*Publisher: Shahid Beheshti University*
*Journal of Financial Management Perspective*
**2023, Vol. 13, No.43**
PP: 9-28
DOI: 10.48308/jfmp.2024.104191
P-ISSN: 2645-4637
E-ISSN: 2645-4645
**Copyright:** © 2023 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/)

# Portfolio optimization with return prediction using LSTM, Random forest, and ARIMA

**Amirali Eghtesad *, Emran Mohammadi ****

Research Paper

## Abstract

In today's world, optimizing investment portfolios has received increasing attention. While predicting the expected returns of investment options and incorporating them into the objective function for profit maximization is a common practice, the most significant innovation in current research is the minimization of prediction error as the objective function. This innovation advises investors to emphasize not only on profit and risk but also on the predictability of investment options when forming an investment portfolio. Integrating return prediction from traditional time series models into portfolio formation can enhance the performance of the primary portfolio optimization model. Since machine learning and deep learning models have demonstrated a significant superiority over time series models, this paper combines return prediction in portfolio formation with machine learning models, namely Random Forest, and deep learning model, Long Short-Term Memory (LSTM). To evaluate the performance of the proposed model, five years of historical data from 2017 to 2021 are used for five industry sectors: banking, automotive, pharmaceutical, metal, and petroleum. The experimental results demonstrate that the mean-variance optimization models perform better when return prediction is done using Random Forest.

**Keywords: Portfolio optimization, Machine learning, Deep learning, Prediction error, Financial risk measure.**

Received: 2023. February. 08, Accepted: 2023. December. 15.
* M.A. Student in Industrial Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. E-Mail: amirali.eghtesad@srbiau.ac.ir
** Assistant Prof., Department of Industrial Engineering, University of Science and Technology, Tehran, Iran (corresponding author). E-Mail: e_mohammadi@iust.ac.ir

**ناشر: دانشگاه شهیدبهشتی**
**نشریه چشم‌انداز مدیریت مالی**
۱۴۰۲، دوره ۱۳، شماره ۴۳
ص‌ص. ۹- ۲۸

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

Copyright: © 2023 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/)

# بهینه‌سازی سبد سرمایه‌گذاری به کمک پیش‌بینی بازده مورد انتظار با استفاده از روش‌های شبکه عصبی LSTM، جنگل تصادفی و ARIMA

**امیرعلی اقتصاد *، عمران محمدی ****

مقاله پژوهشی

## چکیده

در جهان امروز اهمیت مدل‌های بهینه‌سازی سبد سرمایه‌گذاری به صورت فزاینده‌ای مورد توجه قرار گرفته است. هرچند پیش‌بینی بازده مورد انتظار گزینه‌های سرمایه‌گذاری و در نظر گرفتن آن‌ها در تابع هدف بیشینه‌سازی سود امری رایج است لیکن مهم‌ترین نوآوری پژوهش جاری کمینه‌سازی خطای پیش‌بینی به عنوان تابع هدف است. این نوآوری به سرمایه‌گذاران توصیه می‌کند که در تشکیل سبد سرمایه‌گذاری علاوه بر سود و ریسک، بر معیار مهم قابل پیش‌بینی بودن گزینه‌های سرمایه‌گذاری نیز تاکید گردد. ادغام پیش‌بینی بازده مدل‌های سری زمانی سنتی در تشکیل پورتفولیو می‌تواند عملکرد مدل بهینه‌سازی سبد اصلی را بهبود بخشد. از آنجایی که مدل‌های یادگیری ماشین و یادگیری عمیق برتری قابل توجهی نسبت به مدل‌های سری زمانی نشان داده‌اند، این مقاله پیش‌بینی بازده در تشکیل پورتفولیو را با مدل یادگیری ماشین، یعنی جنگل تصادفی و مدل یادگیری عمیق حافظه‌ی کوتاه‌مدت طولانی ترکیب می‌کند. به منظور ارزیابی عملکرد مدل پیشنهادی، داده‌های تاریخی ۵ ساله از سال ۱۳۹۶ تا ۱۴۰۱ از شاخص ۵ صنعت بانکی، خودرویی، دارویی، فلزی و نفتی است. نتایج تجربی نشان می‌دهد که مدل‌های بهینه‌سازی میانگین واریانس با پیش‌بینی بازدهی به‌وسیله جنگل تصادفی ، بهتر عمل می‌کنند.

**کلیدواژه‌ها: بهینه‌سازی پرتفولیو، یادگیری ماشین، یادگیری عمیق، خطای پیش‌بینی، سنجه ریسک مالی.**

---

تاریخ دریافت مقاله: ۱۴۰۱/۱۱/۱۹، تاریخ پذیرش مقاله: ۱۴۰۲/۰۹/۲۴ .

* دانشجوی کارشناسی ارشد مهندسی صنایع، واحد علوم تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.
E-Mail: amirali.eghtesad@srbiau.ac.ir

** استادیار، گروه مهندسی صنایع، دانشگاه علم وصنعت، تهران، ایران (نویسنده مسئول).
E-Mail: e_mohammadi@iust.ac.ir

## ۱. مقدمه

پیش‌بینی بازار سهام یک مشکل چالش برانگیز پیش‌بینی سری‌های زمانی است، زیرا بازار سهام اساساً یک سیستم غیرخطی، پویا، پر سر و صدا و آشفته است. در واقع، قیمت سهام تحت تأثیر عوامل بسیاری مانند رویدادهای سیاسی، سیاست‌ها و اخبار شرکت، موقعیت‌های اقتصادی، نرخ بهره و احساسات سرمایه‌گذاران قرار می‌گیرد. طبق نظریه قدم زدن تصادفی[1] و بازار مالی کارا[2] هیچ سرمایه‌گذاری با استفاده از داده‌های تاریخی و فعلی نمی‌تواند سودی بیشتر از نرخ متوسط بازار به‌دست آورد، بنابراین امکان پیش‌بینی وجود ندارد [۱۶]. با این حال، پژوهشگران و سرمایه‌گذاران با فرضیه بازار کارآمد هم از لحاظ نظری و هم از لحاظ تجربی مخالف می‌باشند و معتقدند که متغیرهای مالی مانند قیمت سهام، ارزش شاخص‌های بازار سهام و قیمت مشتقات مالی تا حدودی قابل پیش‌بینی هستند [۶]. بنابراین به ارائه و آزمایش مدل‌های مربوط به رفتار بازار سهام علاقه و توجه نشان داده و تئوری‌های متعددی در مورد بازارسهام مطرح کرده‌اند. در این میان محبوبیت شبکه‌های عصبی به‌دلیل رفتار ناپارامتریک و غیر خطی و توانایی برآورد هزاران و میلیون‌ها پارامتر، در بازارهای مالی رو به افزایش است، زیرا از توانایی ذاتی در تخمین روابط غیرخطی با دقت بالا، تحلیل‌های چند متغیره بدون هیچ فرض از پیش تعیین‌شده‌ای، و سهولت تعمیم برخوردارند که این امر باعث موفقیت نسبی آن‌ها در مدل‌سازی و پیش‌بینی بخش‌های متعدد بازارهای مالی شده است [۱۵].

