد بهتری را نسبت به RNN نشان داده است (عبدی و همکاران، ۱۴۰۱).

از مدل LSTM در پیش‌بینی قیمت‌های پایانی EWU انگلستان که یک صندوق قابله معامله‌ای می‌باشد استفاده و با مدل ANN [۷]، مدل SVM [۸] و مدلRF[۹]مقایسه گردید. نتایج نشان دهنده

---

1. Taye
2. Long short-term memory
3. Convolutional neural network
4. Recurrent neural network
5. Schmidhuber
6. Hochreiter
7. Artificial neural network
8. Support vector machine
9. Radio frequency

آن است که مدل LSTM بهتر از سایر مدل‌های ذکر شده در مقاله عمل کرده است (نیکو و باقرزاده، ۲۰۱۹).

در مطالعه‌ای دیگر، از مدل LSTM جهت پیش‌بینی روندهای آتی قیمت سهام با استفاده از شاخص‌های قیمت سهام و تحلیل تکنیکال پرداخته شد. داده‌های سهام مورد بررسی، مربوط به بورس برزیل می‌باشد. نتایج تجربی نشان داد که LSTM پیشنهادی نسبت به سایر مدل‌های یادگیری ماشین، مانند: مدل‌های جنگل تصادفی، پرسپترون چندلایه و مدل‌های شبه تصادفی دقیق‌تر بوده است (نلسون[۱] و همکاران،۲۰۱۷).

در کاری دیگر به مقایسه مدل LSTM و SVM برای پیش‌بینی بازار سهام پرداخت شده است. در این مطالعه، از قیمت سهام شاخص داوجونز (DJI) به همراه قیمت نفت خام و طلا برای پیش‌بینی قیمت شاخص داوجونز (DJI) استفاده شده است. هدف این مطالعه بررسی افزودن پارامترهای خارجی برای بهبود اثر بخشی پیش‌بینی بازار سهام بوده است. بر اساس نتایج، مدل بدون افزودن پارامترهای خارجی عملکرد بهتری دارد (لاکشمینارایانان[۲]و مک‌کری[۳]، ۲۰۱۹).

CNN

CNN (شبکه عصبی کانولوشنال) نوعی مدل یادگیری عمیق است که برای طبقه بندی تصاویر یا تشخیص اشیاء در تصاویر استفاده می‌شود (یوشا بنجیو[۴]و یان لیکون[۵]، ۱۹۹۵).

در پژوهشی دیگر یک مدل یادگیری عمیق برای پیش‌بینی ۱۵ سهم S&P 500 مبتنی بر رویداد پیشنهاد داده شد. رویدادها از طریق متن کاوی از رویترز و بلومبرگ شناسایی شدند. اطلاعات استخراج شده برای پیش‌بینی قیمت و معاملات سهام، با مدل CNN مورد بررسی قرار گرفته‌اند (ژاو دینگ[۶] و همکاران، ۲۰۱۵).

در مطالعه‌ای به بررسی مدلCNN با استفاده از نظریه گراف پرداخت شده است که به مدل‌سازی بازار سهام آمریکا به عنوان یک مدل پیچیده می‌پردازد که از اطلاعات روابط مکانی و زمانی بین سهام‌های مختلف بهره می‌برد. در همین حال، مدل از هر دو شاخص سهام و اخبار مالی به عنوان ورودی استفاده می‌نماید (پتیل[۷] و همکارن،۲۰۲۰).

---

1. Nelson
2. Lakshminarayanan
3. McCrae
4. Yoshua Bengio
5. Yann LeCun
6. Ding
7. Patil

درکاری دیگر، به مقایسه قدرت انواع مدل‌های MLP، LSTM[1] و CNN درخصوص پیش‌بینی شاخص‌های S&P 500 و Nikkei225 در افق زمانی یک روزه پرداخته شده است. در این پژوهش، ازمتغیرهای ورودی شاخص‌های تکنیکال، داده‌های قیمتی و شاخص‌های کلان اقتصادی جهت پیش‌بینی و برای شناخت بهترین عملکرد از معیار RMSE استفاده شده است که بر اساس نتایج، مدل‌هایCNN، LSTM و MLP به ترتیب بهترین عملکرد را داشته‌اند (گائو [2] و همکاران، ۲۰۲۰).

نتایج استفاده از مدل LSTM، MLP و CNN برای پیش بینی قیمت سهام چهار شرکت بزرگ دولتی ایالات متحده نشان داد که این سه روش در مقایسه با مطالعات مشابهی که جهت تغییر قیمت را پیش بینی می‌کند. نتایج بهتری دارد (کامالوف[3]، ۲۰۲۰).

در پژوهشی دیگر از مدل CNN برای پیش‌بینی نرخ ارز (فارکس) استفاده شد. شاخص‌های فنی به عنوان یک ماتریس تصویر در نظر گرفته شدند. ماتریس تصویر، ابتدا با استفاده از روش Wavelet شناسایی شد و سپس توسط CNN-ResNet پردازش شد. در نهایت، از LightGBM برای جایگزینی لایه softmax برای خروجی یک پیش‌بینی استفاده شد (ژاو[4] و خوشی[5]، ۲۰۲۰).

CNN-LSTM

مدل CNN-LSTM، از ترکیب دو مدلCNN و LSTM برای استخراج و یادگیری وابستگی‌ها بین ویژگی‌های ورودی و متغیر هدف به‌دست آمده است.

از مدل CNN-LSTM جهت پیش‌بینی شاخص S&P500 ، KOSPI200 و FTSE100 استفاده شد که داده‌های پرداخته شده به صورت روزانه می‌بودند. مدل مورد نظر در یک شرایط شبیه سازی شده معاملاتی عملکرد بهتر نسبت B&H (خرید و نگهداری) داشته است. به طوری که،۲/۵۷ برابر سودآورتر از استراتژی B&H بوده است (لی[6] و کیم[7]، ۲۰۲۰).

همچنین، از مدل CNN-LSTM برای پیش‌بینی قیمت شاخص S&P 500 استفاده شد. داده‌ها به صورت یک دقیقه‌ای جمع‌آوری و معیارهای عملکردی MAE، RMSE و MAPE برای ارزیابی مدل‌ها استفاده شده‌اند. نتایج نشان‌دهنده آن بود که مدل ارائه شده از لحاظ روش‌های مختلف

---

1. Multilayer perceptron
2. Gao
3. Kamalov
4. Zhao
5. Khushi
6. Lee
7. Kim

اندازه گیری خطا، نسبت به شبکه‌های CNN و LSTM به صورت جداگانه خطای کمتری داشته است و توانسته پیش‌بینی بهتری ارائه کند (کیم[۱] و کیم[۲]، ۲۰۱۹).

در بررسی مدل CNN-LSTM در SSE Composite Index که یک شاخص بازار سهام متعلق به بورس شانگهای می‌باشد، با استفاده از داده‌های تاریخی، هشت ویژگی انتخاب و مدل CNN برای استخراج ویژگی‌های داده‌های مربوط به ۱۰ روز قبل اتخاذ شد؛ سپس، مدل LSTM برای پیش‌بینی قیمت سهام، با ویژگی داده‌های استخراج شده به کار گرفته شد. نتایج تجربی نشان داد، مدل CNN-LSTM درمقایسه با مدل‌های MLP ، CNN ، RNN ، LSTM و CNN-RNN عملکرد بهتری داشته است (ونجی لو[۳] و همکاران، ۲۰۲۰).

در پژوهشی دیگر، بررسی مدل CNN-LSTM بر روی داده‌های ارزهای دیجیتال جهت پیش‌بینی قیمت و روند ارزهای دیجیتال بر روی سه ارز دیجیتال با بالاترین ارزش بازار یعنی، بیت‌کوین (BTC)، اتریوم (ETH) و ریپل(XRP) پرداخته شد. برای این کار مدل از داده‌های ارز دیجیتال به عنوان ورودی استفاده و آن‌ها را به طور مستقل مدیریت می‌کند تا از اطلاعات مفید هر ارز دیجیتال به طور جداگانه بهره برداری کند. تجزیه و تحلیل دقیق نشان داد، که مدل توانایی بهره‌برداری مؤثر از داده‌های رمزنگاری و کاهش بیش‌برازش و هزینه‌های محاسباتی را دارد (لوانیس لیورز[۴] و همکاران، ۲۰۲۱).

