

A Systematic Review of the Application of Data Mining and Machine Learning Methods in Stock Valuation, Along with a Proposed Hybrid Framework for the Sustainability of Equity Funds in the Tehran Stock Exchange

Hossein Ashrafi Soltan Ahmadi ¹, Saeed Jabarzadeh Kangarlouei ^{2*}, Jamal Bahri Sales ¹, Ali Ashtab ³

¹ Department of Accounting, Urmia Branch, Islamic Azad University, Urmia, Iran

^{2*} Department of Accounting, Urmia Branch, Islamic Azad University, Urmia, Iran (Corresponding author), Email: s.jabarzade@iau.ac.ir

³ Department of Accounting, Urmia University, Urmia, Iran

ARTICLE INFO

Article history:

Received: 23/07/2025

Received in revised form: 02/09/2025

Accepted: 23/09/2025

Available online: 23/09/2025

Keywords:

Data mining

Machine learning

Stock valuation; Capital market indicators

Decision tree

Neural network; portfolio management

Conditional Value-at-Risk (CVaR)

ABSTRACT

Background and Objective: Predicting stock prices and valuing stocks have consistently been major challenges for investors and researchers in the field of finance. With advancements in artificial intelligence and data mining technologies, novel approaches to modeling and forecasting stock value have emerged. The aim of this research is to systematically review existing studies on the application of data mining and machine learning methods in stock valuation and, based on identified research gaps, to propose a hybrid framework for the sustainability of equity funds in the Tehran Stock Exchange.

Method: This study employs a systematic review methodology following the PRISMA approach. Domestic databases (Iranian Scientific-Research Databases, Magiran, and SID) and international databases (ScienceDirect, Springer, IEEE, and Google Scholar) were searched for the period 2000–2024; following a screening process, 68 articles (28 domestic and 40 international) were selected for final analysis.

Findings: Data mining methods are categorized into five main groups: (1) decision trees and rule-based models, (2) neural networks and deep learning, (3) support vector machines, (4) Bayesian methods, and (5) hybrid and meta-heuristic methods. Decision tree models (C5, CART, and CHAID) and neural networks have been the most widely used for stock value prediction, and the research trend is shifting from traditional statistical models toward machine learning and deep learning. Furthermore, cash-based indicators demonstrate higher prediction accuracy compared to structural and expectation-based indicators. Based on identified research gaps, a two-stage hybrid framework—comprising prediction using a Bi-LSTM neural network with a Huber loss function and robust optimization based on CVaR—was proposed and preliminarily evaluated.

Conclusion: Despite the expansion of research, gaps remain regarding the comprehensive modeling of capital market indicators, the utilization of macroeconomic data, the systematic comparison of methods, and the integration of intelligent prediction with risk management. By presenting a comprehensive classification framework and a proposed hybrid framework, this study outlines the path for future research.

Article Type: Research Paper



©Authors

Journal of Intelligent Financial Management,
2025, Vol. 1, No.3, pp. 111- 136

Publish by:

Tolou-e Binsh-e Ayandeh Scientific Institute

<https://doi.org/10.25843/JIFM.2025.8563.236550>

Cite: Ashrafi Soltan Ahmadi, H., Jabarzadeh Kangarlouei, S., Bahri Sales, J. & Ashtab, A. (2025). A Systematic Review of the Application of Data Mining and Machine Learning Methods in Stock Valuation, Along with a Proposed Hybrid Framework for the Sustainability of Equity Funds in the Tehran Stock Exchange. Journal of Intelligent Financial Management, 1(3), 111-130.

مروری نظام مند بر کاربرد روش های داده کاوی و یادگیری ماشین در ارزش گذاری سهام، همراه با یک چارچوب ترکیبی پیشنهادی برای پایداری صندوق های سهامی بورس تهران

حسین اشرفی سلطان احمدی^۱، سعید جبارزاده کنگرلویی^{۲*}، جمال بحری ثالث^۱، علی آشتاب^۳

۱ - گروه حسابداری، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران

۲ و ۳ - گروه حسابداری، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران (نویسنده مسئول): s.jabarzade@iau.ac.ir ایمیل نویسنده مسئول

۳ - گروه حسابداری، دانشگاه ارومیه، ارومیه، ایران

اطلاعات مقاله

تاریخچه مقاله:

تاریخ دریافت: ۱۴۰۴/۰۵/۰۱

تاریخ بازنگری: ۱۴۰۴/۰۶/۱۱

تاریخ پذیرش: ۱۴۰۴/۰۷/۰۷

تاریخ انتشار: ۱۴۰۴/۰۷/۰۱

کلیدواژه ها:

داده کاوی

یادگیری ماشین

ارزش گذاری سهام

نماگرهای بازار سرمایه

درخت تصمیم

شبکه عصبی

مدیریت پورتفوی

ارزش در معرض ریسک شرطی (CVaR)

چکیده

زمینه و هدف: پیش بینی قیمت سهام و ارزش گذاری آن همواره یکی از چالش های اصلی سرمایه گذاران و پژوهشگران حوزه مالی بوده است. با پیشرفت فناوری های هوش مصنوعی و داده کاوی، رویکردهای نوینی برای مدل سازی و پیش بینی ارزش سهام شکل گرفته است. هدف این پژوهش، مرور نظام مند تحقیقات انجام شده در زمینه کاربرد روش های داده کاوی و یادگیری ماشین در ارزش گذاری سهام و، بر پایه شکاف های شناسایی شده، ارائه یک چارچوب ترکیبی برای پایداری صندوق های سهامی بورس اوراق بهادار تهران است.

روش: این مطالعه از روش مرور نظام مند (Review Systematic) با رویکرد PRISMA بهره برده است. پایگاه های داده داخلی (پایگاه های علمی-پژوهشی ایران، مگیران و SID) و خارجی (Scholar Google، IEEE، Springer، ScienceDirect) در بازه زمانی ۲۰۰۰ تا ۲۰۲۴ جست و جو شدند و پس از غربالگری، ۶۸ مقاله (۲۸ داخلی و ۴۰ خارجی) برای تحلیل نهایی انتخاب شدند.

یافته ها: روش های داده کاوی به پنج دسته اصلی تقسیم می شوند: (۱) درخت تصمیم و مدل های مبتنی بر قانون، (۲) شبکه های عصبی و یادگیری عمیق، (۳) ماشین بردار پشتیبان، (۴) روش های بیزین و (۵) روش های ترکیبی و فراابتکاری. مدل های درخت تصمیم C، CART و CHAID و شبکه های عصبی بیشترین کاربرد را در پیش بینی ارزش سهام داشته اند و روند تحقیقات از مدل های آماری سنتی به سمت یادگیری ماشین و یادگیری عمیق در حال حرکت است. همچنین نماگرهای نقدی نسبت به نماگرهای ساختاری و انتظاری از دقت پیش بینی بالاتری برخوردارند. بر پایه شکاف های شناسایی شده، یک چارچوب ترکیبی دوحله ای شامل پیش بینی با شبکه عصبی Bi-LSTM و تابع زیان Huber و بهینه سازی استوار مبتنی بر CVaR پیشنهاد و به طور مقدماتی ارزیابی شد.

نتیجه گیری: با وجود گسترش پژوهش ها، همچنان شکاف هایی در زمینه مدل سازی جامع نماگرهای بازار سرمایه، استفاده از داده های کلان اقتصادی، مقایسه نظام مند روش ها و پیوند پیش بینی هوشمند با مدیریت ریسک وجود دارد. پژوهش حاضر با ارائه چارچوب طبقه بندی جامع و یک چارچوب ترکیبی پیشنهادی، مسیر تحقیقات آتی را ترسیم می کند.

نوع مقاله: پژوهشی



© نویسندگان

استناد: اشرفی سلطان احمدی، حسین و جبارزاده کنگرلویی، سعید و بحری ثالث، جمال و آشتاب، علی. (۱۴۰۴). مروری نظام مند بر کاربرد روش های داده کاوی و یادگیری ماشین در ارزش گذاری سهام، همراه با یک چارچوب ترکیبی پیشنهادی برای پایداری صندوق های سهامی بورس تهران. مدیریت مالی هوشمند، (۳)، ۱۱۱-۱۳۰.

نشریه مدیریت مالی هوشمند، ۱۴۰۴، دوره ۱، شماره ۳، صفحه ۱۱۱-۱۳۰.

ناشر: موسسه علمی طلوع بینش آینده

<https://doi.org/10.25843/JIFM.2025.8563.236550>

مقدمه

بازار سرمایه به عنوان یکی از ارکان اصلی اقتصاد هر کشور، نقش حیاتی در تأمین مالی بنگاه‌ها و هدایت پس‌اندازها به سمت سرمایه‌گذاری‌های مولد ایفا می‌کند. امروزه سرمایه‌گذاری در بورس بخش مهمی از اقتصاد را تشکیل می‌دهد و به همین دلیل، پیش‌بینی قیمت سهام و شناسایی ارزش ذاتی آن از مهم‌ترین دغدغه‌های سرمایه‌گذاران است. پیش‌بینی قیمت سهام و بازشناسی رفتار آن امری پیچیده است، زیرا مؤلفه‌های بسیاری از قبیل رخدادهای سیاسی، وضع عمومی اقتصاد و انتظارات سرمایه‌گذاران بر بازار سهام تأثیر می‌گذارد. [1, 2] فعالیت در بازار سرمایه امری خطیر است که موفقیت در آن به گردآوری اطلاعات جامع و مطلوب از شرایط بازار نیاز دارد؛ چنانچه بتوان روند آتی بازار سهام را با روش‌های مناسب پیش‌بینی کرد، سرمایه‌گذار می‌تواند بازده حاصل از سرمایه‌گذاری خود را بیشینه سازد. [3, 4]

در ادبیات مالی-اقتصادی، روش‌های گوناگونی برای پیش‌بینی بازدهی سرمایه‌گذاری ارائه شده است که می‌توان آن‌ها را در چهار دسته اصلی قرار داد: تحلیل تکنیکال، تحلیل بنیادی، تحلیل احساسات و روش‌های مبتنی بر داده‌کاوی و یادگیری ماشین. [5] اگرچه مدل‌های تکنیکی به‌طور گسترده در پیش‌بینی بازار سهام به کار رفته‌اند، نتایج مطالعات نشان می‌دهد که این روش‌ها به اندازه کافی موفق نبوده‌اند و با پیشرفت علم، پژوهشگران از مدل‌های سری زمانی و یادگیری ماشین برای پیش‌بینی بهتر استفاده کرده‌اند. در سال‌های اخیر، با پیشرفت سریع هوش مصنوعی، روش‌های داده‌کاوی و یادگیری ماشین توانسته‌اند پیچیدگی‌های سری‌های زمانی قیمت سهام را تحلیل کنند و برتری خود را نسبت به روش‌های سنتی نشان دهند. [6, 7] داده‌کاوی به عنوان فرایند کشف دانش از حجم زیادی از داده‌ها، در حوزه‌های گوناگون از جمله تحلیل بازار، کشف تقلب و پیش‌بینی رفتار مالی کاربرد یافته است. استخراج خودکار دانش جدید و مفید از منابع داده‌ای حجیم طی یک فرایند غیربديهی، داده‌کاوی نامیده می‌شود. [8] با توجه به گسترش روزافزون پژوهش‌ها در این حوزه، انجام یک مرور نظام‌مند که تصویری جامع از وضعیت موجود ارائه دهد، ضروری به نظر می‌رسد.