تجزیه و تحلیل سری‌های زمانی یک حوزه حیاتی در زمینه‌های مختلف است، از امور مالی گرفته تا علوم محیطی، که در آن پیش‌بینی دقیق مقادیر آینده ضروری است. مدل‌های سری زمانی سنتی، مانند میانگین متحرک یکپارچه اتورگرسیو (ARIMA) و هموارسازی نمایی، از دیرباز سنگ بنای پیش‌بینی سری‌های زمانی بوده‌اند. این مدل‌ها در ثبت روندهای خطی و غیر خطی و همچنین فصلی بودن داده‌ها برتری دارند. با این حال، زمانی که با الگوهای پیچیده یا بی نظمی در سری‌های زمانی مواجه می‌شوند، ممکن است با مشکل مواجه شوند. در سال‌های اخیر، مدل‌های یادگیری ماشین به‌عنوان جایگزین‌های مناسب برای تحلیل سری‌های زمانی برجسته شده‌اند. الگوریتم‌های یادگیری ماشین و الگوریتم‌های شبکه عصبی چارچوبی انعطاف‌پذیر ارائه می‌کنند که می‌تواند روابط پیچیده درون داده‌ها را ثبت کند. توانایی آن‌ها در رسیدگی به غیر خطی بودن و گنجاندن طیف گسترده‌ای از ویژگی‌ها، آن‌ها را به ابزارهای قدرتمندی برای پیش‌بینی تبدیل می‌کند [۳].

---

[1] Random Walk Hypothesis
[2] Efficient Market Hypothesis

اخیراً، بسیاری از محققان انواع مختلفی از مدل‌های یادگیری ماشین را برای پیش‌بینی بازار سهام اعمال کرده‌اند و نتایج رضایت‌بخشی مانند رگرسیون بردار پشتیبان[1] و پیش‌بینی جنگل تصادفی ایجاد کرده‌اند . شبکه‌های عصبی مصنوعی[2] به عنوان هسته فناوری یادگیری عمیق نیز به طور گسترده برای پیش‌بینی بازار سهام استفاده شده است. در میان تمام فناوری‌های یادگیری عمیق، شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی اغلب در پیش‌بینی سری‌های زمانی مالی استفاده می‌شوند.

در چارچوب برتری مدل‌های یادگیری ماشین و یادگیری عمیق در پیش‌بینی بازار سهام، بسیاری از محققان این مدل‌ها را در فرآیند پیش‌انتخاب سهام قبل از تشکیل سبد سهام به کار می‌برند و نتایج رضایت‌بخشی را ایجاد می‌کنند. در واقع، پیش انتخاب سهام با کیفیت بالا برای موفقیت مدیریت پرتفولیو بسیار مهم است . در بازار سهام، سرمایه‌گذاران فردی معمولاً سعی می‌کنند بازده آتی سهام سرمایه‌گذاری خود را تعیین کنند و سپس وزن بهینه هر سهم را برای ساختن یک سبد سرمایه‌گذاری تعیین کنند. بنابراین، پس از فرآیند پیش‌انتخاب سهام، سرمایه‌گذاران نیز باید وزن سرمایه‌گذاری بهینه را برای هر سهام انتخابی قبل از انجام سرمایه‌گذاری تجاری محاسبه کنند. این روش عمدتاً مبتنی بر نظریه مدرن بهبود پورتفولیو است.

تئوری مدرن بهینه‌سازی سبد سرمایه‌گذاری شامل مدل‌های مختلفی برای محاسبه وزن بهینه پرتفوی هر دارایی است. تحت دارایی‌های داده شده، مدل‌های پورتفولیو برای بهینه‌سازی یک یا چند تابع هدف تحت شرایط محدودیت‌های مختلف استفاده می‌شود. با حل مسئله بهینه‌سازی پرتفوی، وزن سرمایه‌گذاری بهینه هر دارایی به دست می‌آید.

مدل میانگین واریانس[3] مارکوویتز به عنوان آغازگر تئوری پرتفوی مدرن، یک مدل بهینه‌سازی پرتفوی را با حداکثر کردن بازده مورد انتظار پرتفوی و به حداقل رساندن ریسک سرمایه‌گذاری پرتفوی ایجاد می‌کند. این مدل یک مرز کارآمد را تشکیل می‌دهد که ریسک کل را تحت یک بازده مورد انتظار از پیش تعیین‌شده کاهش می‌دهد. برای هر سطح از بازده مورد انتظار، مرز کارایی استراتژی سرمایه‌گذاری بهینه را ارائه می‌دهد. با این حال، مدل میانگین واریانس محدودیت‌های زیادی برای کاربرد عملی دارد، مانند محدودیت فرضیه‌ها و پیچیدگی محاسباتی برای دارایی‌های مقیاس بزرگتر. بنابراین، مدل‌های متعددی برای حل این مسائل ارائه شده است. به عنوان مثال، کونو و یامازاکی، مدل انحراف میانگین–مطلق را توسعه دادند، که در آن از انحراف مطلق برای جایگزینی واریانس به عنوان معیار ریسک استفاده می‌کنند.

---

[1] Support vector regression
[2] Artificial neural network
[3] Mean–variance

## ۲. مبانی نظری و پیشینه پژوهش

در رابطه با سرمایه‌گذاری در معاملات سهام با استفاده از انواع مدل‌ها کارهای زیادی وجود دارد. این مقاله تنها به ارائه برخی تحقیقات نسبی می‌پردازد.

اولین بار مارکویتز[1] [۱۴] در مقاله‌ای مدل میانگین- واریانس را برای انتخاب سبد سهام معرفی نمود. مارکویتز در ابتدا هدف حداکثر کردن بازده مورد انتظار را درنظر گرفت. در ادامه مارکویتز بازده مورد انتظار را مطلوب و واریانس حاصل از بازده (ریسک) را نامطلوب درنظر گرفت و سپس مسئله را با فرض به حداکثر رساندن بازده مورد انتظار برای سطح معینی از ریسک یا به حداقل رساندن ریسک مورد انتظار در سطح معینی از بازده مورد انتظار پیش برد.

راعی [۱۷] در مقاله‌ای تحت عنوان «تشکیل سبد سهام برای سرمایهگذار مخاطره پذیر» با در نظر گرفتن هدف حداکثرسازی بازده سهام، مبنای مقایسه را مدل مارکویتز قرار داده به ایجاد مدل‌هایی با استفاده از شبکه‌های عصبی مصنوعی پرداخته و در نهایت سبد بهینه تک دوره‌ای را ارائه می‌دهد.

سلیمانی و همکاران [۱۸] مدلی برای پیش‌بینی بازه‌ای مقدار مربوط به کمترین و بیشترین قیمت روزانه به کمک شبکه‌های عصبی، ارائه دادند. آن‌ها در این پژوهش، اطلاعات مربوط به معاملات پنج سهم در دوره زمانی ۱۹۰ روز را به عنوان ورودی‌های شبکه عصبی انتخاب نمودند. نتایج حاصل‌شده بر روی سبد نمونه‌ای پژوهش، سبد بهینه در مقایسه با عملکرد سیستم معاملاتی نوسان‌گیری با دارایی‌های انفرادی دارای نسبت شارپ بهتر (حداقل دو برابر) می‌باشد که نشان از مدیریت ریسک و ریسک به اندازه در سبد بهینه دارد.

حیدری زارع و کردلویی [۹] نیز با استفاده از شبکه عصبی مدلی را برای پیش‌بینی قیمت روز بعد سهام ارائه نمودند. مدل ایشان، اطلاعات روزانه‌ی شاخص کل بورس اوراق بهادار تهران، نرخ ارز، قیمت طلا، قیمت نفت و قیمت سهام را به عنوان ورودی‌های شبکه در نظر می‌گرفت.