طبق پیشینه پژوهش بررسی شده، مدل‌ها به وسیله مقادیر هایپرپارمترهای از پیش تعیین شده اجرا می‌شوند، به طوری که برای سهام‌ها یا شاخص‌های مختلف بررسی شده در یک کار از مقادیر یکسان هایپرپارمتر استفاده شده است. این کار ممکن است با وجود نوع داده‌ها عملکرد مدل را در برخی مناسب و در برخی دیگر نامناسب نشان دهد. برای حل این موضوع در این پژوهش از بهینه‌سازی هایپرپارمترها، جهت افزایش دقت مدل و همچنین، مناسب کردن مقادیر هایپرپارمترها با توجه به نوع سهم استفاده شده است.

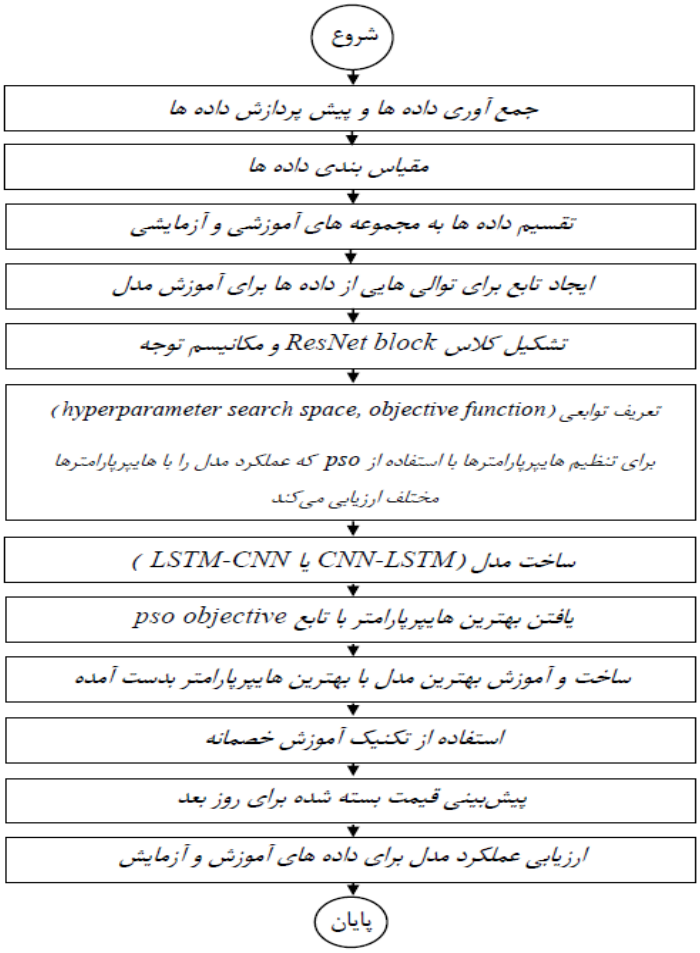
مورد دیگر، بحث استفاده از مدل ترکیبی CNN-LSTM برای پیش‌بینی قیمت‌ها است که تغییر ترکیب این مدل، به مدل LSTM-CNN در ادبیات مشاهده نمی‌شود. در این پژوهش، به بررسی و مقایسه نتایج مدل LSTM-CNN، برای پیش‌بینی قیمت بسته شده، با استفاده از تکنیک‌های تقویت مدل، جهت بهبود مدل پرداخته شده است.

---

[1].Kim, T
[۲]. Kim
[3].Wenjie Lu
[4].Livieris

۳. روش‌شناسی پژوهش

شکل (۱) نمودار فعالیت روش پیشنهادی را نشان می‌دهد.



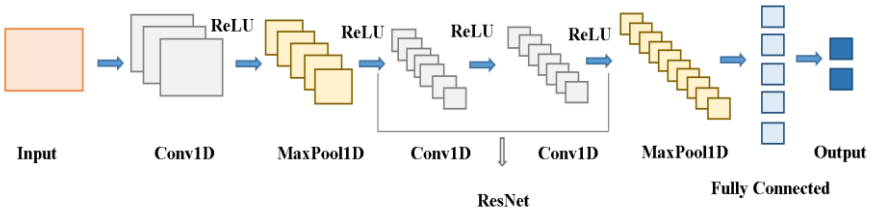
شکل۱. ساختار مدل پیشنهادی

در این بخش به معماری مدل CNN-LSTM و LSTM-CNN پیشنهادی و سایر مدل‌های مورد نظر، الگوریتم بهینه‌سازیPSO، تکنیک‌های بهبود مدل، متغیرهای ورودی مدل‌های یادگیری عمیق و داده‌های پژوهش پرداخته شده است

CNN

CNN (شبکه عصبی کانولوشنال) از چندین لایه و نورون تشکیل شده است که این نورون‌ها، وزن‌ها و سوگیری‌های قابل یادگیری دارند. لایه‌ها، ویژگی‌ها را از تصویر ورودی استخراج می‌کنند، آن‌ها را تبدیل می‌کنند و به لایه بعدی منتقل می‌کنند تا زمانی که یک پیش‌بینی انجام شود. یک مدل CNN به طور کلی، از چهار لایه اصلی تشکیل شده است که عبارتند از: لایه

Convolution، لایه Pooling ، لایه Dropout و لایه Fully Connected هر لایه وظیفه متفاوت و مخصوص به خود را دارد. در این پژوهش از مدل CNN-ResNet استفاده شده ،که با استفاده از این مدل، لایه دیگری به عنوان لایه بلوک باقیمانده نیز به مدل اضافه می‌شود. در شکل (۲) معماری CNN-ResNet مشاهده می‌شود.



شکل ۲. معماری مدل CNN-ResNet

LSTM

معماری اولیه مدل LSTM از گیت فراموشی[۱] (رابطه ۱)، ورودی[۲] (رابطه ۲) و خروجی[۳] (رابطه ۳) ساخته شده است. مدل LSTM از بلوک‌های حافظه متفاوتی به نام سلول تشکیل شده اند. حالت سلول (رابطه ۴) و (رابطه ۵) و حالت پنهان (رابطه ۶) به سلول بعدی منتقل می‌شود و تغییرات این بلوک‌های حافظه از طریق مکانسیم‌هایی که به آن‌ها گیت گفته می‌شود، انجام می‌گیرد. توابع فعال سازی در گیت‌های ورودی، خروجی و فراموشی معمولا تابع تانژانت و تابع سیگموید می‌باشند.

$$f(t) = \sigma(W_f \cdot [h(t-1).x(t)] + b_f) \quad \text{رابطه (۱)}$$

$$i(t) = \sigma(W_i \cdot [h(t-1).x(t)] + b_i) \quad \text{رابطه (۲)}$$

$$g(t) = tanh(W_c \cdot [h(t-1).x(t)] + b_c \quad \text{رابطه (۳)}$$

$$C(t) = f(t) \cdot C(t-1) + i(t) \cdot g(t) \quad \text{رابطه (۴)}$$

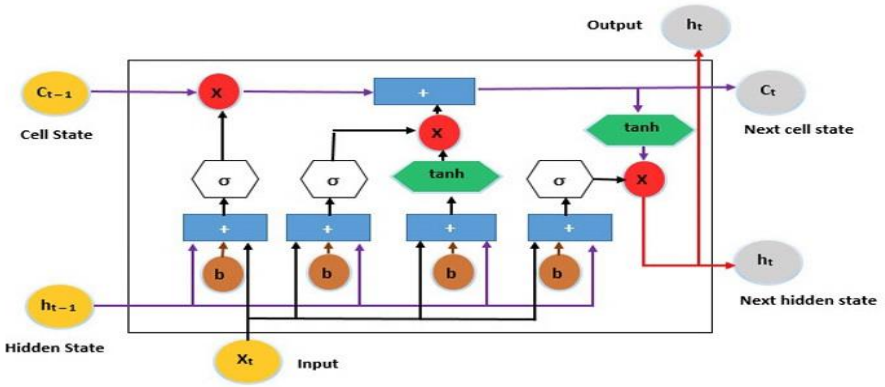
$$o(t) = \sigma(W_o \cdot [h(t-1).x(t)] + b_o) \quad \text{رابطه (۵)}$$

$$h(t) = o(t) \cdot \tanh\big(C(t)\big) \quad \text{رابطه (۶)}$$

در اینجا $f(t)$ دروازه فراموشی ، σ تابع سیگموئید ، W ماتریس‌های وزنی ، $h(t-1)$ حالت پنهان در مرحله زمانی قبلی ، x ورودی ، b بایاس ، $i(t)$ دروازه ورودی ، $g(t)$ دروازه سلول ، $tanh$ تابع تانژانت ، $C(t)$ بردار حافظه ، $o(t)$ دروازه خروجی و $h(t)$ بردار حالت پنهان است. در شکل (۳) معماری مدل LSTM مشاهده می‌شود.