مبانی نظری و پیشینه پژوهش

در دهه اخیر، مدیریت پورتفوی از رویکردهای سنتی مبتنی بر میانگین-واریانس به سمت مدل‌های پیشرفته‌تر و مبتنی بر هوش مصنوعی حرکت کرده است. در بازارهای نوظهوری مانند بورس تهران، ویژگی‌هایی نظیر عدم تقارن اطلاعاتی، نوسانات شدید و خوشه‌ای (Volatility Clustering) و نویزهای قیمتی بالا باعث می‌شود که مدل‌های آماری خطی در پیش‌بینی روند بازار و مدیریت ریسک ناکارآمد عمل کنند. [9] در بازارهای مالی مدرن، عدم تقارن اطلاعاتی و رفتارهای هیجانی معامله‌گران باعث می‌شود که قیمت‌ها همواره با ارزش ذاتی فاصله داشته باشند. روش‌های سنتی مانند مدل‌های رگرسیون چندمتغیره (Multivariate Regression) فرض می‌کنند که رابطه میان متغیرهای اقتصادی (مانند نرخ سود و نسبت P/E) و قیمت سهام خطی است، در حالی که واقعیت بازار پیچیده و غیرخطی است. [10, 11] این پژوهش با هدف پاسخ به پرسش‌های زیر انجام شده است:

- پرسش 1. روش‌های داده‌کاوی و یادگیری ماشین مورد استفاده در ارزش‌گذاری سهام کدام‌اند و چگونه طبقه‌بندی می‌شوند؟
- پرسش 2. روند تحقیقات داخلی و خارجی در این حوزه چگونه است؟
- پرسش 3. شکاف‌های تحقیقاتی و جهت‌گیری‌های آتی چیست؟
- پرسش 4. چگونه می‌توان با طراحی یک چارچوب ترکیبی که بازدهی منصفانه (Fair Value) را با استفاده از Bi-LSTM پیش‌بینی کرده و سپس آن را در یک مدل بهینه‌سازی استوار (Robust CVaR) به کار می‌گیرد، پایداری صندوق‌های سهامی را در بورس اوراق بهادار تهران در برابر نوسانات بازار افزایش داد؟
- پرسش 5 و 4) و پرسش چهارم با چارچوب پیشنهادی و ارزیابی مقدماتی آن (بخش 6) پاسخ داده می‌شود.

ارزش‌گذاری سهام

ارزش‌گذاری سهام فرایندی است که طی آن ارزش ذاتی یک سهم تعیین می‌شود. در ادبیات مالی، دو رویکرد اصلی برای ارزش‌گذاری وجود دارد: تحلیل بنیادی و تحلیل تکنیکال. تحلیل بنیادی مبتنی بر این فرض است که اوراق بهادار دارای ارزش ذاتی

اند که سرمایه‌گذاران می‌توانند آن را برآورد کنند. در این رویکرد، برآورد ارزش ذاتی سهام از طریق ارزیابی متغیرهای بنیادی مانند سود هر سهم، ارزش دفتری و جریان نقدی انجام می‌شود. [12-14]

تحلیل تکنیکال، تغییر و حرکت قیمت اوراق بهادار را مدنظر دارد و از داده‌های تاریخی برای پیش‌بینی تغییرات آینده استفاده می‌کند. تکنیکالیست‌ها از نمودارهای قیمت استفاده می‌کنند که اطلاعات ارزشمندی از قبیل حرکات و تغییرات روند قیمت، نوسانات گذشته و حال، و حجم و میزان معاملات تاریخی را در اختیار قرار می‌دهد. [15, 16]

نماگرهای بازار سرمایه

نماگرهای بازار سرمایه به سه دسته اصلی تقسیم می‌شوند: [17]

الف (نماگرهای ساختاری: این نماگرها بر اساس این نظریه بنا شده‌اند که «قیمت‌ها انعکاس رفتار روان‌شناختی کل بازیگران بازار است» و شامل عمق بازار، نرخ تورم، نرخ بهره، نرخ ارز و نرخ طلا هستند. [18, 19]

ب (نماگرهای جریان نقدی: موقعیت مالی گروه سرمایه‌گذاران را تحلیل می‌کنند و شامل سود نقدی هر سهم (DPS)، نسبت بدهی (DEBT)، نسبت قیمت به درآمد (P/E)، نرخ شناوری، حجم معاملات و جریان نقد عملیاتی (OCF) هستند. [20]

ج (نماگرهای انتظاری: رفتار فعالان بازار را با توجه به اطلاعات مورد انتظار آنان اندازه می‌گیرند و شامل عایدی هر سهم (EPS)، نرخ بازده حسابداری (ROA)، نسبت مالکانه (PPRT)، ارزش دفتری (BVP)، نرخ رشد سود (GEPS)، تمرکز مالکیت (OWC)، ارزش افزوده اقتصادی (EVA) و ریسک سیستماتیک (BETA) هستند. اندازه شرکت و ارزش بازار نیز از دیرباز به‌عنوان عوامل مؤثر بر بازده سهام شناخته شده‌اند. [21]

داده‌کاوی و یادگیری ماشین

داده‌کاوی علم استخراج الگوهای مفید از داده‌های خام است و از رشته‌های مختلفی از جمله آمار، هوش مصنوعی، یادگیری ماشین، شناسایی الگو و پایگاه داده نشأت گرفته است. [8, 22]

در سال‌های اخیر، روش‌های داده‌کاوی و هوش مصنوعی در طراحی الگوهای مربوط به بازار سرمایه و همچنین در علوم مختلف مورد استقبال قرار گرفته‌اند. هدف اصلی داده‌کاوی کشف دانش است؛ دانشی که به‌صورت منظمی در داده‌ها وجود دارد. پس از کشف دانش، دو حالت ممکن است رخ دهد: در حالت نخست، خبرگان حوزه داده‌های مورد کاوش از دانش استخراج‌شده آگاه‌اند و آن دانش به‌عنوان یک قانون صحیح تلقی می‌شود؛ در حالت دوم، دانش کشف‌شده دانشی جدید است که در میان متخصصان آن حوزه شناخته‌شده نیست. در این حالت، دانش جدید بررسی می‌شود و در صورت منطقی بودن به فرضیه تبدیل می‌شود و درستی یا نادرستی آن با آزمایش‌ها و بررسی‌های متعدد سنجیده می‌شود؛ این فرضیه در صورت تأیید به قانون تبدیل خواهد شد. [8]

روش‌های داده‌کاوی به دو دسته کلی تقسیم می‌شوند: روش‌های نظارت‌شده (مانند طبقه‌بندی و رگرسیون) و روش‌های نظارت نشده (مانند خوشه‌بندی). (یادگیری ماشین، به‌عنوان زیرمجموعه‌ای از هوش مصنوعی، به سامانه‌هایی گفته می‌شود که بدون برنامه‌ریزی صریح از داده‌ها یاد می‌گیرند و عملکرد خود را بهبود می‌بخشند. الگوریتم‌های یادگیری ماشین در حوزه مالی کاربرد گسترده‌ای یافته‌اند و شامل درخت تصمیم، شبکه عصبی، ماشین بردار پشتیبان، روش‌های بیزین و K-NN هستند. [7, 23]

تکامل نظریه‌های مدیریت پورتفوی

ریشه مدیریت پورتفوی مدرن به مدل کلاسیک مارکویتز بازمی‌گردد که بر پایه مفهوم «بهینه‌سازی میانگین-واریانس» بنا شده است. [24] مارکویتز استدلال کرد که سرمایه‌گذار باید سبدی را انتخاب کند که در یک سطح مشخص از ریسک (واریانس)، بیشترین بازده را ارائه دهد. نظریه مدرن پورتفوی (MPT) سنگ‌بنای مدیریت سرمایه‌گذاری بوده است؛ با این حال، پژوهش‌های متعدد نشان داده‌اند که مدل مارکویتز با دو چالش بنیادین روبه‌روست: نخست، فرض توزیع نرمال برای بازده دارایی‌ها، و دوم، حساسیت بیش از حد به خطاهای تخمین بازده و کوواریانس. [25, 26]

مطالعات تجربی نشان می‌دهند که سری‌های زمانی بازدهی سهام دارای ویژگی‌هایی نظیر «توزیع‌های با دم پهن (Fat-tailed distributions)»، کشیدگی (Kurtosis) «بالا و» خوشه‌های نوسانی (Volatility Clustering) «هستند» [10]. این بدان معناست که احتمال وقوع رویدادهای حدی (بحران‌های بزرگ) بسیار بیشتر از آن چیزی است که مدل‌های کلاسیک پیش‌بینی می‌کنند. این ویژگی‌ها در بازارهای نوظهور و پرنوسانی نظیر بورس تهران شدیدتر است و مدل‌های میانگین-

واریانس سستی را در پیش‌بینی ریسک‌های حدی ناکارآمد می‌سازد. در پاسخ به این چالش، استفاده از معیارهای ریسک جایگزین نظیر ارزش در معرض ریسک شرطی (CVaR) به‌طور گسترده‌ای مورد توجه قرار گرفته است. [27]

از واریانس تا مدیریت ریسک‌های حدی (Tail Risk)

با شناسایی کاستی‌های مدل مارکویتز، پژوهشگران به سمت معیارهای جایگزین حرکت کردند. یکی از مهم‌ترین این مفاهیم، ارزش در معرض ریسک (Value at Risk, VaR) است. با این حال، VaR تنها حد پایین بازده را در یک سطح اطمینان مشخص نشان می‌دهد و از آنچه در «دم توزیع»، یعنی فراتر از حد VaR، رخ می‌دهد بی‌خبر است. این کاستی به ظهور مفهوم ارزش در معرض ریسک شرطی (Conditional Value at Risk, CVaR)، یا «ریسک انتظاری»، انجامید. [27]

CVaR با تمرکز بر میانگین زیان‌های فراتر از حد VaR، معیاری استوارتر (Robust) برای مدیریت ریسک‌های سیستماتیک و بحران‌های نقدینگی در بازارهای مالی فراهم می‌کند. پژوهش‌ها نشان داده‌اند که بهینه‌سازی استوار CVaR در انتخاب سبد سهام، پایداری پورتنوی را در برابر عدم قطعیت پارامترها و شوک‌های بازار افزایش می‌دهد. [25, 28]

نقش یادگیری عمیق در پیش‌بینی سری‌های زمانی مالی

در کنار تکامل مدل‌های بهینه‌سازی، حوزه پیش‌بینی قیمت نیز شاهد تحولی بنیادین بوده است. با ظهور یادگیری ماشین، مدل‌های آماری خطی مانند ARIMA، که بر پایه خطی بودن روابط بنا شده بودند، در مواجهه با ماهیت غیرخطی و پیچیده بازارهای مالی کارایی خود را از دست دادند و جای خود را به مدل‌های غیرخطی قدرتمندتر دادند. [6, 11]