آذر و همکاران [۴]، به مقایسه روش‌های کلاسیک و هوش مصنوعی در پیش‌بینی شاخص قیمت سهام و طراحی مدل ترکیبی پرداختند. آن‌ها با استفاده از چندین روش آماری و طراحی یک مدل شبکه عصبی و یک مدل شبکه عصبی فازی به پیش‌بینی شاخص قیمت پرداختند. در نهایت نیز یک مدل ترکیبی استفاده نمودند که خروجی هر یک از مدل‌های شبکه عصبی، شبکه عصبی فازی، مدل آماری ARIMA به همراه داده‌های شاخص قیمت سهام واقعی به یک شبکه

---

[1] Markowitz

عصبی فازی دیگر وارد می‌شد. نتایج نشان داد که به ترتیب عملکرد رویکرد ترکیبی، شبکه فازی عصبی، شبکه عصبی و مدل ARIMA به طور قابل توجهی از بقیه روش‌ها بهتر است.

هوانگ[1][۱۱] با استفاده از رگرسیون برداری پشتیبان (SVR) و الگوریتم‌های ژنتیک، یک مدل برای انتخاب سهام توسعه داد. این مدل از رگرسیون برداری پشتیبان برای پیش‌بینی بازده آتی هر سهم استفاده کرد که الگوریتم‌های ژنتیک برای بهینه‌سازی پارامترها و ویژگی‌های ورودی مدل به کار برد. سپس، سهام با رتبه بالا و وزن‌دهی مساوی برای ساخت پورتفولیو استفاده شدند.

لی[2]و یو[3][۱۳] سه نوع شبکه عصبی بازگشتی را برای پیش‌بینی بازده سهام مقایسه کردند، شبکه عصبی بازگشتی،[4] واحد بازگشتی گیتی[5]و شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی. نتایج تجربی ارائه شده نشان داده است شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی عملکرد بهتری دارد. همچنین، آن‌ها سبدهای سرمایه‌گذاری مبتنی بر مدل مارکویتز را بر اساس نتایج پیش‌بینی شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی ساختند. نتایج این تحقیق نشان داد که این پرتفوی بازدهی بهتری به دست آورده است.

فیشر[6]و کراوس[7][۸] از سال ۱۹۹۲ تا ۲۰۱۵، شبکه‌های عصبی حافظه‌ی کوتاه‌مدت طولانی را برای پیش‌بینی حرکات سهام‌های اصلی تشکیل‌دهنده S&P500 به کار بردند. آن‌ها دریافتند که سبدهای سرمایه‌گذاری مبتنی بر شبکه‌های عصبی حافظه‌ی کوتاه‌مدت طولانی از سبدهای سرمایه‌گذاری مبتنی بر پیش‌بینی بازدهی بدون حافظه مانند: جنگل تصادفی، رگرسیون لجستیک و شبکه عصبی عمیق عملکرد بهتری دارد.

کاستی رایج مدل‌های فوق این است که این روش‌های ساخت پرتفوی، ریسک هر سهام را تجزیه و تحلیل نمی‌کنند، که باعث عدم تعادل بازده و ریسک مورد انتظار پرتفوی می‌شود.

علیزاده و همکاران [۲] با استفاده از سیستم استنتاج فازی نرم‌افزاری تطبیقی مبتنی بر شبکه‌های عصبی، یک مدل بهینه‌سازی پورتفولیو توسعه دادند که برای پیش‌بینی بازده پورتفولیو از آن استفاده کردند و از شاخص واریانس برای ارزیابی ریسک استفاده نمودند. نتایج این پژوهش نشان دادند که این مدل بهینه‌سازی پورتفولیو عملکرد بهتری نسبت به مدل میانگین–واریانس و

---

1 Huang
2 Lee
3 Yoo
4 Recurrent neural network
5 Gated recurrent unit
6 Fischer
7 Krauss

روش شبکه عصبی داشت. این تحقیق به ما اطلاع داد که ترکیب تکنیک هوش مصنوعی با بهینه‌سازی پورتفولیو مدرن می‌تواند عملکرد بهتری نسبت به هر یک از مدل‌ها به تنهایی برای سرمایه‌گذاری در بازار ارائه دهد.

دنگ[1]و مین[2][۷] از مدل رگرسیون خطی با ده متغیر برای انتخاب سهام از بورس‌های آمریکا و جهان استفاده کردند و سپس بر اساس مدل بهینه‌سازی میانگین–واریانس (MV) به همراه برخی محدودیت‌های عملی مانند تحمل مخاطره، خطای پیگیری سیستماتیک و نرخ گردش سرمایه، پورتفولیو را ساختند. آن‌ها مشاهده کردند که بازدهی تعدیل‌شده به مخاطره مدل پیشنهادی در دسته دارایی‌های خارجی بهتر از دسته دارایی‌های داخلی بود و بازدهی پورتفولیو با خطای پیگیری سیستماتیک و تحمل مخاطره سیستماتیک افزایش یافت.

وانگ[3]و همکاران [۲۱] یک مدل پورتفولیو را با استفاده از شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی برای انتخاب سهام و مدل بهینه‌سازی میانگین–واریانس برای بهبود پورتفولیو توسعه دادند. در این مدل، ابتدا شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی تعدادی سهم را از مجموعه کل سهام انتخاب کرد، سپس سهام انتخاب‌شده برای ساختن مدل پورتفولیو میانگین–واریانس استفاده شدند. آن‌ها شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی را با مدل‌های ماشین بردار پشتیبانی (SVM)، جنگل تصادفی (RF) و ARIMA در فرآیند انتخاب سهام مقایسه کردند و سپس از مدل میانگین–واریانس برای بهینه‌سازی پورتفولیو استفاده کردند. نتایج آزمایشات نشان داد که مدل پیشنهادی توسعه‌یافته توسط آن‌ها از سایر مدل‌ها عملکرد بهتری داشت.

تا[4]و همکاران [۱۹] پورتفولیوها را با استفاده از شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی و سه تکنیک بهینه‌سازی پورتفولیو، روش وزن‌دهی مساوی، شبیه‌سازی مونت کارلو و مدل میانگین–واریانس ساختند. همچنین، آن‌ها از رگرسیون خطی و مدل ماشین بردار پشتیبانی به عنوان مقایسات در فرآیند انتخاب سهام استفاده کردند. نتایج آزمایشات نشان دادند که شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی دقت پیش‌بینی بالاتری نسبت به رگرسیون خطی و مدل ماشین بردار پشتیبانی داشته و پورتفولیوهای ساخته‌شده توسط آن از سایر مدل‌ها بهتر بودند. این مدل‌ها از روش‌های مختلفی برای انتخاب سهام استفاده می‌کنند و سپس مدل‌های بهینه‌سازی پورتفولیو با سهام انتخاب‌شده برای سرمایه‌گذاری تشکیل می‌دهند. این روش‌ها به ما نشان می‌دهند که به سمت ساخت مدل‌های پورتفولیو در عمل می‌توانیم حرکت کنیم. با این حال، مدل‌های بهینه‌سازی پورتفولیوهای کلاسیک معمولاً برای سرمایه‌گذاری کوتاه مدت در عمل

[1] Deng
[2] Min
[3] Wang
[4] Ta

مناسب نیستند. بنابراین، بررسی رویکرد کارآمدتر برای ترکیب نتایج پیش‌بینی بازده با مدل‌های بهینه‌سازی پورتفولیو مهم است.