---

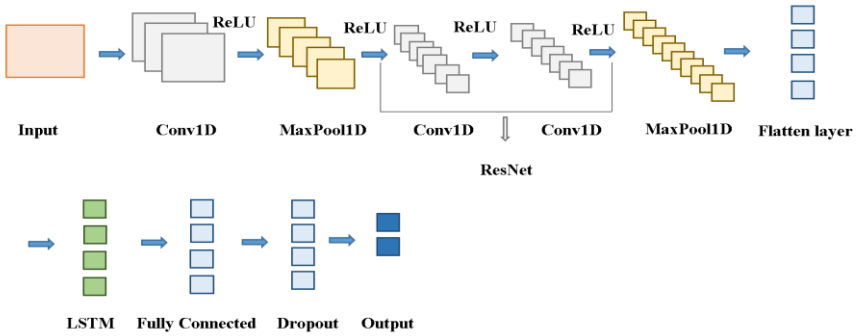
۱. Forget Gate
۲. Input Gate
۳. Output Gate



شکل ۳. معماری مدل LSTM

CNN-LSTM پیشنهادی

مدل CNN-LSTM پیشنهادی، از CNN برای استخراج ویژگی‌های داده‌ها و تبدیل داده‌های خام به مجموعه‌ای از ویژگی‌های مرتبط، و از LSTM برای یادگیری وابستگی‌ها بین ویژگی‌های ورودی و متغیر هدف استفاده می‌نماید و به وسیله آن پیش‌بینی قیمت بسته شدن سهام در روز بعد انجام می‌گیرد. با این کار مدل از مزیت هر دو مدل CNN و LSTM بهره‌مند می‌شود. مدل CNN-LSTM پیشنهادی از داده‌های سری زمانی به جای تصویر استفاده می‌نماید. در شکل (۴) معماری CNN-LSTM پیشنهادی مشاهده می‌شود.

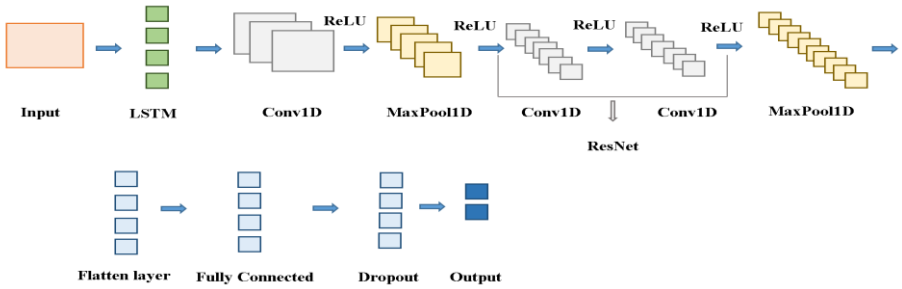


شکل ۴. معماری مدل CNN-LSTM پیشنهادی

LSTM-CNN پیشنهادی

مدلLSTM-CNN پیشنهادی، برعکس مدل CNN-LSTM عمل می‌کند. به طوری که در این روش با توجه به نوع داده‌ها که سری زمانی می‌باشند از قابلیت مدل LSTM برای استخراج ویژگی داده‌ها و الگوها از توالی داده‌ها و از CNN، برای شناسایی و درک وابستگی‌ها در ویژگی‌های

استخراج شده توسط LSTM استفاده می‌شود. در شکل (۵) معماری LSTM-CNN پیشنهادی مشاهده می‌شود.



شکل ۵. معماری مدل LSTM-CNN پیشنهادی

بهینه سازی هایپرپارامترها توسط PSO

PSO (Particle Swarm Optimization) یک الگوریتم بهینه‌سازی تصادفی مبتنی بر جمعیت است که معمولاً برای حل مسائل بهینه سازی استفاده می‌شود، انعطاف‌پذیری، سرعت همگرایی و قابلیت جستجو جهانی از مزایای این الگوریتم می‌باشد. درPSO ، جمعیتی از راه حل‌های بالقوه، به نام ذرات، در فضای جستجو برای یافتن راه حل بهینه حرکت می‌کنند. هر ذره نشان‌دهنده یک راه حل بالقوه است و یک موقعیت و یک سرعت دارد. موقعیت نشان‌دهنده مجموعه‌ای بالقوه از راه حل‌های پارامترها است و سرعت، جهت و بزرگی حرکت ذره را تعیین می‌کند. رفتار ذرات تحت تأثیر موقعیت شناخته شده خود (personal best) و بهترین موقعیت شناخته شده در بین کل جمعیت (global best) است (کیم [۱] و چو[۲]،۲۰۲۱). فرمول به‌روزرسانی سرعت در رابطه (۷) و موقعیت ذره در رابطه (۸) آمده است.

$$v_{ij}^{t+1} = wv_{ij}^{t} + c_1 r_1 \left(pbest_{ij}^{t} - x_{ij}^{t}\right) + c_2 r_2 \left(gbest_{j}^{t} - x_{ij}^{t}\right) \quad (۷)$$

$$x_{ij}^{t+1} = x_{ij}^{t} + v_{ij}^{t+1} \quad \text{رابطه (۸)}$$

در این رابطه v_{ij}^{t+1} سرعت ذره i برای هایپرپارامتر j در زمان $t+1$ و x_{ij}^{t+1} موقعیت ذره i برای هایپرپارامتر j در زمان $t+1$ است. w ، وزن اینرسی است که تاثیر سرعت قبلی ذره را تعیین می‌کند. c_1 و c_2 ، به ترتیب ضرایب شناختی و اجتماعی هستند که رفتار ذره را نسبت به موقعیت شناخته شده خود (pbest) و بهترین موقعیت شناخته شده در بین کل جمعیت (gbest) کنترل می‌کند. r_1 و r_2 ، یک عدد تصادفی بین ۰ و ۱ تولید می کند.

۱.kim
۲.cho

آموزش خصمانه

آموزش خصمانه [۱] تکنیکی است که در یادگیری ماشین برای بهبود استحکام و تعمیم مدل‌ها، به ویژه در زمینه یادگیری عمیق و شبکه‌های عصبی استفاده می‌شود. در پژوهشی با ایجاد اغتشاشات جزئی اما دقیق طراحی‌شده بر روی تصاویر به آموزش خصمانه پرداخته شد (زگدی[۲] و همکاران، ۲۰۱۴). هدف از این کار، این است که مدل در برابر حملات متخاصم و رفتارهای غیر منتظره بازار انعطاف‌پذیرتر شود و توانایی آن برای پیش‌بینی دقیق قیمت سهام در حضور اغتشاشات بهبود یابد.

بلوک باقی مانده‌ها

شبکه ResNet[۳] نوع خاصی از شبکه عصبی است (هی[۴] و همکاران، ۲۰۱۶). در شبکه ResNet برای حل مشکل زیاد شدن تعداد لایه‌ها ، بلوک جدیدی به نام Residual Block معرفی شد.