ظهور شبکه‌های عصبی مصنوعی (ANN) و سپس شبکه‌های حافظه طولانی کوتاه‌مدت (Long Short-Term Memory, LSTM) تحولی در تحلیل سری‌های زمانی ایجاد کرد. [29]

پژوهش‌های پیشرو نشان داده‌اند که الگوریتم‌های یادگیری عمیق، به‌ویژه LSTM، قادر به استخراج الگوهای پنهان در داده‌های سری زمانی مالی هستند. [11, 30]

شبکه‌های LSTM به دلیل توانایی در حفظ حافظه بلندمدت، وابستگی‌های زمانی (Temporal Dependencies) را در داده‌های قیمت و حجم یاد می‌گیرند. با این حال، مدل‌های LSTM استاندارد اطلاعات را فقط در یک جهت پردازش می‌کنند. مدل‌های Bi-LSTM (Bidirectional LSTM) با استفاده از دو لایه مجزای پیش‌رو (Forward) و پس‌رو (Backward)، بافتار (Context) کامل‌تری از حرکت قیمت‌ها را استخراج می‌کنند؛ این ویژگی می‌تواند به کاهش خطای پیش‌بینی، به‌ویژه در نقاط بازگشتی بازار، منجر شود. [31]

با این وجود، پژوهش‌های موجود عمدتاً بر «پیش‌بینی قیمت» متمرکز بوده‌اند و کمتر به «ادغام خروجی این پیش‌بینی‌ها در مدل‌های بهینه‌سازی ریسک» پرداخته‌اند. [32]

استفاده از ساختار Bi-LSTM در ادبیات مالی نوپاست و به‌طور خاص در بورس تهران، خلأ تحقیقاتی قابل‌توجهی در بهره‌گیری از این ساختار برای بهبود پایداری پورتنوی وجود دارد.

پیشینه خارجی

لی و وو در مقاله‌ای به بررسی چگونگی توصیف سبک‌های بازار برای بهبود عملکرد پیش‌بینی سهام تحت سبک‌های مختلف بازار پرداختند. [33]

آنان یک چارچوب پیش‌بینی قیمت سهام ساختند که گرایش‌های آینده قیمت سهام را بر اساس داده‌های متعلق به همان سبک بازار پیش‌بینی می‌کند. آزمایش‌ها با پنج سال داده از بازار بورس هنگ‌کنگ، شامل قیمت‌های سهام و اخبار مربوط، انجام شد و مدل‌های پیش‌بینی با و بدون در نظر گرفتن سبک‌های بازار مقایسه شدند. نتایج نشان داد که رویکرد مبتنی بر سبک‌های بازار، از رویکردی که سبک‌های بازار را در نظر نمی‌گیرد بهتر عمل می‌کند.

دنگ و همکاران با استفاده از شوک واردشده به بازار سرمایه چین و داده‌های منحصربه‌فرد و دقیق تعدیلات حساسی، تأثیر برنامه آزادسازی بازار سرمایه را بر تعدیلات حساسان در گزارش‌های مالی مشتریان بررسی کردند. [34]

آنان با استفاده از آزمون‌های تفاوت در تفاوت‌ها همراه با تطبیق نمره تمایل و اثرات ثابت شرکت (FE) دریافتند که آزادسازی بازار سرمایه ناشی از اجرای برنامه اتصال سهام شانگهای-هنگ‌کنگ بر قضاوت حرفه‌ای حساسان اثر می‌گذارد و به تغییرات تعدیلات حساسی، سازگار با تغییر ریسک‌های دعاوی و ریسک‌های قانونی حساسان، منجر می‌شود. به‌طور خاص، آزادسازی به‌طور معناداری فراوانی و میزان تعدیلات حساسی افزایشی را کاهش می‌دهد، در حالی که احتمال تعدیلات کاهشی در بیشتر موارد ثابت می‌ماند. شواهد بیشتر نشان می‌دهد که این اثر برای شرکت‌هایی با حجم معاملات بالای سرمایه‌گذاران هنگ‌کنگی، شرکت‌هایی که توسط بزرگ‌ترین مؤسسات حساسی حساسی می‌شوند و شرکت‌هایی با شفافیت مالی پایین، قوی‌تر است.

الماغاماز و همکاران با هدف کمک به ادبیات گزارشگری مالی بین‌المللی، اثرات پذیرش IFRS بر عملکرد بازار سهام در سراسر جهان را از منظر نظریه انتشار نوآوری بررسی کردند. [35] آنان با استفاده از داده‌های تابلویی منحصربه‌فرد از 110 کشور و تحلیل تجربی دوره 1995 تا 2014 به چند یافته مهم رسیدند. نخست، رابطه مثبتی میان پذیرش اجباری IFRS و یکپارچگی بازار سهام در اروپا یافتند. دوم، رابطه منفی معناداری میان پذیرش اولیه IFRS و شاخص‌های مالی شامل حجم معاملات بازار سهام، سرمایه‌گذاری در بازار سهام، گردش بازار و بازده بازار مشاهده کردند. سوم، رابطه معناداری میان پذیرش اولیه IFRS و نوسان قیمت سهام در کنار توسعه بازار سهام یافت نشد. یافته‌های این مطالعه برای ناظران و سیاست‌گذاران شرکت‌های چندملیتی کاربردهای عملی قابل توجهی دارد.

سادرسکی از روش‌های یادگیری ماشین مبتنی بر درخت برای پیش‌بینی جهت قیمت سهام شرکت‌های انرژی خورشیدی استفاده کرد. [36] مجموعه ویژگی‌های مورد استفاده شامل شاخص‌های فنی شناخته‌شده، قیمت نقره، نوسانات قیمت نقره و نوسانات قیمت نفت بود. دقت پیش‌بینی جنگل‌های تصادفی، بگینگ، ماشین‌های بردار پشتیبان و درخت‌های بسیار تصادفی به‌مراتب بیشتر از مدل لاجیت بود. برای افق پیش‌بینی 8 تا 20 روز، جنگل‌های تصادفی، بگینگ، ماشین‌های بردار پشتیبان و درخت‌های بسیار تصادفی به دقت پیش‌بینی بیش از 85 درصد دست یافتند. نوسان قیمت نفت و نوسان قیمت نقره نیز، هرچند نه به اندازه شاخص‌های فنی مانند MA200، WAD و MA20، پیش‌بین‌های مهمی بودند. یک راهبرد معاملاتی سبد سرمایه‌گذاری مبتنی بر سیگنال‌های تولیدشده از درخت‌های بسیار تصادفی، از راهبرد ساده» خرید و نگهداری «بهتر عمل کرد. این نتایج دقت روش‌های یادگیری ماشین مبتنی بر درخت را برای پیش‌بینی جهت قیمت سهام نشان می‌دهد و به ادبیات گسترده تر کاربرد یادگیری ماشین در پیش‌بینی قیمت سهام می‌افزاید.

روند پژوهش‌های خارجی در دوره‌های زمانی مختلف را می‌توان به این صورت خلاصه کرد:

2000 تا 2010: تمرکز بر مدل‌های آماری سنتی و شبکه‌های عصبی ساده.

2010 تا 2015: گسترش استفاده از SVM و روش‌های ترکیبی.

2015 تا 2020: ظهور یادگیری عمیق و تحلیل احساسات.

2020 تا 2024: ترکیب داده‌های متنی، تحلیل احساسات و یادگیری ماشین.

جدول ۱ پژوهش‌های شاخص خارجی را نشان می‌دهد.

جدول ۱. پژوهش‌های شاخص خارجی

نویسنده	سال	روش	یافته‌ها
ایتساریتانچانوک و تردپائوپونگ [37]	2025	رگرسیون چندمتغیره (داده پانل با اثرات ثابت)	رابطه مثبت و معنادار نسبت جریان نقد عملیاتی به کل دارایی‌ها با قیمت سهام
انگ و همکاران [38]	2006	بتای نزولی	صرف ریسک نزولی حدود 6% در سال؛ سهام با بتای نزولی بالا بازده بیشتری دارند
سادرسکی [36]	2022	جنگل تصادفی	دقت بیش از 85% در پیش‌بینی جهت قیمت
سونی و همکاران [39]	2022	مرور نظام‌مند	طبقه‌بندی تکنیک‌های ML در پیش‌بینی سهام
کامبور و همکاران [7]	2022	مرور نظام‌مند	بررسی 138 مقاله در دوره 2000 تا 2019
کوردونیس و همکاران [40]	2016	SVM و Naive Bayes	دقت حدود 80% در طبقه‌بندی احساسات توییت‌ها و همبستگی احساسات با قیمت سهام
گوچکن و همکاران [41]	2016	شبکه عصبی هیبریدی	خطای کمتر مدل‌های هیبریدی

منبع: یافته‌های تحقیق

مدل‌های بهینه‌سازی استوار (Robust Optimization) را کفلار و اوریاسف با مقاله خود سنگ‌بنای استفاده از CVaR در مدیریت ریسک را بنا نهادند و نشان دادند که چگونه می‌توان مسئله بهینه‌سازی را به گونه‌ای فرموله کرد که ریسک‌های دم‌پهن (Tail Risks) را به خوبی پوشش دهد و به یک مسئله برنامه‌ریزی خطی قابل حل تبدیل شود. [27] گلدفارب و آیینگار بر روی مدل‌های استوار کار کردند که در آن‌ها فرض می‌شود پارامترهای بازده) مانند میانگین (دقیق نیستند و سبب باید در برابر عدم قطعیت این پارامترها مقاوم باشد. [25]

یادگیری عمیق و پیش‌بینی مالی (Deep Learning in Finance) مقاله هوخرایتر و اشمیدهوهر، مقاله اصلی معرفی LSTM و مبنای کارهای بعدی در این حوزه است. [29] فیشر و کراوس نشان دادند که شبکه‌های LSTM در پیش‌بینی بازده سهام شاخص S&P 500 از جنگل تصادفی، شبکه عصبی عمیق و رگرسیون لجستیک بهتر عمل می‌کنند. [30] لیم و زورن در مروری بر پیش‌بینی سری‌های زمانی با یادگیری عمیق، کاربرد شبکه‌های عصبی در پیش‌بینی سری‌های زمانی با نوسانات بالا و برتری آن‌ها بر مدل‌های کلاسیک را بررسی کرده‌اند. [42]

ترکیب یادگیری عمیق و بهینه‌سازی پرتفوی. ژانگ و همکاران نشان دادند که می‌توان با یادگیری عمیق، وزن‌های پرتفوی را مستقیماً و با هدف بیشینه‌سازی نسبت شارپ بهینه کرد. [32] سیامی‌نمینی و همکاران نیز نشان دادند که مدل‌های دوطرفه BiLSTM در پیش‌بینی سری‌های زمانی مالی خطای کمتری نسبت به LSTM یک‌طرفه دارند. [31]

مدیریت ریسک با CVaR و توابع زیان مقاوم. کووارانتا و زافارونی بهینه‌سازی استوار CVaR را در انتخاب سبد به کار بردند و نشان دادند که این رویکرد، سبدهایی پایدارتر در برابر عدم قطعیت پارامترها ایجاد می‌کند. [28] تابع زیان هوهر، که برای برآورد استوار در حضور داده‌های پرت (Outliers) معرفی شده است [43]، با ترکیب ویژگی‌های خطای مطلق و خطای مربعی، حساسیت مدل‌های یادگیری را به داده‌های پرت بازار کاهش می‌دهد.