استون[1] و کاسیمبلی[2] [۲۰] با استفاده از پیش‌بینی بازده سهام، یک رویکرد تعمیم‌یافته برای ساخت مدل بهینه‌سازی پورتفولیو ایجاد کردند. آن‌ها یک مدل میانگین- واریانس- چولگی توسعه‌یافته را با استفاده از بازده‌های پیش‌بینی شده و خطاهای بازده پیش‌بینی شده برای تشکیل یازده تابع هدف ایجاد کردند.

به دنبال این رویکرد، یو وهمکاران [۲۲] پیش‌بینی‌های مدل ARIMA را در پیشبرد شش مدل بهینه‌سازی پورتفولیو (مانند مدل‌های میانگین واریانس، انحراف میانگین–مطلق، ریسک نامطلوب، ارزش در معرض خطر تفاضلی،[3] ارزش در معرض ریسک شرطی و امگا) ترکیب کرد. آن‌ها ابتدا از مدل ARIMA برای پیش‌بینی بازده سهام آتی استفاده کردند، سپس نتایج پیش‌بینی را در گسترش این مدل‌های بهینه‌سازی پرتفوی به کار بردند. نتایج تجربی نشان داد که مدل‌های بهینه‌سازی پرتفوی پیشرفته با پیش‌بینی ARIMA از مدل‌های تک پورتفولیو بهتر عمل کردند و مدل‌های توسعه‌یافته میانگین واریانس و امگا با پیش‌بینی ARIMA بهترین عملکرد را در بین این مدل‌ها داشتند.

این مدل‌ها جهت امیدوارکننده‌ای را برای ترکیب پیش‌بینی بازده سهام با مدل‌های بهینه‌سازی پرتفوی به ما نشان می‌دهند، که به طور کامل مزایای پیش‌بینی سهام را در ساخت مدل‌های بهینه‌سازی پرتفوی اعمال می‌کند.

بنابراین، این پژوهش سعی می‌کند ابتدا با پیش‌بینی بازدهی با مدل‌هایی که در مقالات ذکر شده بهترین عملکرد را داشته‌اند به پیش‌بینی بازدهی پرداخته و سپس خروجی بهترین مدل را در یک مدل تعمیم یافته میانگین–واریانس جهت بهینه‌سازی سبد سرمایه‌گذاری قرار دهد.

## ۳. روش‌شناسی پژوهش

### شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی (LSTM)

شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی نوعی شبکه عصبی بازگشتی است که برای غلبه بر محدودیت شبکه عصبی بازگشتی و حفظ اطلاعات طولانی مدت پیشنهاد شده است. این ویژگی عمدتاً بر اساس سلول‌های حافظه در لایه پنهان است. شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی

---

[1] Ustun
[2] Kasimbeyli
[3] Liquidity-Adjusted Value-at-Risk

معمولا از لایه ورودی، لایه پنهان و لایه خروجی تشکیل شده است. این شبکه عصبی در تجزیه و تحلیل داده‌های سری زمانی عملکرد بهتری را نسبت به شبکه عصبی بازگشتی ساده نشان داده است [۱].

این مقاله از میانگین مربعات خطا برای آموزش شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی استفاده می‌کند. فراپارامترهای[۱] در نظر گرفته شده شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی شامل گره‌های پنهان، لایه‌های پنهان، نرخ یادگیری، اندازه دسته،[۲] نرخ خروج،[۳] تابع فعال‌سازی، بهینه‌ساز و جریمه هستند.

تابع فعال ساز تانژانت هیپربولیک می‌باشد. مقادیر بررسی شده سایر فراپارامترها در جدول (۱) ارائه شده است. برای تعیین فراپارامتر بهینه از سعی و خطا استفاده می‌شود.

با آزمون و خطا، فراپارامترهای بهینه شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی تعیین می‌شود. توپولوژی شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی از ۴ لایه پنهان و هر لایه شامل ۵ گره تشکیل شده است. و ۰/۰۱، ۱۰۰و ۰/۴ به ترتیب برای نرخ یادگیری ، اندازه دسته و نرخ خروج تنظیم شده اند. همچنین جهت بهینه‌سازی یادگیری شبکه عصبی و پیدا کردن وزن‌های مناسب از الگوریتم بهینه‌ساز Adam استفاده شده است. این الگوریتم با ترکیب الگوریتم‌های بهینه‌ساز Momentum وRMSProp هم از حرکت زیگزاگی جلوگیری می‌کند هم شدت جهت حرکت را کنترل می‌نماید و همچنین احتمال پیدا کردن نقطه بهینه سراسری در مقابل نقطه بهینه محلی را نیز بالا می‌برد، به طور کلی این الگوریتم اکثر مشکلات الگوریتم‌های بهینه‌سازی بر پایه گرادیان کاهشی را کاهش می‌دهد [۱۲]. در ادامه، این مقاله از این مدل شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی برای پیش‌بینی بازده مورد انتظار استفاده می‌کند.

**جنگل تصادفی**

جنگل تصادفی یک مدل ناپارامتریک و غیرخطی است که اولین بار توسط هو[۴] [۱۰] ارائه شد. این مدل از مشکل بیش از حد برازش اجتناب می‌کند زیرا همیشه همگرا می‌شود . با توجه به مزایای جنگل تصادفی، اغلب برای پیش‌بینی سهام استفاده می‌شود. پارامترهای اصلی جنگل تصادفی عبارتند از: تعداد درخت تصمیم، حداکثر عمق درخت،حداقل تعداد نمونه‌های موردنیاز برای تقسیم یک گره داخلی، حداقل تعداد نمونه‌های مورد نیاز برای قرار گرفتن در یک گره برگ

---

[۱] Hyperparameters
[۲] Batch size
[۳] Dropout
[4] Ho

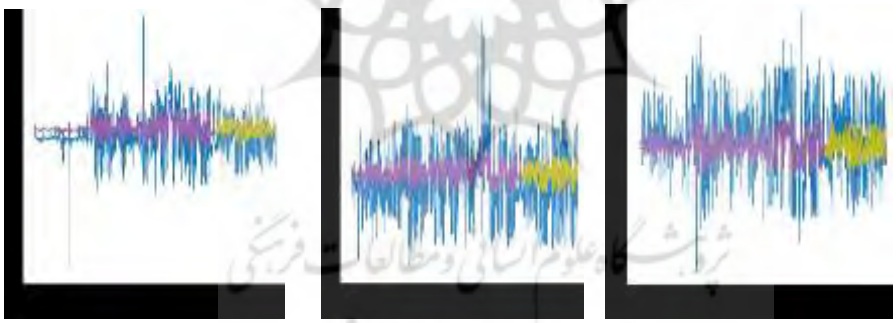
و تعداد ویژگی‌هایی که باید در هنگام جستجو برای بهترین تقسیم در نظر گرفته شوند. در این مقاله پارامترهای هر صنعت به کمک سعی و خطا جداگانه تعیین شده و برای ۵ صنعت ۵ مدل جنگل تصادفی با پارامترهای متمایز شکل می‌دهیم، که شرح کامل این مقادیر در جدول (۲) آورده شده است.