مکانیسم توجه

مکانسیم توجه یک بخش مهم در مدل‌های یادگیری عمیق، به ویژه در وظایف بینایی کامپیوتری هستند (سانگ[۵]و همکاران، ۲۰۱۸). هدف اصلی این مکانسیم کمک به مدل‌ها برای تمرکز بر مرتبط‌ترین بخش داده ورودی در حین انجام پیش‌بینی است. این مکانیسم می‌تواند وزن‌ها را به طور معقول‌تر و موثرتر مطابقت دهد تا مدل یادگیری عمیق بتواند انرژی بیشتری را روی عملیات کارآمدتر متمرکز کند.

داده‌های پژوهش

در این پژوهش، ۱۰ سهم از پنج صنعت، خودرویی ، بانک‌ها و موسسات اعتباری، فراورده‌های نفتی ، فلزات اساسی و محصولات شیمیایی بر اساس بالاترین رتبه ارزش معاملات مورد بررسی قرار گرفته است. داده‌های سهام از تاریخ۲۰ /۱۳۹۲/۰۶ تا تاریخ ۲۰ /۱۴۰۲/۰۶ و از کتابخانه فینپای (Finpy) در محیط پایتون جمع آوری شده است. در مجموع ۲۴۱۰ روز داده در پژوهش استفاده شده است که ۸۰ درصد آن برای آموزش مدل‌ها و ۲۰ درصد برای آزمایش مدل‌ها به کار گرفته می‌شود. جهت هم فرکانس کردن داده تورم از ماهانه به روزانه، تورم اعلام شده در تمام روزهای آن ماه اعمال می‌شود. همچنین، متغیرهای استفاده شده براساس ماهیت هر کدام از آن‌ها در دسته‌های مختلفی تقسیم می‌شود که در جدول (۱) قابل مشاهده است.

---

۱. Adversarial Training
۲. Szegedy
۳. Residual Network
۴. He, K., Zhang
۵. Song

جدول ۱. متغیرهای استفاده شده

متغیرها		
قیمت‌ها	اندیکاتور و اسیلاتورها	اقتصاد کلان
Adj Close	RSI 14	قیمت آزاد دلار آمریکا
Adj Open	EMA 9	نرخ تورم سالانه
Adj High	EMA 12	
Adj Low	EMA 26	
	SMA 10	

۴. تحلیل داده‌ها و یافته‌ها

معیارهای ارزیابی استفاده‌شده جهت بررسی مدل‌ها، شامل چهار معیار RMSE، MAE، R-squared و MAPE می‌باشند. نتایج معیار RMSE در جدول (۲) و جدول (۳)، جدول (۴) و جدول (۵) به ترتیب نتایج معیارMAE، R-squared و MAPE را در دوره آزمایش (برون نمونه‌ای) نشان می‌دهند. بررسی مدل‌های پیشنهادی با استفاده از الگوریتمPSO و بدون استفاده از الگوریتم PSO نشان می‌دهد، PSO در بهینه سازی هایپرپارامترها تاثیر مثبتی داشته و باعث بهبود نتایج مدل می‌شود.

جدول ۲. نتایج معیار RMSE

	RMSE						
	CNN	LSTM	CNN-LSTM	مدل پیشنهادی CNN-LSTM Without PSO	مدل پیشنهادی LSTM-CNN Without PSO	مدل پیشنهادی CNN-LSTM	مدل پیشنهادی LSTM-CNN
خودرو	۱۳۴/۷۷	۱۲۵/۶۵	۱۰۲/۱۵	۸۸/۰۵	۸۵/۹۷	۸۲/۸۷	**۸۰/۱۲**
خساپا	۱۰۹/۰۸	۸۱/۰۶	۷۵/۴۰	۷۵/۰۶	۷۴/۷۶	۷۴/۶۵	**۶۵/۲۷**
وتجارت	۱۳۱/۴۱	۸۴/۴۸	۷۷/۴۶	۷۳/۲۵	۷۰/۹۳	۶۷/۴۵	**۵۱/۰۸**
وبملت	۲۶۶/۶۶	۲۶۳/۵۴	۲۲۲/۱۸	۱۹۷/۳۶	۱۹۳/۶۶	۱۸۸/۱۳	**۱۷۱/۳۴**
شپنا	۵۰۳/۶۸	۴۹۷/۳۷	۳۷۲/۲۱	۲۹۷/۱۱	۲۹۳/۷۱	۲۸۸/۶۳	**۲۰۱/۷۶**
شتران	۱۸۵/۵۶	۱۶۵/۸۱	۱۴۴/۲۷	۱۴۰/۳۵	۱۳۸/۲۱	۱۳۳/۵۲	**۱۳۲/۴۷**
فولاد	۳۸۲/۹۱	۳۷۴/۹۰	۳۶۴/۰۵	۳۵۷/۵۴	۳۵۷/۰۴	۳۵۶/۲۹	**۳۳۷/۱۰**
فملی	۶۴۹/۰۸	۶۱۸/۸۱	۶۱۳/۹۵	۶۱۱/۰۳	۵۵۰/۷۹	۴۶۰/۴۳	**۴۳۲/۱۶**
فارس	۱۶۳۴/۴۹	۱۵۷۳/۵۷	۱۴۵۰/۲۱	۱۳۷۶/۱۹	۱۳۴۰/۹۵	۱۲۸۸/۱۱	**۸۷۸/۴۷**
شپدیس	۷۶۸۸/۷۹	۷۵۶۸/۶۳	۵۸۵۰/۵۸	۵۸۱۹/۷۵	۵۷۹۶/۹۷	۵۷۶۲/۸۰	**۴۸۳۵/۲۸**

جدول ۳. نتایج معیار *MAE*

	MAE						
	CNN	LSTM	CNN-LSTM	مدل پیشنهادی CNN-LSTM Without PSO	مدل پیشنهادی LSTM-CNN Without PSO	مدل پیشنهادی CNN-LSTM	مدل پیشنهادی LSTM-CNN
خودرو	۱۳۳/۰۰	۱۳۵/۲۶	۸۸/۰۰	۸۴/۲۳	۷۸/۵۹	۷۲/۹۵	**۷۲/۶۴**
خساپا	۱۳۰/۵۵	۹۰/۱۵	۶۶/۷۱	۶۶/۶۴	۶۶/۵۵	۶۶/۴۶	**۶۰/۵۲**
وتجارت	۱۴۹/۸۰	۸۱/۸۶	۷۱/۷۷	۷۰/۸۰	۶۹/۳۵	۶۷/۹۱	**۵۴/۴۴**
وبملت	۱۹۸/۰۶	۱۸۰/۱۲	۱۵۶/۱۷	۱۵۶/۰۹	۱۵۵/۹۷	۱۵۵/۸۵	**۹۹/۹۳**
شپنا	۴۷۷/۴۳	۳۸۶/۲۲	۲۷۰/۵۵	۲۵۳/۳۸	۲۲۷/۶۴	۲۰۱/۹۰	**۱۵۵/۹۸**
شتران	۱۹۳/۰۶	۱۸۳/۴۶	۱۲۱/۱۶	۱۲۰/۷۴	۱۲۰/۱۲	۱۱۹/۵۰	**۱۰۹/۲۶**
فولاد	۲۴۴/۸۶	۲۳۹/۹۳	۲۲۶/۷۳	۲۲۱/۵۵	۲۱۳/۷۸	۲۰۶/۰۱	**۱۸۷/۷۰**
فملی	۴۷۰/۵۷	۴۲۶/۱۶	۳۹۵/۹۳	۳۹۲/۴۹	۳۸۷/۳۴	۳۸۲/۲۰	**۳۵۲/۵۵**
فارس	۱۰۷۶/۰۱	۹۷۸/۱۵	۸۷۷/۷۰	۸۴۷/۸۶	۸۰۳/۱	۷۵۸/۳۵	**۶۹۰/۵۸**
شپدیس	۶۵۹۴/۴۳	۶۱۲۹/۰۱	۵۸۹۲/۸۷	۵۵۷۴/۲۹	۵۰۹۶/۴۳	**۴۶۱۸/۵۸**	۴۷۱۵/۷۱