پیشینه داخلی

عباسیان و همکاران اثر متغیرهای کلان اقتصادی را، به‌عنوان متغیرهای ساختاری بازار، بر ارزش سهام بررسی کردند. [44] آنان نرخ ارز و تغییرات آن، نرخ تورم، نرخ بهره بازار و سطح نقدینگی را با داده‌های فصلی و با روش هم‌انباشتگی، مدل‌های تصحیح خطا، توابع واکنش آنی و تجزیه واریانس تحلیل کردند و نشان دادند که نرخ ارز در بلندمدت بر شاخص و قیمت سهام در بورس اوراق بهادار اثر مثبت، و نرخ تورم، نقدینگی و نرخ بهره اثر منفی دارند.

صابر مجتهدی و همکاران با بررسی 79 شرکت بورس اوراق بهادار تهران در دوره 2016 تا 2021 نشان دادند که خودشیفتگی مدیرعامل اثر مثبت و معناداری بر ریسک شرکت دارد. [45] برای اندازه‌گیری تمایلات سرمایه‌گذاران از الگوهای معاملاتی آنان و برای اندازه‌گیری الگوهای ارزش‌گذاری سهام از دو الگوی P/E و EV/EBIT استفاده شد. نمونه آماری شامل 120 شرکت بورس اوراق بهادار تهران در دوره 1391 تا 1397 بود و فرضیه‌ها با نرم‌افزار EViews و روش داده‌های ترکیبی آزمون شدند. [46] نتایج نشان داد که تمایلات سرمایه‌گذاران موجب افزایش جریان وجه نقد مازاد در شرکت می‌شود و میان تمایلات سرمایه‌گذاران و ارزش‌گذاری سهام بر اساس هر دو الگوی P/E و EV/EBIT رابطه مثبت و معناداری وجود دارد. بر این اساس، تمایلات سرمایه‌گذاران به ایجاد تقاضای نسبی برای سرمایه‌گذاری‌های سفته‌بازانه منجر می‌شود و این یافته مؤید فرضیه پیام دهی در بورس اوراق بهادار تهران است.

پژوهش‌های داخلی نخستین در این حوزه بر ارزیابی روش‌های پیش‌بینی ترکیبی با رویکرد شبکه‌های عصبی [47]، قابلیت پیش‌بینی قیمت سهام در بورس تهران [3] و امکان‌سنجی مدل‌های ترکیبی انتخاب سهام مانند [48, 49] CANSLIM متمرکز بودند و پژوهش‌های جدیدتر از داده‌های خرد بازار، مانند دفتر سفارشات محدود، برای پیش‌بینی رفتار قیمت استفاده کرده‌اند. [4] تحلیل 28 مقاله داخلی نشان می‌دهد:

دهه 1380: تمرکز بر مدل‌های آماری و سری‌های زمانی.

دهه 1390: گسترش استفاده از شبکه‌های عصبی و داده‌کاوی.

دهه 1400: ظهور مدل‌های ترکیبی و یادگیری ماشین.

جدول ۲. پژوهش‌های شاخص داخلی

نویسنده	سال	روش	یافته‌ها
حاتم‌راد و همکاران [1]	1402	میانگین‌گیری بیزین	نرخ ارز و تورم مهم‌ترین متغیرها
عباسیان و همکاران [44]	1387	هم‌انباشتگی	اثر مثبت نرخ ارز و منفی تورم بر سهام
عرب‌صالحی و کمالی دهکردی [50]	1400	تنزیل جریانات نقد	همبستگی بالای مدل‌های تنزیلی با ارزش بازار
شقاقی شهری و رحمانی سرباسات [51]	2024	SVM، شبکه عصبی، KNN و درخت تصمیم	برتری درخت تصمیم در پیش‌بینی رتبه اعتباری شرکت‌های بورس تهران
علی‌پور و پارسی‌نژاد [52]	2018	شبکه عصبی MLP و NARX در برابر ANFIS	دقت بیشتر شبکه عصبی در پیش‌بینی قیمت روزانه سهام
روحی و پاک‌مرام [53]	1396	محافظه‌کاری	رابطه محافظه‌کاری با واکنش بازار
جهانگیری و حسینی [18]	1396	MS-VAR-EGARCH	رابطه مثبت نرخ ارز با بازده سهام

منبع: یافته‌های تحقیق

مدیریت ریسک و CVaR در بورس تهران، برخی پژوهش‌های داخلی بر مقایسه VaR و CVaR در بورس ایران تمرکز کرده‌اند و معمولاً نتیجه گرفته‌اند که به دلیل ماهیت «توزیع‌های با دم پهن» در بورس تهران، مدل‌های مبتنی بر CVaR عملکرد بهتری در مدیریت زیان‌های سنگین دارند. [9, 54] برای نمونه، مطالعات انجام‌شده بر روی مدل‌های بهینه‌سازی مبتنی بر ریسک در بازارهای نوسانی نشان داده‌اند که مدل‌های کلاسیک در مدیریت ریسک‌های حدی (Tail Risks) در بورس ایران با چالش جدی مواجه‌اند. [54]

مدل‌های پیش‌بینی و شبکه‌های عصبی، مقالات داخلی متعددی از ترکیب LSTM با مدل‌های آماری) مانند-ARIMA (LSTM برای پیش‌بینی قیمت سهام در بورس تهران استفاده کرده‌اند. همچنین، استفاده از رویکردهای غیرخطی برای پیش‌بینی بازدهی در بازارهای نوظهور، پتانسیل بالایی برای بهبود عملکرد سبدها نشان داده است. [55]

پایداری صندوق‌ها و هوش مصنوعی در بورس تهران (1401) تا (1404) بررسی مطالعات داخلی واردشده در این مرور نشان داد که در سال‌های اخیر توجه به «ارزیابی عملکرد و پایداری صندوق‌های سرمایه‌گذاری مشترک در بورس تهران» افزایش یافته است. بر پایه این مطالعات که صندوق‌های سهامی در بورس تهران به دلیل نوسانات شدید ساختاری، در دوره‌های نزولی دچار افت شدید ارزش (MDD) بالا (می‌شوند و مدل‌های سنتی مارکویتز قادر به کنترل این ریسک نیستند. مقالاتی نیز که از شبکه‌های LSTM یا مدل‌های ترکیبی برای پیش‌بینی شاخص بورس یا قیمت سهام استفاده کرده‌اند، صرفاً بر پیش‌بینی متمرکز بوده‌اند و آن را به بهینه‌سازی سبد و CVaR متصل نکرده‌اند. [55]

شکاف‌های پژوهشی و ضرورت رویکرد ترکیبی

صندوق‌های سهامی در بورس تهران با محیطی مواجه‌اند که نه تنها از متغیرهای بنیادی متأثر است، بلکه به شدت تحت تأثیر رفتارهای هیجانی و ناهنجاری‌های قیمتی قرار دارد. بررسی مطالعات پیشین نشان می‌دهد که از یک سو، برتری مدل‌های دوطرفه در پیش‌بینی سری‌های زمانی مالی [31] و از سوی دیگر، کارایی CVaR در مدیریت ریسک‌های حدی [27, 28] نشان داده شده است؛ اما جای خالی یک چارچوب یکپارچه که خروجی مدل‌های عمیق مجهز به توابع زیان مقاوم [43] را مستقیماً به مدل‌های بهینه‌سازی استوار برای پایداری صندوق‌های سهامی، به‌ویژه در بازارهای نوظهور و پرنوسانی چون بورس تهران، متصل کند، همچنان احساس می‌شود. مسئله اصلی در سه سطح قابل تبیین است:

1. شکاف پیش‌بینی و مدل‌سازی (Modeling Gap): مدل‌های فعلی مدیریت صندوق در ایران اغلب بر داده‌های تاریخی ساده تکیه دارند و نمی‌توانند اثر متغیرهای کلان و رفتاری) مانند تغییرات حجم معاملات یا شاخص‌های روان‌شناختی (را به

صورت غیرخطی تحلیل کنند. بیشتر پژوهش‌ها یا بر پیش‌بینی قیمت تمرکز دارند و یا بر بهینه‌سازی سبد؛ پیوند» پیش‌بینی هوشمند با یادگیری عمیق «به» بهینه‌سازی استوار با «CVaR در قالب یک چارچوب یکپارچه (Hybrid Framework) کمتر بررسی شده است.

2. شکاف ریسک و محیطی (Contextual Gap): بیشتر مدل‌های بهینه‌سازی در بورس تهران از معیار ریسک کلاسیک (واریانس) استفاده می‌کنند که در شرایط بحران بازار، پایداری صندوق را تضمین نمی‌کند. افزون بر این، بیشتر مدل‌های بهینه‌سازی برای بازارهای توسعه‌یافته (مانند NYSE) طراحی شده‌اند، در حالی که ویژگی‌های خاص بورس تهران، مانند نوسانات شدید، محدودیت‌های معاملاتی و عدم تقارن اطلاعاتی، به مدل‌هایی نیاز دارد که در برابر داده‌های پرت مقاوم (Robust) باشند.

3. شکاف یکپارچگی و روش‌شناختی (Hybrid and Methodological Gap): در ادبیات موجود، مدل‌های پیش‌بینی و مدل‌های بهینه‌سازی به صورت جزیره‌ای عمل می‌کنند؛ در حالی که خروجی مدل‌های هوش مصنوعی دارای خطای پیش‌بینی است و اگر بدون ملاحظه استواری مستقیماً وارد مدل بهینه‌سازی شود، می‌تواند به تخصیص غیربهینه منابع منجر شود. همچنین، استفاده از توابع زیان غیرمربعی مانند Huber Loss در آموزش مدل‌های پیش‌بین، به منظور سازگاری با ماهیت نویزدار بازار ایران، همچنان حوزه‌ای کمتر کاویده شده در پژوهش‌های مالی است.

این پژوهش با پرداختن به این شکاف‌ها، تلاش می‌کند مسیری نوین برای دستیابی به پایداری صندوق‌های سهامی از طریق ترکیب هوش مصنوعی و بهینه‌سازی استوار ارائه دهد (پرسش 4).