*جدول ۱ . پارامترهای شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی*

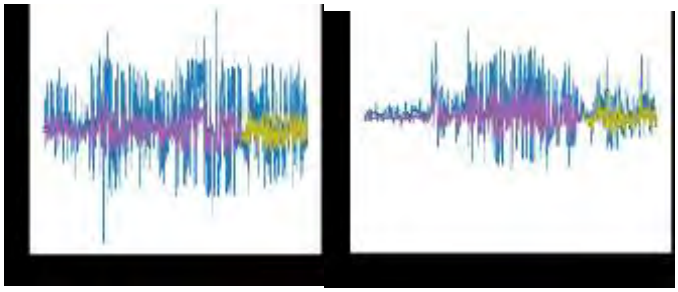
| | | |
|---|---|---|
| گره پنهان | | ۵–۱۰–۱۵–۲۰–۲۵–۳۰ |
| لایه پنهان | | ۱و۲و۳و...و۱۰ |
| نرخ یادگیری | | ۰/۱–۰/۰۱–۰/۰۰۱–۰/۰۰۰۱ |
| اندازه دسته | | ۵۰–۱۰۰–۲۰۰ |
| نرخ خروج | | ۰/۱و۰/۲و۰/۳و...و۰/۵ |
| تابع جریمه | | میانگین مربعات خطا |
| تابع بهینه ساز | | SGD,RMSprop,Adam |

*جدول۲ . پارامترهای برآورد شده جنگل تصادفی گروه‌ها*

| پارامتر | بانکی | خودرویی | دارویی | فلزات اساسی | نفتی |
|---|---|---|---|---|---|
| N_estimator | ۵۰ | ۲۰ | ۵۰ | ۲۰ | ۲۰ |
| Max-depht | ۴ | ۱۲ | ۲ | ۳ | ۲ |
| Min-samples-split | ۱۰ | ۱۰ | ۹ | ۹ | ۹ |
| Min-samples-leaf | ۳ | ۹ | ۱۱ | ۷ | ۹ |
| Random State | ۱ | ۳۰ | ۴۲ | ۴۲ | ۳۰ |



پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی



*شکل۱. نحوه برازش شبکه عصبی بر بازده شاخص‌های بانکی،خودرویی،فلزی،دارویی ونفتی(چپ به راست*

شکل (۱) نشان‌دهنده نحوه برازش شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی بر روی داده‌های آموزش و سنجش مدل می‌باشد. خطوط بنفش نشان‌دهنده داده‌هایی هستند که مدل در طول آموزش و راه‌اندازی برآورد کرده است. خطوط زرد بیانگر برآورد مدل در طول ۳۰۴ روز منتهی به تیر ۱۴۰۱ می‌باشد، که جهت سنجش عملکرد مدل مورد استفاده قرار گرفته است. خطوط آبی نیز بیانگر بازدهی واقعی هر صنعت در بازه زمانی فروردین ۱۳۹۶ الی تیر ۱۴۰۱ می‌باشد. این شکل به وضوح نشان می‌دهد که مدل دچار برازش بیش از حد نشده است.

**میانگین متحرک خود همبسته یکپارچه (ARIMA)**

مدل‌های سری زمانی مختلفی وجود دارند که می‌توانند برای تجزیه و تحلیل داده‌ها مورد استفاده قرار گیرند که یکی از آن‌ها میانگین متحرک یکپارچه اتورگرسیو (ARIMA) است. دلایل متعددی برای محبوبیت مدل ARIMA وجود دارد، از جمله ویژگی‌های آماری آن و همچنین روش شناخته‌شده باکس-جنکینز[۱] برای ساخت مدل‌ها [۵]. در یک مدل میانگین متحرک یکپارچه اتورگرسیو، مقدار آینده یک متغیر تابعی خطی از چندین مشاهدات گذشته و خطاهای تصادفی فرض می‌شود. فرآیند زیربنایی که سری‌های زمانی را تولید می‌کند در معادله (۱) آورده شده است.

رابطه (۱) $$y = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \varepsilon_0 - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}$$

در این معادله $y_t$ و $\varepsilon_t$ به ترتیب نشان‌دهنده مقادیر واقعی مشاهدات و خطای تصادفی در زمان $t$ می‌باشند. $\phi_i$ و $\theta_j$ پارامترهای مدل هستند که به ترتیب نشان‌دهنده بخش

[۱] Box-Jenkins

خودهمبستگی (AR) و میانگین متحرک (MA) می‌باشند. p و q اعداد صحیح هستند و اغلب به عنوان مرتبه مدل شناخته می‌شوند. $\varepsilon_t$ که خطاهای تصادفی مدل فرض می‌شود، به طور مستقل از یکدیگر و با میانگین صفر و واریانس ثابت $\sigma^2$ توزیع شده اند. اگر $q = 0$ باشد، رابطه (۱) به یک مدل مدل AR از مرتبه p تبدیل می‌شود و وقتی $p = 0$، مدل به یک مدل MA از مرتبه q تبدیل می‌شود. یکی از وظایف اصلی مدل ARIMA، تعیین ترتیب مدل مناسب (p, q) است.

**نظریه مدرن سبد سهام**

اگر اوراق بهادار ریسک‌دار باشند، مسئله اصلی هر سرمایه‌گذار، تعیین سبدی از اوراق بهادار است که مطلوبیت آن حداکثر شود. این مسله معادل انتخاب سبد سهام بهینه از مجموعه سبد سهام‌های ممکن می‌باشد که تحت عنوان مسله انتخاب سبد سهام از آن یاد می‌شود. پیدایش تئوری مدرن سبد سهام به سال ۱۹۵۲ برمی‌گردد یعنی زمانی که هری مارکوتیز مقاله ای تحت عنوان «انتخاب سبد سهام» را منتشر کرد. رویکرد مارکوتیز برای انتخاب پرتقوی با این فرض شروع می‌شود که شخصی، مقدار معینی پول برای سرمایه‌گذاری در اختیار دارد. این مبلغ برای یک مدت زمان مشخص که «دوره نگهداری سرمایه» نامیده می‌شود سرمایه‌گذاری خواهد شد. سپس مبلغ مورد نظر، مصرف و یا سرمایه‌گذاری مجدد خواهد شد.

مارکویتز بیان می‌کند که سرمایه‌گذاران به صورت همزمان به دو پدیده ریسک و بازده توجه می‌کنند. ریسک با نوسانات بازده مرتبط است و نوسانات توسط واریانس بازده اندازه‌گیری می‌شود. مطابق تئوری مدرن سبد سهام، سرمایه‌گذاری که در پی حداکثر نمودن بازده مورد انتظار و حداقل کردن عدم اطمینان یا ریسک است، دو هدف متضاد در پیش رو دارد که بایستی آن‌ها را با متنوع‌سازی و تشکیل سبد سهام در برابر یکدیگر موزانه نماید. از دیدگاه مارکوتیز، تنوع بخشی شامل ترکیب اوراق با حداقل همبستگی مثبت به منظور کاهش ریسک در سبد سهام، بدون از دست دادن بازده سبد سهام است.