جدول ۴. نتایج معیار *R-squared*

	R-squared						
	CNN	LSTM	CNN-LSTM	مدل پیشنهادی CNN-LSTM Without PSO	مدل پیشنهادی LSTM-CNN Without PSO	مدل پیشنهادی CNN-LSTM	مدل پیشنهادی LSTM-CNN
خودرو	۰/۹۱۷	۰/۹۲۰	۰/۹۳۸	۰/۹۴۰	۰/۹۴۴	۰/۹۴۹	**۰/۹۵۴**
خساپا	۰/۸۵۹	۰/۹۴۵	۰/۹۷۲	۰/۹۷۱	۰/۹۷۲	۰/۹۷۱	**۰/۹۷۶**
وتجارت	۰/۸۶۵	۰/۹۰۸	۰/۹۲۳	۰/۹۲۵	۰/۹۲۸	۰/۹۳۱	**۰/۹۳۸**
وبملت	۰/۹۴۰	۰/۹۴۳	۰/۹۴۵	۰/۹۴۶	۰/۹۴۷	۰/۹۴۸	**۰/۹۵۱**
شپنا	۰/۸۶۳	۰/۹۱۷	۰/۹۳۵	۰/۹۳۹	۰/۹۴۶	۰/۹۵۲	**۰/۹۵۷**
شتران	۰/۹۱۲	۰/۹۳۱	۰/۹۵۶	۰/۹۵۷	۰/۹۵۸	۰/۹۶۰	**۰/۹۶۲**
فولاد	۰/۸۳۴	۰/۸۴۰	۰/۸۴۳	۰/۸۴۸	۰/۸۵۵	۰/۸۶۲	**۰/۸۸۳**
فملی	۰/۷۵۹	۰/۷۸۴	۰/۷۹۲	۰/۸۰۰	۰/۸۱۳	۰/۸۲۷	**۰/۸۳۹**
فارس	۰/۷۰۲	۰/۷۱۵	۰/۷۳۱	۰/۷۴۷	۰/۷۷۰	۰/۷۹۴	**۰/۸۰۸**
شپدیس	۰/۷۲۴	۰/۷۳۳	۰/۷۳۴	۰/۷۵۰	۰/۷۷۳	۰/۷۹۶	**۰/۸۰۴**

جدول ۵. نتایج معیار MAPE

	MAPE						
	CNN	LSTM	CNN-LSTM	مدل پیشنهادی CNN-LSTM Without PSO	مدل پیشنهادی LSTM-CNN Without PSO	مدل پیشنهادی CNN-LSTM	مدل پیشنهادی LSTM-CNN
خودرو	۵/٪۵۹	۴/٪۸۹	۳/٪۲۲	۳/٪۱۲	۲/٪۸۹	۲/٪۸۲	**۲/٪۵۸**
خساپا	۶/٪۶۳	۳/٪۹۶	۳/٪۰۶	۳/٪۰۹	۳/٪۰۴	۲/٪۹۹	**۲/٪۷۲**
وتجارت	۷/٪۵۵	۴/٪۸۳	۳/٪۹۳	۳/٪۷۵	۳/٪۵۴	۳/٪۲۲	**۲/٪۸۰**
وبملت	۵/٪۹۲	۵/٪۱۵	۴/٪۴۶	۴/٪۴۳	۴/٪۳۹	۴/٪۳۵	**۳/٪۸۸**
شپنا	۷/٪۱۰	۶/٪۲۹	۴/٪۰۷	۳/٪۸۸	۳/٪۴۸	۳/٪۲۹	**۲/٪۶۰**
شتران	۴/٪۹۵	۴/٪۹۳	۳/٪۲۸	۳/٪۳۱	۳/٪۲۲	۳/٪۲۳	**۲/٪۹۶**
فولاد	۵/٪۶۳	۵/٪۱۸	۴/٪۶۳	۴/٪۵۲	۴/٪۲۶	۴/٪۱۹	**۳/٪۰۲**
فملی	۷/٪۸۰	۶/٪۵۵	۶/٪۴۶	۶/٪۱۸	۵/٪۶۶	۵/٪۳۴	**۴/٪۲۸**
فارس	۶/٪۱۷	۵/٪۹۴	۵/٪۷۱	۵/٪۷۴	۵/٪۷۰	۵/٪۶۷	**۵/٪۳۷**
شپدیس	۹/٪۴۶	۸/٪۶۲	۵/٪۸۶	۵/٪۸۳	۵/٪۷۷	۵/٪۷۲	**۵/٪۵۲**

برای بررسی معنادار بودن دقت پیش‌بینی مدل پیشنهادی با سایر مدل‌ها از آزمون دیبولد–ماریانو استفاده شده است. (دیبولد[1] وماریانو[2]، ۱۹۹۵). میانگین مربعات خطای پیش‌بینی (MSFE) به عنوان معیاری برای سنجش دقت پیش‌بینی می‌باشد. که از رابطه (۹) به‌دست می‌آید.

رابطه (۹)
$$MSFE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_{i.t})^2$$
$$\Delta MSFE_{ij} = MSFE_i - MSFE_j$$

در اینجا n تعداد مشاهدات و y_t مقدار واقعی مشاهده شده در زمان t و $\hat{y}_{i.t}$ مقدار پیش‌بینی شده توسط مدل i *(مدل اول)* در زمان t و $\Delta MSFE_{ij}$ تفاوت میانگین مربعات خطای پیش بینی بین دو مدل i *(مدل اول)* و j *(مدل دوم)*می‌باشد. آمار آزمون برای آزمون دیبولد–ماریانو با استفاده از تفاوت‌های MSFE محاسبه می‌شود و از توزیع نرمال تبعیت می‌کند که از رابطه (۱۰) به‌دست می‌آید.

رابطه (۱۰)
$$DM = \frac{\Delta MSFE_{i.j}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(\Delta e_{i.t} - \Delta e_{j.t})^2}}$$

در اینجا $\Delta e_{i.t}$ تفاوت خطای در مدل i در زمان t و $\Delta e_{j.t}$ تفاوت خطای در مدل j در زمان t می‌باشد. جدول (۶) . مقایسه نتایج آزمون دیبولد–ماریانو مدل LSTM-CNN و سایر مدل‌ها را نشان می‌دهد.