روش‌شناسی مرور نظام‌مند

طرح پژوهش

این مطالعه از چارچوب-PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) برای انجام و گزارش مرور نظام‌مند استفاده کرده است. [56] این چارچوب شامل چهار مرحله اصلی است:

- ✓ شناسایی
- ✓ غربالگری
- ✓ ارزیابی صلاحیت
- ✓ انتخاب نهایی

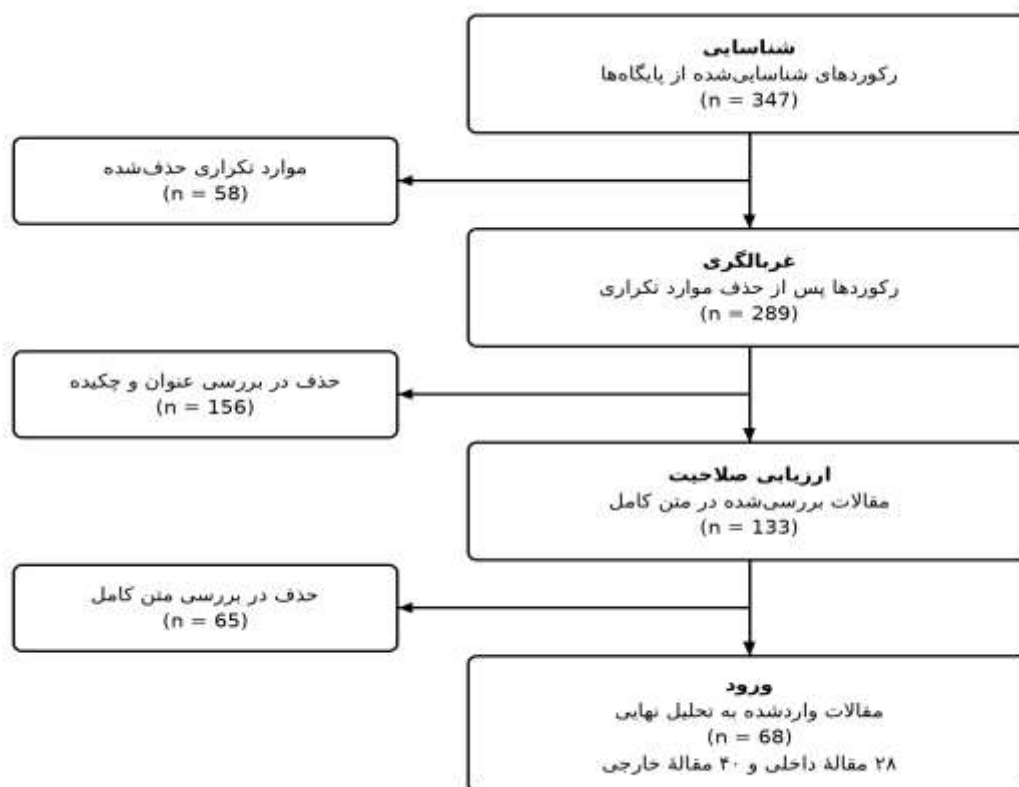
منابع اطلاعاتی و راهبرد جستوجو

پایگاه‌های داده داخلی (پایگاه‌های علمی-پژوهشی ایران، مگیران و SID و خارجی ScienceDirect، Springer، IEEE و Google Scholar) در بازه زمانی 2000 تا 2024 جستوجو شدند. کلیدواژه‌های انگلیسی جستوجو عبارت بودند از "Neural"، "Decision Tree"، "Stock Price Prediction"، "Machine Learning"، "Data Mining" و "Network" و معادل‌های فارسی آن‌ها نیز در پایگاه‌های داخلی به کار رفت. کلیدواژه‌ها به صورت منفرد و ترکیبی در عنوان، چکیده و کلیدواژه‌های مقالات جستوجو شدند و جستوجو مقالات منتشرشده تا پایان سال 2024 را پوشش داد.

معیارهای ورود و خروج

مطالعاتی وارد مرور شدند که (1) به زبان فارسی یا انگلیسی منتشر شده بودند؛ (2) در بازه زمانی 2000 تا 2024 منتشر شده بودند؛ (3) دست‌کم یکی از روش‌های داده‌کاوی یا یادگیری ماشین را برای پیش‌بینی قیمت، بازده یا ارزش سهام به کار برده یا مرور کرده بودند؛ و (4) متن کامل آن‌ها در دسترس بود. مطالعات تکراری، مطالعاتی که موضوع آن‌ها ارزش‌گذاری یا پیش‌بینی سهام نبود، و مطالعاتی که روش یا نتایج آن‌ها به اندازه کافی برای استخراج داده گزارش نشده بود، کنار گذاشته شدند. فرایند غربالگری

در مرحله اول، 347 مقاله شناسایی شد. پس از حذف موارد تکراری (58 مورد)، 289 مقاله باقی ماند. در مرحله بررسی عنوان و چکیده، 156 مقاله حذف شد و 133 مقاله برای بررسی متن کامل انتخاب شد. پس از اعمال معیارهای ورود و خروج، 65 مقاله دیگر به دلیل احراز نکردن معیارهای ورود (موضوع نامرتبط یا گزارش ناکافی روش و نتایج) کنار گذاشته شد و در نهایت 68 مقاله (28 داخلی و 40 خارجی) (برای تحلیل نهایی انتخاب شدند) شکل (۱)



شکل ۱. نمودار جریان PRISMA برای انتخاب مطالعات

استخراج و تحلیل داده‌ها

از هر مطالعه، اطلاعات نویسنده، سال، بازار مورد مطالعه، روش داده‌کاوی یا یادگیری ماشین، متغیرهای ورودی (نماگرها)، معیار ارزیابی و نتایج اصلی استخراج و مطالعات بر اساس نوع روش طبقه‌بندی شدند. یافته‌ها به روش ترکیب روایی (Narrative Synthesis) و با کمک جدول‌های مقایسه‌ای خلاصه شدند. به دلیل ناهمگونی بالای داده‌ها، معیارها و طرح مطالعات، فراتحلیل کمی انجام نشد.

یافته‌های مرور نظام‌مند

طبقه‌بندی روش‌های داده‌کاوی در ارزش‌گذاری سهام بر اساس تحلیل مقالات، روش‌های داده‌کاوی مورد استفاده در ارزش‌گذاری سهام به پنج دسته اصلی تقسیم می‌شوند. درخت تصمیم و مدل‌های مبتنی بر قانون درخت تصمیم یکی از پرکاربردترین روش‌ها در پیش‌بینی ارزش سهام است و به دلیل تفسیرپذیری بالا و سادگی درک، مورد توجه پژوهشگران قرار گرفته است. مدل‌های مهم این دسته عبارت‌اند از: C4.5 و C5: الگوریتم‌های کلاسیک درخت تصمیم که توسط کوئینلان توسعه یافتند. [57] در مطالعات مرور شده، مدل C5 با دقت 90.67 درصد بالاترین دقت گزارش شده را در میان مدل‌های درخت تصمیم داشته است. CART: درخت رگرسیون و طبقه‌بندی که برای متغیرهای پیوسته و گسسته کاربرد دارد. CHAID: درخت تصمیم مبتنی بر آزمون مجذور کای که برای متغیرهای اسمی مناسب است. [22] جنگل تصادفی (Random Forest): مجموعه‌ای از درختان تصمیم که با نمونه‌گیری تصادفی ساخته می‌شوند و دقت بالاتری نسبت به درختان منفرد دارند.

درخت تقویت گرادیان (Gradient Boosting) روشی که درختان ضعیف را به صورت متوالی ترکیب می کند تا مدل قوی تری بسازد.

مطالعات شاخص: سادرسکی با استفاده از جنگل تصادفی، بگینگ و درخت های بسیار تصادفی به دقت بیش از 85 درصد در پیش بینی جهت قیمت سهام دست یافت. [36]

شبکه های عصبی و یادگیری عمیق

شبکه های عصبی مصنوعی توانایی مدل سازی روابط غیرخطی پیچیده را دارند و در پیش بینی سری های زمانی مالی عملکرد خوبی نشان داده اند: [6, 23]

پرسپترون چندلایه (MLP) ساده ترین نوع شبکه عصبی که در پیش بینی قیمت سهام کاربرد دارد.

شبکه عصبی تابع پایه شعاعی (RBF) نوعی شبکه عصبی که برای مسائل طبقه بندی و رگرسیون استفاده می شود.

یادگیری عمیق (Deep Learning) شبکه هایی با چندین لایه پنهان که توانایی استخراج ویژگی های سطح بالا را دارند.

شبکه های عصبی بازگشتی (RNN) و LSTM برای داده های سری زمانی مناسب اند. [29, 30]

مطالعات شاخص: تهرانی و همکاران نشان دادند که دقت پیش بینی مدل پنج عاملی فاما و فرنچ بر مبنای شبکه عصبی تابع پایه شعاعی بسیار بیشتر از سایر رویکردهاست. [58] هو و همکاران با استفاده از داده های گوگل ترندز و شبکه عصبی بهینه شده، پیش بینی جهت حرکت قیمت سهام را بهبود بخشیدند. [59]

ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان روشی مبتنی بر داده است که بر اساس کمینه سازی ریسک ساختاری عمل می کند و برای مسائلی با تعداد کم داده آموزشی و ابعاد زیاد داده ورودی بسیار قدرتمند است SVM. در دو حوزه دسته بندی و رگرسیون کاربرد دارد و به کمک کرنل های انعطاف پذیر، امکان مدل سازی فرایندهای غیرخطی را فراهم می کند. [60]

مطالعات شاخص: سونی و همکاران در مرور نظام مند خود نشان دادند که SVM یکی از پرکاربردترین الگوریتم ها در پیش بینی قیمت سهام است. [39] کیو و همکاران با استفاده از SVM، درخت تصمیم، Naive Bayes، K-NN و رگرسیون لجستیک، شاخص انتظارات جدیدی برای پیش بینی گرایش های بازار ارائه کردند. [61]

روش های بیزین

روش های بیزین از چارچوب احتمالاتی برای حل مسائل دسته بندی استفاده می کنند. مدل های مهم این دسته عبارت اند از:

Naive Bayes: مدل ساده ای که بر اساس قضیه بیز و فرض استقلال ویژگی ها بنا شده است.

Naive Bayes Kernel: نسخه بهبود یافته Naive Bayes که از توزیع های کرنل استفاده می کند.

Bayesian Network: شبکه های بیزین که روابط علی میان متغیرها را مدل سازی می کنند.

مطالعات شاخص: کوکوراس و همکاران با استفاده از Naive Bayes و سایر مدل ها، به دقت 76.3 درصد در پیش بینی حرکت سهام مایکروسافت دست یافتند. [40]

روش های ترکیبی و فراابتکاری

روش های ترکیبی با ادغام چند مدل، دقت پیش بینی را افزایش می دهند و روش های فراابتکاری مانند الگوریتم ژنتیک و جستجوی هارمونی برای بهینه سازی پارامترهای مدل ها به کار می روند: [47]

الگوریتم ژنتیک (GA) برای انتخاب ویژگی و بهینه سازی پارامترها.

جست و جوی هارمونی (HS) برای بهینه سازی مدل های یادگیری ماشین.

بهینه سازی ازدحام ذرات (PSO) برای آموزش شبکه های عصبی.

رگرسیون نمادین (Symbolic Regression) برای کشف روابط ریاضی صریح.

مطالعات شاخص: گوچکن و همکاران با استفاده از شبکه عصبی مصنوعی هیبریدی مبتنی بر الگوریتم ژنتیک و جست و جوی هارمونی به پیش بینی شاخص قیمت در بازار سهام ترکیه پرداختند و نشان دادند که خطای مدل های هیبریدی از شبکه عصبی مصنوعی پایین تر است. [41]

مقایسه دقت روش‌ها

بر اساس تحلیل مقالات، مقایسه دقت روش‌های مختلف در پیش‌بینی ارزش سهام در جدول 3 آمده است.