**مدل میانگین واریانس مارکویتز**

مارکوویتز به عنوان پیشرو نظریه مالی مدرن، مدل میانگین واریانس را معرفی کرد که یک راه حل ریاضی برای حل و فصل مبادله بین حداکثرکردن بازده مورد انتظار و به حداقل رساندن ریسک ارائه کرد. این مقاله نتایج پیش‌بینی سود مورد انتظار را در پیشبرد مدل میانگین واریانس برای ساخت مدل میانگین واریانس با پیشبینی[۱] ترکیب می‌کند. مدل میانگین واریانس با

[۱] Mean–variance with forecasting

پیش‌بینی در واقع یک مسئله بهینه‌سازی چند هدفه است. معادلات زیر مدل میانگین واریانس با پیش‌بینی را ارائه می‌دهد. که در آن $x_i$ به معنای نسبت دارایی $i$ در پرتفوی است، $n$ تعداد دارایی‌های موجود در پرتفوی است. $\sigma_{ij}$ کوواریانس دارایی $i$ و $j$ است، $\hat{r}$ نشان‌دهنده بازده پیش‌بینی‌شده دارایی است و $\overline{\varepsilon_i}$ نشان‌دهنده میانگین خطاهای پیش‌بینی دارایی $i$ در دوره نمونه است که آن را به صورت $\varepsilon_i = r_i - \hat{r}_i$ نشان می‌دهیم.

$$Max \sum_{i=1}^{n} x_i r$$ رابطه (۲)

$$Min \sum_{i=1}^{n} x_i \overline{\varepsilon_i}$$ رابطه (۳)

$$Min \sum_{i,j=1}^{n} x_i x_j \sigma_{ij}$$ رابطه (۴)

Subject to

$$\sum_{i=1}^{n} x_i = 1$$ رابطه (۵)

$$0 \le x_i \le 1 \qquad i = 1,2,...,n$$ رابطه (۶)

رابطه (۲)-(۳) نشان‌دهنده حداکثرکردن بازده مورد انتظار پرتفوی و میانگین بازده غیرعادی دوره نمونه می‌باشد. رابطه (۴) نشان‌دهنده حداقل کردن سنجه ریسک نامطلوب می‌باشد. رابطه (۶) نیز نشان‌دهنده این موضوع هست که فروش استقراضی مجاز نمی‌باشد.

در این مقاله از پیش‌بینی بازده مورد انتظار دوره ۳۰۴ روزه، یعنی فروردین ۱۴۰۰ تا تیر ۱۴۰۱ استفاده شده است.

روش مجموع وزندار اغلب برای تبدیل بهینه‌سازی نمونه کارها چند هدف فوق به یک مدل هدف واحد استفاده می‌شود. بنابراین، با در نظر گرفتن وزن‌های یکسان مدل میانگین واریانس با پیش‌بینی به رابطه (۷) تبدیل می‌شود.

$$Max \sum_{i=1}^{n} x_i r_i - \sum_{i,j=1}^{n} x_i x_j \sigma_{ij} - \sum_{i=1}^{n} x_i \overline{\varepsilon_i}$$ رابطه (۷)

Subject to

$$\sum_{i=1}^{n} x_i = 1$$ رابطه (۸)

$$0 \le x_i \le 1 \qquad i = 1,2,....,n$$ رابطه (۹)

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

## ۴. تحلیل داده‌ها و یافته‌ها

پژوهش حاضر به دنبال ارائه مدلی برای انتخاب بهینه سهام با استفاده از پیش‌بینی بازدهی مورد انتظار به کمک شبکه عصبی و جنگل تصادفی و به دست آوردن نسبت تخصیص پرتفوی به هر صنعت به کمک ترکیب بازدهی‌های به‌دست آمده با مدل مارکوییتز می‌باشد.

اطلاعات مربوط به پژوهش از طریق بانک اطلاعات کتابخانه بورس اوراق بهادار تهران و نرم افزار رهاورد نوین گردآوری شده است، همچنین جهت پردازش داده‌ها و تشکیل مدل‌های مورد نیاز از نرم افزار پایتون استفاده شده است. از مجموع ۱۲۰۱ عدد داده به‌دست آمده حد فاصل تاریخ ۱۳۹۶ تا ۱۴۰۱، ۸۹۷ داده ابتدایی یعنی سال‌های ۱۳۹۶ تا ابتدای سال ۱۴۰۰ برای آموزش مدل استفاده شده و ۳۰۴ عدد داده باقی مانده برای سنجش عملکرد مدل و همچنین ارائه مدلی جهت پیدا کردن سبد سهام بهینه می‌باشد.

### سنجش عملکرد مدل‌های پیش‌بینی

به منظور اندازه‌گیری جامع عملکرد مدل‌های مختلف در فرآیند پیش‌بینی بازده سهام، پنج معیار، یعنی میانگین مربعات خطا ، میانگین خطای مطلق $H_R$، $H_{R+}$ و $H_{R-}$ در این مقاله اعمال می‌شوند. همانطور که این معیارها به وضوح مدل را نشان می‌دهند توانایی پیش‌بینی، به طور گسترده ای به عنوان معیارهای عملکرد مورد استفاده قرار می‌گیرند. این معیارها به صورت زیر تعریف می‌شوند.

رابطه (۱۰)
$$\mathrm{MSE}=\frac{1}{N}\sum_{t=1}^{N}(r_t-r_t)^2$$

رابطه (۱۱)
$$\mathrm{MAE}=\frac{1}{N}\sum_{t=1}^{N}|r_t-r_t|$$

رابطه (۱۲)
$$\mathrm{H_R}=\frac{\mathrm{Count}_{t=1}^{n}(r_t r_t>0)}{\mathrm{Count}_{t=1}^{n}(r_t r_t\neq 0)}$$

رابطه (۱۳)
$$\mathrm{H_{R+}}=\frac{\mathrm{Count}_{t=1}^{n}(r_t>0 \text{ And } r_t>0)}{\mathrm{Count}_{t=1}^{n}(r_t>0)}$$

رابطه (۱۴)
$$\mathrm{H_{R-}}=\frac{\mathrm{Count}_{t=1}^{n}(r_t<0 \text{ And } r_t<0)}{\mathrm{Count}_{t=1}^{n}(r_t<0)}$$

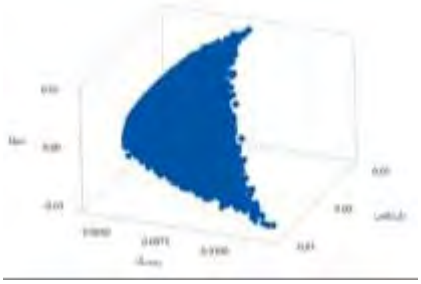
که در آن $r$، $\hat{r}$ به ترتیب نشان‌دهنده بازده واقعی و بازده پیش‌بینی‌کننده در زمان $t$ هستند. رابطه (۱۰) و رابطه (۱۱) به ترتیب بیانگر میانگین مربعات خطا و میانگین مطلق خطا می‌باشند علاوه بر این رابطه (۱۲) نشان‌دهنده نرخ کل پیش‌بینی درست از نظر مثبت با منفی بودن نماد در آن روز، رابطه (۱۳) به معنای دقت پیش‌بینی روزهای مثبت و رابطه (۱۴) به معنی دقت پیش‌بینی روزهایی که بازدهی منفی است می‌باشد.

با توجه به جدول (۳) عملکرد الگوریتم‌های یادگیری ماشین و شبکه عصبی در پیش‌بینی جهت بازار و میزان خطا به مراتب بهتر از الگوریتم سری زمانی بوده و همچنین اگرچه میزان درستی پیش‌بینی از منظر جهت بازار در جنگل تصادفی بهتر بوده اما میزان خطا در پیش‌بینی به وسیله شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی کمتر از جنگل تصادفی می‌باشد و این امر نشان‌دهنده برتری این مدل در پیش‌بینی می‌باشد.