[1].Diebold
[2].Mariano

جدول ۶. مقایسه نتایج آزمون دیبولد–ماریانو مدل LSTM-CNN پیشنهادی با سایر مدل‌ها

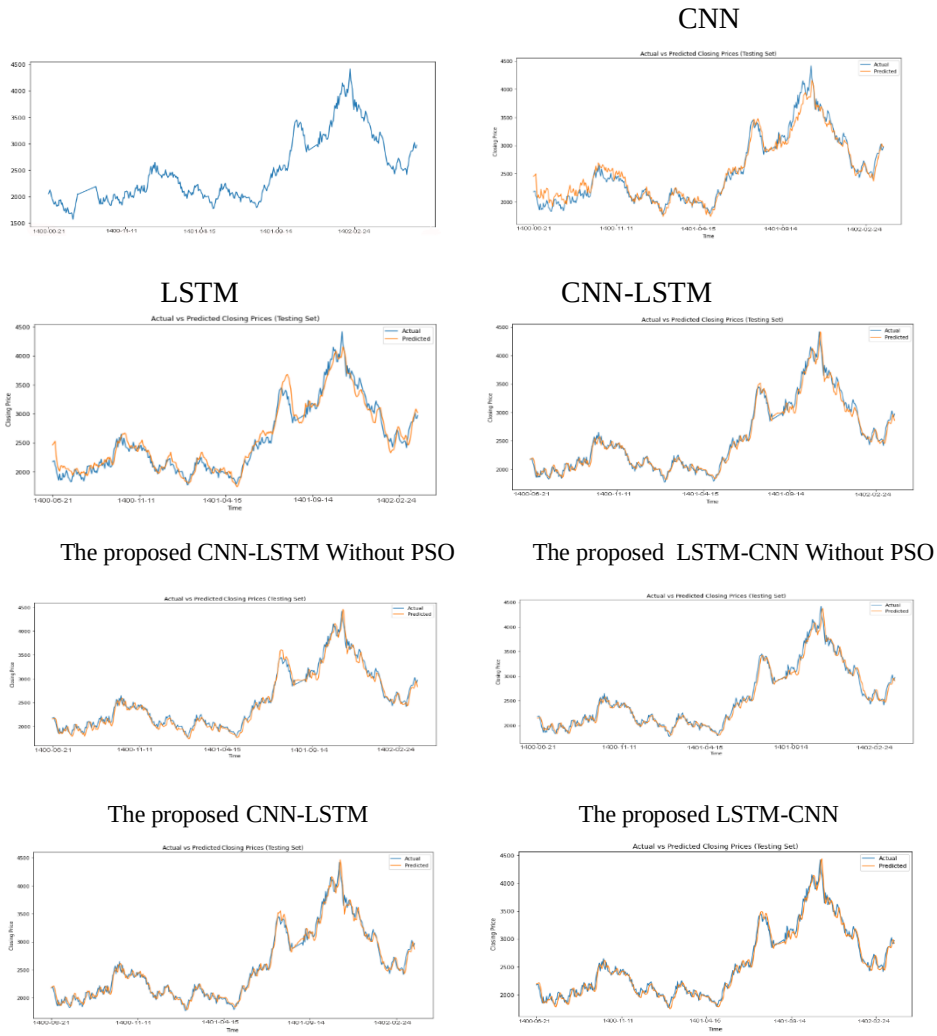
	LSTM-CNN مدل پیشنهادی		
	Observations	**DM Stat**	**p-value**
CNN	۱۰	۱۰۹۴/۱۷	۱۴ E -۱/۲۰
LSTM	۱۰	۵۶۱/۱۲	۱۳ E -۳/۴۱
CNN-LSTM	۱۰	۳۲۸/۳۱	۱۲ E -۴/۹۸
CNN-LSTM مدل پیشنهادی Without PSO	۱۰	۳۰۱/۹۰	۱۲ E -۷/۰۱
LSTM-CNN مدل پیشنهادی Without PSO	۱۰	۲۳۱/۱۷	۱۱ E -۲/۷۹
CNN-LSTM مدل پیشنهادی	۱۰	۱۸۹/۸۰	۱۱ E -۷/۷۰

برای آموزش مدل‌های CNN ، LSTM و CNN-LSTM از مقدار هایپرپارامترهای یکسانی استفاده شده است که این مقادیر در جدول (۷) مشاهده می شوند. اما، در مدل‌های پیشنهادی، مقادیر بهینه هایپرپارامترها به وسیله PSO به‌دست می‌آیند.

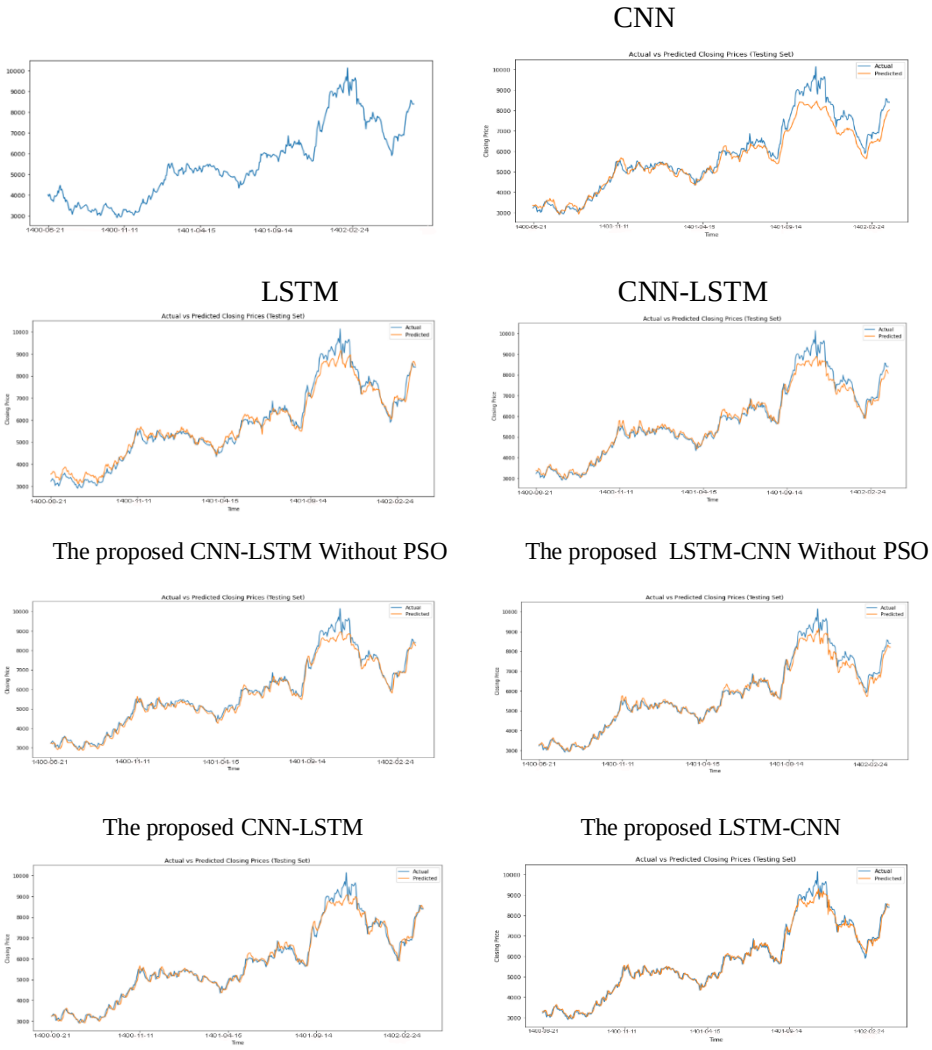
جدول ۷. مقادیر هایپرپارامترهای مدل‌ها

Hyperparameter	**CNN**	**LSTM**	**CNN-LSTM**
Learning Rate	۰/۰۰۱	۰/۰۰۱	۰/۰۰۱
Number of hidden units in LSTM layer	-	۳۲	۳۲
Number of layer	-	۳	۳
Convolution layer Activation function	Relu	-	Relu
LSTM layer Activation function	-	tanh	Tanh
Convolution layer filter	۶۴	-	۶۴
Convolution layer kernel size	۳	-	۳
Convolution layer padding	Same	-	Same
Pooling layer pool_size	۲	-	۲
Batch_size	۷۵	۷۵	۷۵
Epochs	۱۰۰	۱۰۰	۱۰۰
Dropout	۰/۱	۰/۱	۰/۱
Time_step	۶۰	۶۰	۶۰
Optimizer	Adam	Adam	Adam
Loss function	mean_squared_error	mean_squared_error	mean_squared_error

در شکل (۶) نمودار قیمت‌های پیش‌بینی شده تو سط مدل‌ها در دوره آزمایش (برون نمونه‌ای) برای سهم ایران خودرو (خودرو) و در شکل (۷) نمودار قیمت‌های پیش‌بینی شده تو سط مدل‌ها در دوره آزمایش (برون نمونه‌ای) برای سهم پالایش نفت اصفهان (شپنا)، به طور نمونه مشاهده می‌گردد. نمودار نارنجی رنگ قیمت پیش‌بینی شده و نمودار آبی رنگ قیمت واقعی می‌باشند.