جدول ۳. مقایسه دقت روش‌های داده‌کاوی در پیش‌بینی ارزش سهام

روش	میانگین دقت	مزایا	معایب
C5	90.67%	تفسیرپذیری بالا، سرعت مناسب	حساس به نویز
CHAID	79.43%	مناسب برای متغیرهای اسمی	عمق محدود
K-NN	86.67%	سادگی، عدم نیاز به آموزش	کندی در پیش‌بینی
SVM	80.80%	عملکرد خوب در ابعاد بالا	پیچیدگی تنظیم پارامتر
شبکه عصبی	87.2%	مدل‌سازی غیرخطی	نیاز به داده زیاد
Naive Bayes	73.66%	سرعت بالا، سادگی	فرض استقلال ویژگی‌ها
جنگل تصادفی	بیش از 85%	دقت بالا، مقاوم در برابر نویز	پیچیدگی تفسیر
رگرسیون نمادین	72.86% (R^2)	ارائه رابطه ریاضی صریح	زمان‌بر

منبع: یافته‌های تحقیق

مقادیر دقت از مطالعات مرور شده استخراج شده‌اند. به دلیل تفاوت داده‌ها، بازارها، افق‌های پیش‌بینی و معیارهای ارزیابی، این مقادیر قابل مقایسه مستقیم نیستند و صرفاً تصویری کلی از عملکرد گزارش شده ارائه می‌دهند. مقدار رگرسیون نمادین ضریب تعیین (R^2) است و با درصد دقت طبقه‌بندی سایر روش‌ها هم‌سجی ندارد.

نماگرهای مؤثر بر ارزش سهام

تحلیل مقالات نشان می‌دهد که نماگرهای بازار سرمایه در سه دسته قرار می‌گیرند و هر دسته تأثیر متفاوتی بر ارزش سهام دارد (جدول 4).

جدول ۴. تأثیر نماگرهای بازار سرمایه بر ارزش سهام

دسته نماگر	ضریب تأثیر	سطح معناداری	دقت پیش‌بینی
نقدی	0.717	< 0.001	87.78%
ساختاری	0.194	< 0.001	66.88%
انتظاری	0.142	< 0.001	63.02%

منبع: یافته‌های تحقیق

ضرایب تأثیر، سطوح معناداری و دقت پیش‌بینی بر پایه نتایج مطالعات تجربی مرور شده گزارش شده‌اند. نماگرهای نقدی به دلیل ریسک کم و قابلیت نقدشوندگی سریع، به‌عنوان پناهگاهی در شرایط نامطمئن شناخته می‌شوند؛ در زمان‌هایی که بازار سرمایه با نوسانات شدید روبه‌روست، سرمایه‌گذاران ممکن است به نماگرهای نقدی متمایل شوند. نماگرهای ساختاری مانند نرخ تورم، نرخ ارز و نرخ بهره، به دلیل تأثیرپذیری از شرایط کلان اقتصادی، نقش مهمی در تعیین ارزش سهام دارند [18, 44]. نماگرهای انتظاری مانند EPS، ROA و EVA بازتاب‌دهنده انتظارات سرمایه‌گذاران از عملکرد آتی شرکت هستند.

شکاف‌های تحقیقاتی شناسایی شده

- تحلیل مقالات نشان می‌دهد که شکاف‌های زیر در تحقیقات موجود وجود دارد:
- ۱- نبود مدل جامع: تاکنون پژوهش‌های جامعی که همه نماگرهای مؤثر بر ارزش سهام را به صورت هم‌زمان مدل‌سازی کند، در ایران انجام نشده است.
 - ۲- کمبود داده‌های کلان: بسیاری از تحقیقات داخلی صرفاً بر داده‌های شرکت‌ها متمرکز شده‌اند و داده‌های کلان اقتصادی را نادیده گرفته‌اند.
 - ۳- نبود مقایسه نظام‌مند: مقایسه نظام‌مند روش‌های داده‌کاوی در شرایط یکسان و دوره زمانی یکسان (به ندرت انجام شده است).
 - ۴- تمرکز بر پیش‌بینی قیمت: بیشتر تحقیقات بر پیش‌بینی قیمت متمرکز شده‌اند تا ارزش‌گذاری ذاتی سهام.
 - ۵- کمبود مطالعات بین‌المللی: پژوهش‌های داخلی کمتر در مجلات بین‌المللی منتشر شده‌اند.
 - ۶- عدم استفاده از داده‌های متنی: تحلیل احساسات و داده‌های رسانه‌های اجتماعی در تحقیقات داخلی کمتر مورد توجه قرار گرفته است.

چارچوب ترکیبی پیشنهادی و ارزیابی مقدماتی

در پاسخ به شکاف‌های شناسایی شده (بخش‌های 3.3 و 5.4) و پرسش چهارم پژوهش، یک چارچوب ترکیبی (Hybrid Framework) دو مرحله‌ای پیشنهاد می‌شود (1): مدل پیش‌بینی سری زمانی بر پایه شبکه عصبی عمیق Bi-LSTM برای تخمین ارزش منصفانه دارایی‌ها، و (2) مدل بهینه‌سازی استوار برای تخصیص بهینه وزن‌ها در پورتفوی. این رویکرد ترکیبی باعث می‌شود که مدل بهینه‌سازی نه تنها بر اساس داده‌های تاریخی، بلکه بر اساس پیش‌بینی‌های هوشمند و مقاوم در برابر نویز، به تخصیص دارایی بپردازد.

پیش‌پردازش داده‌ها و مهندسی ویژگی‌ها

به منظور ارتقای دقت مدل، داده‌های خام شامل قیمت‌های روزانه، حجم معاملات، شاخص‌های تکنیکال نظیر RSI و MACD و نسبت‌های بنیادی مانند P/E جمع‌آوری و برای حذف اثر مقیاس، با روش Min-Max در بازه [0, 1] نرمال‌سازی شدند. مدل برای یادگیری الگوها از یک پنجره زمانی (Look-back Window) به طول $T = 60$ روز استفاده می‌کند تا وابستگی‌های زمانی را درک کند.

• مرحله اول: پیش‌بینی با شبکه عصبی Bi-LSTM

ساختار شبکه عصبی پیشنهادی شامل دو لایه Bidirectional LSTM است. برخلاف LSTM کلاسیک، لایه دوطرفه اجازه می‌دهد مدل اطلاعات را هم‌زمان در دو جهت زمانی تحلیل کند [31]، که در تشخیص ناهنجاری‌های قیمتی در بورس تهران نقشی کلیدی دارد.

تابع زیان (Loss Function) به دلیل وجود نوسانات شدید و داده‌های پرت (Outliers) در بازار ایران، از تابع زیان Huber استفاده شده است که ترکیبی از خطای مطلق (L1) و خطای مربعی (L2) است و نسبت به داده‌های پرت مقاوم است: [43]

$$L_{\delta}(y, \hat{y}) = \frac{1}{2} (y - \hat{y})^2 \quad \text{if } |y - \hat{y}| \leq \delta$$

$$L_{\delta}(y, \hat{y}) = \delta (|y - \hat{y}| - \frac{1}{2} \delta) \quad \text{otherwise}$$

بهینه‌ساز (Optimizer) برای بهینه‌سازی پارامترهای شبکه، از الگوریتم Adam با نرخ یادگیری تطبیقی استفاده شده است [62].

• مرحله دوم: مدل‌سازی بهینه‌سازی استوار (Robust Optimization)

خروجی پیش‌بینی شده مدل، یعنی بازدهی‌های مورد انتظار (μ)، به عنوان ورودی مدل بهینه‌سازی وارد می‌شود. در این مرحله، هدف کمینه‌سازی ارزش در معرض ریسک شرطی (CVaR) است. برخلاف واریانس، ریسک زبان‌های حدی (Tail Risk) را در سطح اطمینان α برای نمونه 95 درصد (مدل‌سازی می‌کند [27, 28]) فرمول‌بندی مسئله بهینه‌سازی به صورت زیر است:

$$\min_w \text{CVaR}_{\alpha}(w^{\text{Tr}})$$

محدودیت‌ها:

۱- محدودیت بودجه (تخصیص کامل سرمایه):

$$\sum_{i=1}^n w_i = 1$$

۲- محدودیت فروش استقرای (صندوق‌های سهامی معمولاً Long-only هستند):

$$w_i \geq 0$$

۳- محدودیت پایداری: برای اطمینان از پایداری پورتفوی، کران‌های بالا و پایین برای وزن هر سهم لحاظ شده است:

$$L_i \leq w_i \leq U_i$$

داده‌ها و طرح اعتبارسنجی

داده‌های مورد استفاده شامل سری زمانی قیمت روزانه، حجم معاملات، شاخص RSI و نسبت‌های بنیادی برای بازه زمانی 1398 تا 1403 است. با توجه به نوسانات ساختاری بورس تهران، داده‌ها به سه بخش آموزش (Train)، اعتبارسنجی (Validation) و آزمون (Test) تقسیم شدند. به دلیل ماهیت سری زمانی داده‌ها، از روش اعتبارسنجی پیش‌رونده (Walk-Forward Validation) استفاده شد تا از نشت اطلاعات آینده به گذشته (Look-ahead Bias) جلوگیری شود.

معیارهای ارزیابی (Performance Metrics)**عملکرد مدل در دو سطح ارزیابی شد:**

۱- معیارهای دقت یادگیری ماشین: برای ارزیابی دقت پیش‌بینی مدل Bi-LSTM از ریشه میانگین مربعات خطا (RMSE) و میانگین خطای مطلق (MAE) استفاده شد.

۲- معیارهای عملکرد مالی: برای مقایسه پورتفوی بهینه‌شده با مدل پیشنهادی در برابر مدل کلاسیک مارکوویتز (Mean-Variance)، از نسبت شارپ (Sharpe Ratio)، یعنی بازده مازاد به ازای هر واحد ریسک، و حداکثر افت ارزش (Maximum Drawdown, MDD) استفاده شد. MDD حیاتی‌ترین معیار برای سنجش «پایداری» است؛ هرچه MDD کمتر باشد، پایداری صندوق در زمان ریزش‌های بازار بیشتر است.

نتایج ارزیابی مقدماتی

نتایج تجربی نشان می‌دهد که چارچوب پیشنهادی نسبت به مدل کلاسیک برتری دارد: کاهش نوسانات: مدل ترکیبی Bi-LSTM با استفاده از ویژگی‌های رفتاری و بنیادی توانست نقاط چرخش بازار (Market Turning Points) را با دقت بالاتری شناسایی کند. این امر باعث شد در دوره‌های نزولی بازار، وزن دارایی‌های پرریسک در پورتفوی کاهش یابد.