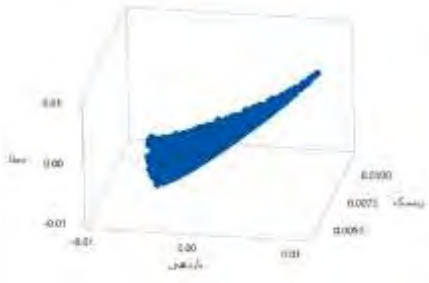
در ادامه جهت بررسی دقیق‌تر بازدهی‌های پیش‌بینی‌شده را در مدل میانگین واریانس با پیش‌بینی قرار می‌دهیم تا به شبیه‌سازی ۲۰۰۰۰ پرتفوی با توجه به پیش‌بینی به دست آمده از شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی را انجام دهیم. نتایج این شبیه سازی در شکل (۲) آورده شده است، این موضوع می‌تواند به ما کمک نماید تا مرز کارایی را مشخص نماییم و ترکیب بهینه مناسب را پیدا نماییم.

این امر به ما کمک می‌کند تا بتوانیم مقایسه‌ای بین عملکرد سبد تشکیل شده به‌وسیله بازدهی‌های پیش‌بینی شده با سبد تشکیل شده به‌وسیله بازدهی‌های دوره مشابه قبل انجام دهیم.

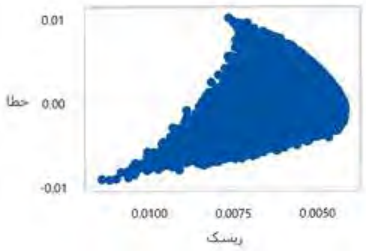
*نمودار۱. ارزش لحظه ای پرتفوهای به دست آمده با مدل نیم واریانس*



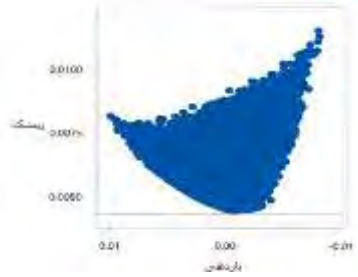
(الف) دید ۳ بعدی: بازدهی، ریسک، خطا



(ب) دید ۳ بعدی: ریسک، بازدهی، خطا




(ج) دید ۲ بعدی: ریسک، خطا



(د) دید ۲ بعدی: بازدهی، ریسک

*شکل۲. شبیه سازی ۲۰۰۰۰ پرتفوی به کمک پیش‌بینی بازدهی با جنگل تصادفی*

*جدول۳. خطاهای بدست آمده از مدل‌ها*

| مدل | | MAE | MSE | $H_R$ | $H_{R^+}$ | $H_{R^-}$ |
|---|---|---|---|---|---|---|
| ARIMA | mean | ۰/۰۱۳۳۰۷۷ | ۰/۰۰۰۶۰۶۳ | ۵۳/۴۴٪ | ۵۳/۶۳٪ | ۵۵/۲۸٪ |
| | SD | ۰/۰۰۵۰۰۷۷ | ۰/۰۰۰۱۹۳۸ | ۰/۰۷۷۷۰۱۸ | ۰/۰۹۸۷۶۳۴ | ۰/۱۴۸۰۷۷۷ |
| LSTM | mean | ۰/۰۱۴۲۴۱۵ | ۰/۰۰۰۳۵۲۷ | ۵۵/۵۶٪ | ۵۴/۸۷٪ | ۶۶/۲۷٪ |
| | SD | ۰/۰۰۴۵۷۱۴ | ۰/۰۰۰۲۳۸۹ | ۰/۰۲۱۸۸۹۷ | ۰/۰۶۷۰۳۹۴ | ۰/۱۱۸۹۹۳۵ |
| RF | mean | ۰/۰۱۴۴۹۰۹ | ۰/۰۰۰۴۹۹۹ | ۶۰/۴۶٪ | ۵۷/۹۵٪ | ۶۳/۰۳٪ |
| | SD | ۰/۰۰۴۶۹۹۵ | ۰/۰۰۰۴۹۶۷ | ۰/۰۴۲۱۱۹۴ | ۰/۰۶۰۷۷۴۵ | ۰/۰۴۲۰۳۳۷ |

در نهایت این مقاله جهت مقایسه عملکرد عملی روش‌های مختلف پیش‌بینی با فرض سرمایه‌گذاری بلند مدت یک ساله، به کمک بازدهی پیش‌بینی شده توسط روش‌های مختلف و ترکیب آن‌ها با مدل بهینه‌ساز میانگین واریانس با پیش‌بینی، اوزان بهینه سبد را با در نظر گرفتن سطح ریسک‌پذیری جهت سرمایه‌گذاری محاسبه می‌کنیم.

نمودار (۱) نشان می‌دهد با فرض سرمایه‌گذاری ۱ واحدی در صنایع پیشنهادی در ابتدای دوره، عملکرد نهایی سبدهای محاسبه شده با ۳ روش مختلف پیش‌بینی در انتهای دوره به چه

صورت خواهد بود. در این نمودار به وضوح مشخص است اگر چه در کوتاه مدت عملکرد مدل به‌دست آمده به کمک شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی مناسب نبوده اما در بلند مدت و انتهای دوره سود به دست آمده به کمک شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی بسیار بالاتر از باقی روش‌ها بوده است.

مدل به دست آمده توسط الگوریتم جنگل تصادفی اگر چه سود کمتری نسبت به شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی دارد اما نوسان کمتری نیز داشته است.

## ۵. بحث و نتیجه‌گیری

در این پژوهش، یک مدل جدید جهت پیش‌بینی بازده شاخص‌های ۵ صنعت به کمک یادگیری عمیق و یادگیری ماشین و ترکیب این بازدهی‌ها با مدل میانگین واریانس مارکویتز جهت تشکیل پرتفوی بهینه مورد بررسی قرار گرفت.

یک مطالعه تجربی به کمک داده‌های شاخص ۵ صنعت بانکی، خودرویی، دارویی، فلزی و نفتی حد فاصل فروردین۱۳۹۶ تا تیر ۱۴۰۱ مورد بررسی قرار گرفت.

در اکثر موارد معاملگران جهت انجام معامله سعی در بررسی سهام و نمادهای خاص دارند بی‌آنکه صنایع مربوط به هر نماد به طور جداگانه مورد بررسی قرار گیرند. در این مقاله سعی بر انتخاب صنعت مناسب جهت سرمایه‌گذاری بلندمدت شده است.

مدل‌های استفاده‌شده در تحقیق جهت پیش‌بینی بازدهی مدل شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی، جنگل تصادفی و روش سری زمانی ARIMA می‌باشد که با مقایسه نتایج به دست آمده این موضوع آشکار که می‌شود که مدل شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی به مراتب عملکرد بهتری در پیش‌بینی بازدهی داشته اند.

مدل بهینه‌سازی سبد سرمایه‌گذاری ارائه‌شده در این مقاله ترکیب مدل میانگین واریانس مارکویتز با بازده‌های پیش‌بینی‌شده می‌باشد. این مدل دارای دو تابع هدف بیشینه بازده مورد انتظار پرتفوی و میانگین بازده غیرعادی دوره نمونه وکمینه‌سازی سنجه ریسک نامطلوب می‌باشد. محدودیت‌های مدل شامل: محدودیت نقدینگی و محدودیت سقف و کف سرمایه‌گذاری است.