شکل ۶ نمودار قیمت‌های پیش‌بینی شده توسط مدل‌ها در دوره آزمایش برای سهم خودرو



شکل ۷. نمودار قیمت‌های پیش‌بینی شده توسط مدل‌ها در دوره آزمایش برای سهم شپنا

در نهایت، استراتژی معاملاتی مبتنی بر مدل بر روی سهم خودرو نشان می‌دهد، استراتژی‌های پیشنهادی توانسته اند عملکرد به مراتب بهتری نسبت به سایر استراتژی‌ها به‌دست آورند. همچنین، استراتژی LSTM-CNN بهترین عملکرد را به ثبت رسانده است. مبنای استراتژی به این صورت است که اگر قیمت پیش‌بینی فردا از قیمت واقعی امروز بیشتر باشد وارد موقعیت خرید شده و تا زمانی که قیمت پیش‌بینی فردا کمتر از قیمت واقعی امروز نشود، از موقعیت خرید خارج نمی‌شویم. حال اگر قیمت پیش‌بینی فردا از قیمت واقعی امروز کمتر باشد و موقعیت معاملاتی باز وجود نداشته باشد وارد موقعیت خرید نخواهیم شد. این استراتژی با سه استراتژی خرید و نگهداری (Buy& Hold)، RSI وSMA مقایسه شده است. مقدار سرمایه اولیه ۱۰۰/۰۰۰/۰۰۰ تومان و مجموع

هزینه معاملات خرید و فروش ۱/۲۵درصد مبلغ نهایی معامله می‌باشد. معاملات به صورت سرمایه گذاری مجدد (سود و زیان کسب شده + سرمایه اولیه) صورت می‌گیرد. بررسی استراتژی‌ها بر روی پنج سهم و در چهار دوره زمانی انجام گرفته است. مقایسه عملکرد مالی استراتژی‌ها در جدول (۸) مشاهده می‌گردد.

جدول ۸. مقایسه عملکرد مالی استراتژی‌ها

	دوره اول (۱۴۰۱/۰۱/۰۱ تا ۱۴۰۱/۰۳/۳۱)									
	Buy & Hold		RSI		SMA		مدل پیشنهادی CNN-LSTM		مدل پیشنهادی LSTM-CNN	
	بازدهی مالی(٪)	نسبت شارپ	بازدهی مالی(٪)	نسبت شارپ	بازدهی مالی(٪)	نسبت شارپ	بازدهی مالی(٪)	نسبت شارپ	بازدهی مالی(٪)	نسبت شارپ
خودرو	−۱۶/۹۴	−۱/۴۷	−*	−	−	−	۱۱/۰۶	−۰/۱۵	**۱۱/۱۹**	**−۰/۱۹**
وتجارت	−۴/۲۸	**−۰/۳۷**	−	−	−	−	**۵/۲۲**	−۰/۴۱	۱/۸۳	−۰/۷۲
شپنا	۲۸/۶۶	۰/۶۰	−	−	−	−	۳۴/۴۰	۰/۷۱	**۴۲/۵**	**۰/۷۶**
فولاد	−۰/۵۵	**−۰/۶۴**	−	−	−	−	۶/۴۶	−۱/۴۸	**۷/۵۵**	−۰/۷۱
فارس	**۱۷/۱۳**	**۰/۱۱**	−	−	−	−	۳/۲۰	−۳/۲۹	۵/۹۴	−۲/۰۸
دوره دوم (۱۴۰۱/۰۶/۲۰ تا ۱۴۰۱/۰۹/۲۰)										
خودرو	۲۸/۲۸	۰/۵۹	−	−	−	−	۲۴/۵۷	۰/۶۴	**۳۳/۵۴**	**۰/۸۳**
وتجارت	−۱/۷۸	−۰/۷۹	۱۴/۳۰	−۰/۰۳	−	−	۱۵/۲۳	−۰/۰۱	**۲۲/۴۷**	**۰/۲۷**
شپنا	۱۸/۴۵	۰/۰۹	**۲۶/۰۶**	۰/۳۱	۱۷/۵۴	۰/۰۵	۱۹/۹۶	۰/۲۸	۲۲/۲۳	**۰/۳۸**
فولاد	−۵/۴۷	−۱/۲۸	۵/۳۶	−۰/۳۱	−	−	۱۳/۰۲	−۰/۰۸	**۱۴/۵۵**	**−۰/۰۵**
فارس	۶/۶۵	−۰/۴۰	۱۲/۵۷	−۰/۱۸	۳/۳۹	−۲/۰۴	۷/۴۶	−۰/۷۴	**۱۹/۲۹**	**۰/۱۹**
دوره سوم (۱۴۰۲/۰۱/۰۱ تا ۱۴۰۲/۰۳/۳۱)										
خودرو	−۶/۴۷	−۰/۴۰	−	−	−	−	۱۳/۸۴	−۰/۰۵	**۱۵/۵۳**	**۰/۰۲**
وتجارت	۲۶/۵۵	۰/۴۲	−	−	−	−	۲۲/۱۷	۰/۳۸	**۲۸/۶۶**	**۰/۵۴**
شپنا	−۳/۸۳	−۰/۳۶	−	−	−	−	۱۱/۱۰	−۰/۲۵	**۱۳/۶۳**	**−۰/۱۵**
فولاد	−۱۱/۰۶	−۰/۹۱	−۳/۱۱	−۳/۴۳	−	−	۱۰/۷۳	−۰/۲۸	**۱۷/۳۶**	**۰/۱۳**
فارس	**۲۱/۶۰**	**۰/۲۷**	−	−	−	−	۱۳/۵۵	−۰/۲۸	۱۶/۶۶	۰/۱۴
*سیگنالی دریافت نشده است **دوره چهارم (۱۴۰۲/۰۶/۲۰تا ۱۴۰۲/۰۹/۲۰**										
خودرو	−۷/۲۶	−۰/۶۱	۱۶/۰۹	۰/۰۵	−۱۰/۲۷	−۲/۶۹	۱۸/۰۶	۰/۱۸	**۲۴/۳۰**	**۰/۲۶**
وتجارت	−۱۱/۵۲	−۱/۳۰	۴/۸۰	−۰/۹۶	−	−	۱۰/۰۹	−۰/۳۲	**۱۶/۴۷**	**۰/۰۷**
شپنا	۲/۷۵	−۰/۵۳	−	−	−	−	۷/۱۸	−۰/۴۲	**۱۱/۳۳**	**−۰/۲۶**
فولاد	۷/۸۵	−۰/۳۲	۱۴/۴۴	−۰/۰۹	۹/۰۱	−۰/۶۵	۲۰/۹۲	۰/۲۸	**۲۵/۱۶**	**۰/۴۲**
فارس	−۱/۹۵	−۳/۰۸	۰/۱۶	−۳/۳۷	−	−	۲/۰۱	−۳/۲۵	**۴/۴۱**	**−۰/۹۷**

۵. بحث و نتیجه‌گیری

در این پژوهش، به پیش‌بینی قیمت با استفاده از مدل‌های یادگیری عمیق در بورس اوراق بهادار تهران پرداخته شده است. پیش‌بینی قیمت با این روش، این امکان را به افراد می‌دهد که بتوانند تصمیمات بهتری اتخاذ کنند و برنامه ریزی بهتری داشته باشند. برای این کار دو مدل پیشنهادی ارائه شده است که هایپرپارامترهای این مدل‌ها بر اساس الگوریتم PSO که یک الگوریتم، بهینه‌سازی تصادفی مبتنی بر جمعیت می‌باشد، بهینه شده‌اند. همچنین، در این پژوهش از تکنیک‌های تقویت مدل همچون، آموزش خصمانه، جهت انعطاف پذیری مدل در برابر رفتار

غیرمنتظره بازار، مکانسیم توجه، جهت تمرکز بر مرتبط‌ترین بخش‌های داده ورودی در حین انجام پیش‌بینی و بلوک باقی‌مانده، برای حل مشکل زیاد شدن تعداد لایه‌ها بهره‌گرفته شده است. داده‌های ورودی مدل‌ها، قیمت‌های تعدیل شده سهم، اندیکاتور و اسیلاتورها و همچنین، متغیرهای اقتصاد کلان می‌باشند. مدل‌های پیشنهادی از ترکیب دو مدل LSTM و CNN به‌دست می‌آیند؛ مدلLSTM-CNN و CNN-LSTM پیشنهادی. این مدل‌ها بر روی ۱۰ سهم از پنج صنعت مختلف بورس اوراق بهادار تهران بررسی شده و بر اساس نتایج معیارهای ارزیابی، مدل‌های پیشنهادی در مقایسه با سایر مدل‌ها توانسته‌اند عملکرد بهتری از خود بجای بگذارند و موفق به ثبت کمترین خطا نسبت به سایر مدل‌ها شوند. به طور مثال، در سهم خودر و معیار MAPE، دو مدل پیشنهادی LSTM-CNN و CNN-LSTM توانسته‌اند به ترتیب عدد ۲/۵۸ درصد و ۲/۸۲ درصد را به ثبت رسانند که این نتیجه نشان دهنده این است که، مدل LSTM-CNN پیشنهادی توانسته از مدل CNN-LSTM پیشنهادی بهتر عمل کند. همچنین، بررسی نتایج دو مدل پیشنهادی، با الگوریتم PSO و بدون الگوریتم PSO مشخص کرد که این الگوریتم به بهینه‌سازی هایپرپارامترها و بهبود دقت مدل‌ها کمک می‌کند و بر خروجی مدل می‌تواند تاثیر مثبتی بگذارد. علاوه بر این، با استفاده از مدل‌ها، استراتژی‌های ساخته، که جهت ارزیابی بر روی پنج سهم، در چهار دوره زمانی اجرا و بر اساس بازده مالی و نسبت شارپ مقایسه شده‌اند. نتایج حاکی ازآن است که مدل LSTM-CNN توانسته بهترین عملکرد را نسبت به سایر استراتژی‌ها به‌دست آورد.