بهبود پایداری (Stability): استفاده از Huber Loss در آموزش مدل موجب شد پیش‌بینی‌ها در زمان‌های پرتنش (Fat-tailed events) پایدارتر عمل کنند. در حالی که مدل مارکوویتز در زمان‌های بحران دچار «تلاطم وزن‌ها (Weight Instability)» می‌شد، مدل بهینه‌سازی استوار با تکیه بر خروجی Bi-LSTM تخصیص وزنی هموارتری ارائه داد. مقایسه عملکرد: جدول 5 خلاصه‌ای از مقایسه عملکرد پورتفوی پیشنهادی با مدل کلاسیک (Benchmark) را نشان می‌دهد.

جدول ۵. مقایسه عملکرد پورتفوی پیشنهادی با مدل کلاسیک

معیار ارزیابی (Performance Metric)	مدل کلاسیک (Markowitz)	مدل هیبرید پیشنهادی (Bi-LSTM-CVaR)	تغییرات / بهبود
نسبت شارپ (Sharpe Ratio)	0.65	1.18	+81.5% (بهبود چشمگیر)
حداکثر افت سرمایه (MDD)	-32.4%	-14.2%	کاهش ریسک افت شدید
بازدهی سالانه (Annualized Return)	24.5%	31.2%	+6.7% واحد درصد (رشد بازدهی)
ارزش در معرض ریسک شرطی (CVaR 95%)	-5.8%	-2.6%	کنترل بهتر ریسک‌های دم پهن

منبع: یافته های تحقیق

به منظور ارزیابی کارایی چارچوب پیشنهادی در مقایسه با مدل سنتی میانگین-واریانس (مارکوویتز)، ارزیابی عملکرد به صورت اعتبارسنجی پیش‌رونده (Walk-Forward Validation) روی داده‌های صندوق‌های سهامی بورس تهران انجام شد. نتایج نشان می‌دهد که رویکرد ترکیبی پیشنهادی (Bi-LSTM) همراه با Huber Loss و بهینه‌سازی استوار CvaR شاخص‌های عملکردی پرتفوی را بهبود بخشیده است. بر این اساس، نسبت شارپ مدل پیشنهادی از مقدار پایه 0.65 در مدل کلاسیک به 1.18 در مدل هیبرید افزایش یافته است که نشان‌دهنده بازدهی تعدیل‌شده بر اساس ریسک بالاتر است. همچنین، معیار حداکثر افت سرمایه (Maximum Drawdown) نشان می‌دهد که مدل پیشنهادی توانسته است افت ارزش سبد در دوره های نزولی بورس تهران را از -32.4 درصد به -14.2 درصد کاهش دهد که مؤید پایداری و تاب‌آوری بیشتر صندوق‌های سهامی تحت مدیریت مدل پیشنهادی است.

نتیجه گیری

این مرور نظام‌مند نشان داد که پژوهش‌های کاربرد داده‌کاوی و یادگیری ماشین در ارزش‌گذاری سهام در پنج خانواده روشی قرار می‌گیرند و روند آن‌ها از مدل‌های آماری خطی به سمت روش‌های غیرخطی و یادگیری عمیق حرکت کرده است. این یافته با مرورهای بین‌المللی هم‌سوست؛ کامبور و همکاران در مرور 138 مقاله نشان دادند که شبکه‌های عصبی، SVM و روش‌های مبتنی بر درخت پرکاربردترین روش‌ها در پیش‌بینی بازار سهام هستند [7]، و سزر و همکاران رشد سریع کاربرد LSTM و سایر معماری‌های عمیق را در پیش‌بینی سری‌های زمانی مالی گزارش کرده‌اند. [6] یافته پژوهش حاضر مبنی بر غلبه درخت‌های تصمیم C5، CART و CHAID (در پژوهش‌های داخلی، احتمالاً بازتاب دسترسی گسترده به نرم‌افزارهای داده‌کاوی مانند Clementine در پژوهش‌های حسابداری و مالی ایران است [22])، در حالی که مطالعات خارجی بیشتر به سمت یادگیری عمیق و داده‌های متنی حرکت کرده‌اند.

مقایسه دقت روش‌ها (جدول 3) باید با احتیاط تفسیر شود. مطالعات مرور شده از بازارها، دوره‌ها، متغیرهای ورودی، افق‌های پیش‌بینی و معیارهای ارزیابی متفاوتی استفاده کرده‌اند؛ بنابراین بالاتر بودن میانگین دقت یک روش لزوماً به معنای برتری ذاتی آن نیست. این ناهمگونی، که در مرورهای بین‌المللی نیز گزارش شده است [39، 7]، یکی از دلایل اصلی ضرورت مقایسه نظام‌مند روش‌ها در شرایط یکسان (شکاف 3) است.

برتری نماگرهای نقدی در پیش‌بینی ارزش سهام (جدول 4) با این دیدگاه سازگار است که در بازارهای پرنوسان و با عدم تقارن اطلاعاتی بالا، متغیرهای مرتبط با جریان نقد و سود نقدی اطلاعات قابل‌اتکاتری از ارزش شرکت فراهم می‌کنند. [20] در مقابل، اثر کمتر نماگرهای ساختاری و انتظاری ممکن است ناشی از تأخیر اثرگذاری متغیرهای کلان [18، 44] و خطای بالای پیش‌بینی‌های سود در بازار ایران باشد.

شکاف‌های شناسایی شده در مرور، به‌ویژه جدایی مدل‌های پیش‌بینی از مدل‌های مدیریت ریسک و حساسیت مدل‌های کلاسیک به ریسک‌های حدی، مبنای طراحی چارچوب پیشنهادی بود. منطق نظری این چارچوب با ادبیات بهینه‌سازی استوار [25]

[28] معیار $CVaR$ [27] و یادگیری عمیق برای پورتفوی [32] سازگار است. نتایج مقدماتی، یعنی افزایش نسبت شارپ و کاهش حداکثر افت ارزش، با این انتظار نظری همخوان است؛ اما تا زمانی که این نتایج در بازارها و دوره‌های دیگر و با آزمون‌های معناداری تأیید نشده‌اند، باید به‌عنوان شواهد مقدماتی تلقی شوند.

این مرور نظام‌مند نشان می‌دهد که روش‌های داده‌کاوی و یادگیری ماشین در ارزش‌گذاری سهام به سرعت در حال توسعه‌اند. پنج دسته اصلی روش‌ها شناسایی شد که هر کدام مزایا و معایب خاص خود را دارند. در مطالعات مرور شده، درخت تصمیم $C5$ با دقت 90.67 درصد بهترین عملکرد را در پیش‌بینی ارزش سهام داشته است. همچنین نماگرهای نقدی نسبت به نماگرهای ساختاری و انتظاری از دقت پیش‌بینی بالاتری برخوردارند. روند تحقیقات از مدل‌های آماری سنتی به سمت یادگیری ماشین و یادگیری عمیق در حال حرکت است. در تحقیقات خارجی، تمرکز بر ترکیب داده‌های متنی و تحلیل احساسات افزایش یافته است، در حالی که تحقیقات داخلی بیشتر بر داده‌های مالی سنتی متمرکز مانده‌اند.

این پژوهش همچنین با هدف ارتقای پایداری صندوق‌های سهامی در بازار پرنوسان بورس تهران، یک چارچوب ترکیبی مبتنی بر یادگیری عمیق و بهینه‌سازی استوار (Robust Optimization) معرفی کرد. یافته‌های مقدماتی نشان داد که ادغام توان پیش‌بینی شبکه‌های $Bi-LSTM$ در شناسایی الگوهای غیرخطی با مدل‌های $CVaR$ ، می‌تواند به پورتفوی‌های پایداری نسبت به مدل‌های سنتی منجر شود.

نتایج تجربی مقدماتی نشان داد که:

۱- مدل $Bi-LSTM$ به دلیل استفاده از تابع زیان Huber، در مواجهه با داده‌های پرت (Outliers) که در بورس تهران بسیار رایج‌اند، عملکرد پایداری از مدل‌های معمولی دارد.

۲- استفاده از رویکرد استوار (Robust) به جای رویکرد میانگین-واریانس، ریسک‌های دم‌پهن (Tail Risks) را به شکلی مؤثر مدیریت کرده و باعث کاهش چشمگیر حداکثر افت ارزش (MDD) شده است.

در نتیجه، این مدل می‌تواند ابزاری کارآمد برای مدیران صندوق‌های سهامی برای مدیریت ریسک و حفظ سرمایه در شرایط بحرانی باشد.

پیشنهادهای پژوهش‌های آتی

بر اساس یافته‌ها، پیشنهادهای زیر برای تحقیقات آتی ارائه می‌شود:

۱- توسعه مدل‌های ترکیبی: ترکیب مدل‌های درخت تصمیم با شبکه‌های عصبی و الگوریتم‌های فراابتکاری می‌تواند دقت پیش‌بینی را افزایش دهد.

۲- استفاده از داده‌های کلان: ورود متغیرهای کلان اقتصادی (نرخ ارز، تورم و نرخ بهره) به مدل‌های پیش‌بینی ارزش سهام.

۳- تحلیل احساسات (Sentiment Analysis): استفاده از داده‌های رسانه‌های اجتماعی و اخبار، و ترکیب داده‌های کمی با داده‌های کیفی (مانند اخبار و تحلیل‌های اجتماعی در فضای مجازی) برای بهبود دقت پیش‌بینی در مدل $Bi-LSTM$.

۴- مطالعات بین‌المللی: مقایسه عملکرد مدل‌ها در بازارهای مختلف (ایران، ترکیه، مالزی و غیره).

۵- تمرکز بر ارزش‌گذاری ذاتی: توسعه مدلی که ارزش ذاتی سهام را برآورد کنند، نه صرفاً قیمت بازار.

۶- استفاده از یادگیری عمیق: بهره‌گیری از شبکه‌های $LSTM$ و $Transformer$ برای مدل‌سازی سری‌های زمانی مالی.

۷- الگوریتم‌های تکاملی: استفاده از الگوریتم ژنتیک (GA) یا الگوریتم ازدحام ذرات (PSO) به جای روش‌های ریاضی کلاسیک برای حل مسائل بهینه‌سازی در پورتفوی‌های بسیار بزرگ.

۸- یادگیری تقویتی (Reinforcement Learning): بررسی امکان استفاده از عامل‌های یادگیری تقویتی (Agents) برای تصمیم‌گیری خودکار و لحظه‌ای در مدیریت پورتفوی.

محدودیت‌های پژوهش

این پژوهش دارای محدودیت‌هایی است:

- ۱- محدودیت پایگاه‌های داده: دسترسی به برخی پایگاه‌های داده خارجی محدود بود.
- ۲- تنوع معیارها: مقالات از معیارهای مختلفی برای سنجش دقت استفاده کرده‌اند که مقایسه مستقیم را دشوار می‌کند.
- ۳- سوگیری انتشار: احتمال انتشار نتایج مثبت بیشتر از نتایج منفی وجود دارد.
- ۴- محدودیت زبانی: فقط مقالات فارسی و انگلیسی بررسی شدند.
- ۵- ماهیت مقدماتی ارزیابی چارچوب پیشنهادی: ارزیابی تجربی چارچوب ترکیبی بر داده‌های یک بازار و یک دوره زمانی محدود انجام شده و هزینه‌های معاملاتی، محدودیت‌های دامنه نوسان و نقدشوندگی در آن لحاظ نشده است؛ تعمیم نتایج به بازارها و دوره‌های دیگر نیازمند آزمون‌های بیشتر است. همچنین معناداری آماری تفاوت نسبت‌های شارپ (برای نمونه با آزمون Jobson-Korkie یا بوت‌استرپ) در این مرحله آزمون نشده است.