در نهایت با ضرایب به دست آمده جهت تخصیص پرتفوی به ۵ صنعت سعی در سنجش عملکرد مدل‌های این مقاله با مدل کلاسیک ترکیب بازده گذشته با مدل میانگین واریانس مارکویتز در دوره ای مابین فروردین۱۴۰۰ تا تیر ۱۴۰۱ شد.

نتایج به دست آمده از برتری کامل مدل میانگین واریانس به کمک پیش‌بینی نسبت به مدل‌های کلاسیک دارند. در مدل به دست آمده به‌وسیله شبکه عصبی حافظه‌ی کوتاه‌مدت طولانی سود بیشتری حاصل گردید. اما این پرتفوی نوسان بیشتری نیز داشت. در مدل به‌دست‌آمده به وسیله جنگل تصادفی گرچه سود ما کمتر از روش جنگل تصادفی بوده است اما این پرتفوی نوسان کمتری را در طول بازه ۳۰۴ روزه تجربه کرده است.

## ۶. پیشنهادها و محدودیت‌ها

پیشنهاد می‌شود، در پژوهش‌های آتی برخی عوامل مالی، رفتاری نیز علاوه بر شاخ‌های درنظر گرفته شده در این پژوهش، مورد بررسی قرار گیرد. علاوه بر این در بخش تبدیل مسئله چند هدفه بازدهی، ریسک و خطا به مدل میانگین واریانس با پیش‌بینی از روش هم وزن جهت تک هدفه نمودن استفاده شده است. توصیه می‌گردد در پژوهشات آتی از سایر روش‌های برنامه ریزی چند هدفه استفاده گردد. در آخر با توجه به ماهیت بازدهی، ریسک و خطا استفاده از روش‌های عدم قطعیت در مدل ارائه شده می‌تواند موضوعی برای پژوهش‌های آتی نیز باشد.

## 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.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

## منابع


1. Abdi, N., Moradzadeh Fard, M., Ahmadzadeh, H., & Khoddam, M. 2021. "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), 119-143.* (*in Persian*)
2. Alizadeh, Meysam, Roy Rada, Fariborz Jolai, and Elnaz Fotoohi. 2011. "An Adaptive Neuro-Fuzzy System for Stock Portfolio Analysis." *International Journal of Intelligent Systems* 26(2):99–114. doi: 10.1002/INT.20456.
3. Assimakopoulos, Spyros, Evangelos Makridakis, and Vassilios Spiliotis. 2018. "Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward." *PLoS ONE* 13:26.
4. Azar, Adel, Amir Afsar, and Parviz Ahmadi. 2007. "Comparison of Classical Methods and Artificial Intelligence in Predicting the Stock Price Index and Designing a Hybrid Model." *Lecturer of Humanities*.
5. Box, G.E.P. and Jenkins, G. M. 1976. "Time Series Analysis: Forecasting and Control". *Holden Day, San Francisco.* (1970).
6. Chong, Eunsuk, Chulwoo Han, and Frank C. Park. 2017. "Deep Learning Networks for Stock Market Analysis and Prediction: Methodology, Data Representations, and Case Studies." *Expert Systems with Applications* 83:187–205. doi: 10.1016/j.eswa.2017.04.030.
7. Deng, Shijie, and Xinyu Min. 2013. "Applied Optimization in Global Efficient Portfolio Construction Using Earning Forecasts." *The Journal of Investing* 22(4):104–14. doi: 10.3905/JOI.2013.22.4.104.
8. Fischer, Thomas, and Christopher Krauss. 2018. "Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions." *European Journal of Operational Research* 270(2):654–69. doi: 10.1016/j.ejor.2017.11.054.
9. Heydari Zare, Behzad, and Hamidreza Kordloui. 2011. "Stock Price Prediction Using Artificial Neural Network." *Journal of Industrial Strategic Management* 7(17). (*in Persian*)
10. Ho, Tin Kam. 1995. "Random Decision Forests." *Proceedings of the International Conference on Document Analysis and Recognition, ICDAR* 1:278–82. doi: 10.1109/ICDAR.1995.598994.
11. Huang, Chien Feng. 2012. "A Hybrid Stock Selection Model Using Genetic Algorithms and Support Vector Regression." *Applied Soft Computing* 12(2):807–18. doi: 10.1016/J.ASOC.2011.10.009.
12. Kingma, Diederik P., and Jimmy Lei Ba. 2015. "Adam: A Method for Stochastic Optimization." in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*.
13. Lee, Sang Il, and Seong Joon Yoo. 2018. "Threshold-Based Portfolio: The Role of the Threshold and Its Applications." *The Journal of Supercomputing 2018 76:10* 76(10):8040–57. doi: 10.1007/S11227-018-2577-1.
14. Markowitz, Harry. 1952. "PORTFOLIO SELECTION." *The Journal of Finance* 7(1):77–91. doi: 10.1111/J.1540-6261.1952.TB01525.X.
15. Mohebbi, Somayeh, Mohamad Esmaeil Fadaeinejad, and Mohammad reza Hamidizadeh. 2021. "The Proposed Algorithm to Select Appropriate Features for Predicting Tehran Stock Exchange Index." *Journal of Financial*

*Management Perspective* 11(34):35–67. doi: 10.52547/jfmp.11.34.35. (*in Persian*)

16. Oussar, Yacine, and Gérard Dreyfus. 2000. "Initialization by Selection for Wavelet Network Training." *Neurocomputing* 34(1–4):131–43. doi: 10.1016/S0925-2312(00)00295-2.
17. Reza, Rai. 2003. "Stock Portfolio Formation for the Venture Capitalist Comparing Neural Network and Markowitz." *Business Management Perspective* (2):77-. (*in Persian*)
18. Soleymani Sarvestani, Sajad, Sayyed Mohammad Reza davoodi, and Ali Kheradmand. 2022. "Optimal Daily scalping trading portfolio based on interval-valued prediction with ANN approach." *Journal of Financial Managment Perspective* 12(39):103–20. doi: 10.52547/jfmp.12.39.103. (*in Persian*)
19. Ta, Van Dai, Chuan Ming Liu, and Direselign Addis Tadesse. 2020. "Portfolio Optimization-Based Stock Prediction Using Long-Short Term Memory Network in Quantitative Trading." *Applied Sciences (Switzerland)* 10(2).
20. Ustun, Ozden, and Refail Kasimbeyli. 2012. "Combined Forecasts in Portfolio Optimization: A Generalized Approach." *Computers & Operations Research* 39(4):805–19. doi: 10.1016/J.COR.2010.09.008.
21. Wang, Wuyu, Weizi Li, Ning Zhang, and Kecheng Liu. 2020. "Portfolio Formation with Preselection Using Deep Learning from Long-Term Financial Data." *Expert Systems with Applications* 143:113042. doi: 10.1016/j.eswa.2019.113042.
22. Yu, Jing Rung, Wan Jiun Paul Chiou, Wen Yi Lee, and Shun Ji Lin. 2020. "Portfolio Models with Return Forecasting and Transaction Costs." *International Review of Economics and Finance* 66:118–30.

**استناد**

اقتصاد ، امیرعلی و محمدی، عمران (۱۴۰۲). بهینه‌سازی سبد سرمایه‌گذاری به کمک پیش‌بینی بازده مورد انتظار با استفاده از روش‌های شبکه عصبی LSTM، جنگل تصادفی و ARIMA. *چشم‌انداز مدیریت مالی*، ۱۳(۴۳)، ۹-۲۸.

پژوهشگاه علوم انسانی و مطالعات فرهنگی

**Citation**

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*)