طبق نتایج، به وسیله بهینه‌سازی هایپرپارامترها، می‌توان بجای استفاده از هایپرپارامترهای یکسان برای هر سهم، آن‌ها را بر اساس رفتار آن سهم بهینه کرد که در نهایت منجر به پیش‌بینی دقیق‌تری می‌شود. همچنین، با استفاده از مدل LSTM-CNN استخراج ویژگی‌ها از داده‌ها به‌طور مناسب‌تری از CNN-LSTM صورت می‌گیرد که می‌تواند در بهبود دقت پیش‌بینی کمک کند.

۶. پیشنهادها و محدودیت‌ها

کاهای آینده حول این موضوع ، می‌تواند در شناسایی و استفاده از متغیرهای تاثیرگذار هر سهام و استفاده از آن در مدل‌ها، خوشه‌بندی سهام و آموزش مدل‌ها بر روی نماینده هر خوشه، بررسی دقت مدل‌ها بر روی دیگر صنعت‌ها و صندوق‌های قابل معامله بورس اوراق بهادار تهران، بررسی مدل‌ها در مقاطع زمانی کوتاه‌تر و به هم پیوسته و مقایسه عملکرد مدل‌ها در دوره‌های زمانی متفاوت، افزودن متغیر اخبار بازار و تحلیل احساسات به مدل‌ها، استفاده از تصاویر کندل‌ها و داده‌های سری زمانی به عنوان ورودی در مدل‌ها انجام گیرد.

سپاسگزاری

از کلیه افرادی که ما را در انجام این پژوهش یاری نمودند تشکر می‌نماییم. در این پژوهش از سازمان، نهاد یا شخصی کمک مالی دریافت نشده است.

References

Abdi,Nasimeh; Moradzadeh Fard ,mehdi ; Ahmadzadeh,Hamid & Khoddam ,Mahmoud (2022). A Hybrid Model for Portfolio Optimization Based on Stock Price Forecasting with LSTM Recurrent Neural Network using Cardinality Constraints and Multi-Criteria Decision Making Methods (Case study of Tehran Stock Exchange). *Journal of Financial Management Perspective*, 11(36),119-136. (*in Persian*)

Christian, Szegedy. Wojciech, Zaremba. Ilya, Sutskever. Joan Bruna, Dumitru Erhan. Ian, Goodfellow.& Rob, Fergus.(2014). Intriguing properties of neural networks. In *ICLR*.

Diebold, F,X.& and R,S, Mariano. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, 13: 253-63.

Ding, X. Zhang, Y. Liu, T.& Duan, J.(2015). Deep learning for event-driven stock prediction. *Artificial Intelligence*, Buenos Aires, Argentina,2327-2333.

Eghtesad, Amirali. & Mohammadi, Emran. (2023). Portfolio optimization with return prediction using LSTM, Random forest, and ARIMA. *Journal of Financial Management Perspective*, 13(43), 9 - 28. (*in Persian*)

F, Kamalov.(2020). Forecasting significant stock price changes using neural networks.*Neural Computing and Applications.*

Gao, Penglei. Zhang, Rui.& Yang, Xi.(2020). The Application of Stock Index Price Prediction with Neural Network. *Mathematical and Computational Applications*, 25*(3).*

Ghazi Asgari Naeini, Saba; Neshat, Najmeh & Jafari Nodoushan, Abbasali (2022). Sustainable Policy-Making of Financial Systems in Crisis Situations With Modelling Based on Artificial Neural Networks. *Journal of Financial Management Perspective*, 12(38), 103 - 129. (*in Persian*)

Hamayel, M.J.& Owda, A,Y.(2021). A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. *AI* **,** 2, 477–496.

He, K. Zhang, X. Ren, S.& Sun, J.(2016). Deep residual learning for image recognition. *In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-778.

Hochreiter, S.(1998). The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions, *International Journal of Uncertainty Fuzziness and Knowledge-Based*, 6(2):107-116.

Khaidem,L. Saha,S. & Roy Dey,S.(2016). Predicting the direction of stock market prices using random forest. *in Applied Mathematical Finance*,1-2.

Kim, T.& Kim, H, Y .(2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PloS one*.

Lakshminarayanan, S,K.,& McCrae, J.(2019). A comparative study of svm and lstm deep learning algorithms for stock market prediction .*Artificial Intelligence and Cognitive Science,*Galway, Ireland,5–6.

Lee,S,W.& H,Y, Kim.(2020). Stock market forecasting with super-high dimensional time-series data using ConvLSTM, trend sampling, and specialized data augmentation.*Expert Systems with Applications,*161:113704.

Li, P. Song, Y. McLoughlin, I.V, Guo,W.& Dai, L.-R.(2018). An Attention Pooling based Representation Learning Method for Speech Emotion Recognition. *In Proceedings of the Interspeech* ,Hyderabad, India.

Livieris, I.E., Kyriakidou, N., Stavroyiannis, S.,Pintelas, P. (2021). An Advanced CNN-LSTM Model for Cryptocurrency Forecasting. *Electronics*, MDPI. 2021.

Nelson, D,M,Q. Pereira, A,C,M. & De Oliveira, R,A.(2017). Stock market's price movement prediction with LSTM neural Networks. *Proceedings of the International Joint Conference on Neural Networks*.

Nikou, M. Mansourfar, G.& Bagherzadeh, J.(2019). Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. *Intell*,26, 164–174.

Patil, P. Wu, C, S. Potika, K.& Orang, M.(2020). Stock market prediction using ensemble of graph theory, machine learning and deep learning models. *Engineering and Information Management*,Sydney, Australia, 12–15.

Tae-Young Kim.& Sung-Bae Cho.(2021). Optimizing CNN-LSTM neural networks with PSO for anomalous query access control. *Neurocomputing,*666-677.

Taye, M.M.(2023). Understanding of Machine Learning with Deep Learning: Architectures,Workflow,Applications and Future Directions.*Computers*, 12, 91.

Wenjie, Lu. Jiazheng, Li. Yifan, Li. Aijun, Sun.& Jingyang, Wang.(2020). A CNN-LSTM-Based Model to Forecast Stock Prices, *Complexity*, 6622927.

Yann, LeCun.& Yoshua, Bengio.(1995). Convolutional networks for images, speech, and time series.*in The Handbook of Brain Theory and Neural Networks.*

Zhao, Y.& Khushi, M.(2020). Wavelet denoised-ResNet CNN and LightGBM method to predict forex rate of change.*IEEE International*, Sorrento, Italy, 11–17 .

استناد

غلامی، نیما و شمس قارنه، ناصر (۱۴۰۳). ارائه مدلی برای پیش بینی قیمت سهام مبتنی بر CNN-LSTM بهینه شده در بورس اوراق بهادار تهران. *چشم‌انداز مدیریت مالی*، ۱۴(۴۵)، ۱۲۲-۱۴۷.

Citation

Gholami, Nima & Shams Gharneh, Naser (2024). Presenting an Optimized CNN-LSTM Model for Stock Price Forecasting in the Tehran Stock Exchange. *Journal of Financial Management Perspective*, 14(45) 123 - 147 (*in Persian*)