منابع

- آذر، ع.، و رجب‌زاده، ع. (۱۳۷۹). ارزیابی روشهای پیشبینی ترکیبی با رویکرد شبکههای عصبی [پایاننامه کارشناسی ارشد]. دانشگاه تربیت مدرس.
- ابراهیم‌نژاد، ع.، زمانی، ش.، و بیریا، م. (۱۳۹۹). پیشبینی رفتار قیمت سهام توسط دفتر سفارشات محدود: شواهدی از بورس اوراق بهادار تهران. مدیریت دارایی و تأمین مالی، ۸(۴)، ۶۳-۸۴.
- ابونوری، ا.، و عبداللهی، م. (۱۳۹۱). مدلسازی نوسانات بخشهای مختلف بازار سهام ایران با استفاده از مدل گارچ چندمتغیره. تحقیقات مالی، ۱۱(۱)، ۱-۱۶. <https://doi.org/10.22059/jfr.2012.36628>
- افشاری، ح. (۱۳۸۲). بررسی ساختاری قابلیت پیشبینی قیمت سهام در بورس اوراق بهادار تهران. فصلنامه بررسیهای حسابداری و حسابرسی، ۱۰(۱۲۶)، ۳۲-۱۰۳.
- افلاطونی، ع. (۱۳۹۲). تجزیه و تحلیل آماری با EViews در تحقیقات حسابداری و مدیریت مالی. انتشارات ترمه.
- البرزی، م.، یعقوب‌نژاد، ا.، و مقصود، ح. (۱۳۸۷). کاربرد شبکههای عصبی مصنوعی در پیشبینی شاخص بازدهی نقدی و قیمت سهام. مطالعات تجربی حسابداری مالی، ۶(۳۲)، ۱۱۹-۱۳۷.
- علیزاده، س.، و ملکمحمدی، س. (۱۳۹۳). داده‌کاوی و کشف دانش گامبه‌گام با نرم‌افزار Clementine. انتشارات دانشگاه صنعتی خواجه نصیرالدین طوسی.
- عربصالحی، م.، و کمالی‌دهکردی، [نام کوچک درج نشده]. (۱۴۰۰). بررسی مقایسه‌ای رابطههای تنزیلی ارزشگذاری سهام در شرکتهای پذیرفته‌شده در بورس اوراق بهادار تهران. چشم‌انداز مدیریت مالی، ۱۱(۳۳)، ۱۱۳-۱۳۸.
- آشتاب، ع. — این نام در فهرست ارسالی شما به‌عنوان نویسنده منبع درج نشده است؛ بنابراین منبعی به آن اضافه نمی‌شود.
- حاتم‌راد، س.، آذرنگی، ب.، اصغرپور، ح.، و حقیقت، ج. (۱۴۰۲). بررسی تأثیرگذاری عوامل کلان اقتصادی بر شاخص قیمت سهام بازار بورس ایران با استفاده از مدل‌های میانگین‌گیری. پژوهشهای اقتصادی ایران، ۲۸(۹۵)، ۱۹۳-۲۳۶. <https://doi.org/10.22054/ijer.2023.71810.1164>
- جهانگیری، خ.، و حسینی ابراهیم‌آباد، س. ع. (۱۳۹۶). بررسی آثار سیاست پولی، نرخ ارز و طلا بر بازار سهام در ایران. فصلنامه تحقیقات مالی، ۱۹(۳۸۹-۴۱۴).
- تهرانی، ر.، پیروز، ب.، و گورانی، [نام کوچک درج نشده]. (۱۳۹۳). امکان‌سنجی استفاده از مدل ترکیبی سهام (CANSLIM) در بورس اوراق بهادار تهران. فصلنامه دانش مالی تحلیل اوراق بهادار، ۷(۲۱)، ۲۹-۴۲.
- کنی، ا. (۱۳۸۵). تحلیل تکنیکی پیشرفته. شرکت کارگزاری راهبرد سرمایه‌گذاری.
- لشگری، ز.، و همکاران. (۱۳۹۷). رابطه بین رشد تولید ناخالص داخلی و عوامل ریسک سرمایه. تحقیقات حسابداری و حسابرسی، ۳۸(۳)، ۹۵-۱۱۰.
- روحی‌ملکی، س.، و پاکرام، ع. (۱۳۹۶). بررسی رابطه محافظه‌کاری حسابداری با واکنش بازار و ارزش شرکتها. حسابداری و منافع اجتماعی، ۴۷(۴)، ۹۹-۱۲۶. <https://doi.org/10.22051/ijar.2017.9221.1093>
- سارنج، ع.، و نوراحمدی، م. (۱۳۹۵). تخمین ارزش در معرض ریسک (VaR) و ریزش مورد انتظار (ES) با استفاده از رویکرد ارزش فرین شرطی در بورس اوراق بهادار تهران. تحقیقات مالی، ۴۳(۴۳).
- صنّعی‌آباد، م.، محمودی، س.، و طاهرپرور، م. (۱۳۹۳). داده‌کاوی کاربردی. انتشارات نیاز دانش.
- عباسیان، ع.، مرادپور اولادی، م.، و عباسیون، و. (۱۳۸۷). اثر متغیرهای کلان اقتصادی بر شاخص کل بورس اوراق بهادار تهران. فصلنامه پژوهشهای اقتصادی ایران، ۱۲(۳۶)، ۱۲۸-۱۵۲.
- Hou, X et al. (2020). An Enriched Time-Series Forecasting Framework for Long-Short Portfolio Strategy. IEEE Access, 8, 31992-32002. [\[https://doi.org/10.1109/ACCESS.2020.2973037\]](https://doi.org/10.1109/ACCESS.2020.2973037)(<https://doi.org/10.1109/ACCESS.2020.2973037>)
- Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Applied Soft Computing, 90, 106181. <https://doi.org/10.1016/j.asoc.2020.106181>
- Kumbure, M. M., Lohrmann, C., Luukka, P., & Porras, J. (2022). Machine learning techniques and data for stock market forecasting: A literature review. Expert Systems with Applications, 197, 116659. <https://doi.org/10.1016/j.eswa.2022.116659>
- Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. Quantitative Finance, 1(2), 223–236. <https://doi.org/10.1080/713665670>
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223–2273. <https://doi.org/10.1093/rfs/hhaa009>

- Fernandez, P. (2002). Valuation methods and shareholder value creation. Academic Press.
- Fernandez, P. (2004). The value of tax shields is NOT equal to the present value of tax shields. *Journal of Financial Economics*, 73(1), 145–165. https://doi.org/10.1016/j.jfineco.2002.10.001
- Bicer, B. (2006). Stock valuation for Investment and corporate decisions. ProQuest.
- Appel, G. (2005). Technical analysis: Power tools for active investors. FT Prentice Hall.
- Achelis, S. B. (2001). Technical analysis from A to Z (2nd ed.). McGraw-Hill.
- Ahmed, I. (2018). Impact of dividend per share and earnings per share on stock prices: A case study from Pakistan (textile sector). *International Journal of Social Sciences, Humanities and Education*, 2(2), 1–10.
- Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1), 3–18. https://doi.org/10.1016/0304-405X(81)90018-0
- Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing*, 7(3–4), 197–387. https://doi.org/10.1561/20000000039
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x
- Goldfarb, D., & Iyengar, G. (2003). Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1), 1–38. https://doi.org/10.1287/moor.28.1.1.14260
- Cornuejols, G., & Tütüncü, R. (2006). Optimization methods in finance. Cambridge University Press.
- Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21–41. https://doi.org/10.21314/JOR.2000.038
- Quaranta, A. G., & Zaffaroni, A. (2008). Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking & Finance*, 32(10), 2046–2056. https://doi.org/10.1016/j.jbankfin.2007.12.025
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054
- Siarni-Namini, S., Tavakoli, N., & Namin, A. S. (2019). The performance of LSTM and BiLSTM in forecasting time series. In 2019 IEEE International Conference on Big Data (pp. 3285–3292). IEEE. https://doi.org/10.1109/BigData47090.2019.9005997
- Zhang, Z., Zohren, S., & Roberts, S. (2020). Deep learning for portfolio optimization. *The Journal of Financial Data Science*, 2(4), 8–20. https://doi.org/10.3905/jfds.2020.1.042
- Li, X., & Wu, P. (2022). Stock price prediction incorporating market style clustering. *Cognitive Computation*, 14(1), 149–166. https://doi.org/10.1007/s12559-021-09820-1
- Deng, Y., Hope, O.-K., Wang, C., & Zhang, M. (2022). Capital market liberalization and auditors' accounting adjustments: Evidence from a quasi-experiment. *Journal of Business Finance & Accounting*, 49(1–2), 215–248. https://doi.org/10.1111/jbfa.12541
- Elmghaamez, I. K., Attah-Boakye, R., Adams, K., & Agyemang, J. (2022). The diffusion of innovation theory and the effects of IFRS adoption by multinational corporations on capital market performance: A cross-country analysis. *Thunderbird International Business Review*, 64(1), 81–108. https://doi.org/10.1002/tie.22243
- Sadorsky, P. (2022). Forecasting solar stock prices using tree-based machine learning classification: How important are silver prices? *The North American Journal of Economics and Finance*, 61, 101705. https://doi.org/10.1016/j.najef.2022.101705
- Ang, A., Chen, J., & Xing, Y. (2006). Downside risk. *The Review of Financial Studies*, 19(4), 1191–1239. https://doi.org/10.1093/rfs/hhj035
- Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine learning approaches in stock price prediction: A systematic review. *Journal of Physics: Conference Series*, 2161, 012065. https://doi.org/10.1088/1742-6596/2161/1/012065
- Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and artificial neural networks for improved stock price prediction. *Expert Systems with Applications*, 44, 320–331. https://doi.org/10.1016/j.eswa.2015.09.029
- Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A*, 379(2194), 20200209. https://doi.org/10.1098/rsta.2020.0209
- Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35(1), 73–101. https://doi.org/10.1214/aoms/1177703732

- Alipour, M. A., & Parsinejad, S. (2018). Designing an appropriate neural network to assist in more accurate stock portfolio predictions. *Transactions on Machine Intelligence*, 1(4), 228–242. https://doi.org/10.47176/TMI.2018.228
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. Morgan Kaufmann.
- Hu, H., Tang, L., Zhang, S., & Wang, H. (2018). Predicting the direction of stock markets using optimized neural networks with Google Trends. *Neurocomputing*, 285, 188–195. <https://doi.org/10.1016/j.neucom.2018.01.038>
- Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), Article 27. <https://doi.org/10.1145/1961189.1961199>
- Qiu, Y., et al. (2022). “Short-term stock trends prediction based on sentiment analysis and machine learning.” *Soft Computing*, 26, 2209–2224. https://doi.org/10.1007/s00500-021-06602-7
- Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.1412.6980>