Journal of Advances in Finance and Investment
Volume 5, Issue 3, 2024
pp. 27-64.
Paper type: Research paper

The modeling the fixed asset investing with a machine learning approach by emphasizing the role of regulatory criteria

Farzaneh Shams Doost[1], Omid Mahmoudi Khoshro[2], Ataollah Mohammadi Malgharni[3], Amir Sheikhahmadi[4]

Received: 12/11/2023 Accepted: 15/09/2024

Extended Abstract

Introduction

The purpose of this research is to model fixed asset investment with the role of regulatory criteria and artificial intelligence approach of companies admitted to the Tehran Stock Exchange. Due to the characteristics of the specific market and lack of sufficient knowledge of the investors towards the market and behavioral trends, it causes that the country's capital market, which is the heart of the country's economy, is not efficient and dynamic enough and it causes investors to make mistakes in their investment decisions due to not knowing enough about this and this has provided the ground for their exit from the capital market and has inappropriate consequences for the country. Therefore, the introduction of investment opportunities will encourage market participants to invest and reduce financial crises and avoid large losses in the stock market.

Literature Review

The purpose of corporate governance is to create an appropriate framework based on which an appropriate balance is established between the interests of various stakeholders, the freedom of management and the accountability of management. In other words, corporate governance is a mechanism that, through its application, problems of representation disarise and the quality of information disclosed by companies improves and respecting the rights of shareholders and all stakeholders, including the government and the general public, will be the focus of attention. In industries where there is a lot of competition, stronger corporate governance mechanisms prevail over it and the existence of competition in the

1. Department of Finance, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran.
2. Department of Accounting, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran. (Corresponding Author). omkhoshro@iausdj.ac.ir
3. Department of Accounting, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran.
4. Department of Computer Engineering, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran.

How to cite this paper: Shams Doost, F., Mahmoudi Khoshro, O., Mohammadi Malgharni, A., & Sheikhahmadi, A. (2024). Modeling fixed asset investment using a machine learning approach with emphasis on the role of regulatory criteria. *Advances in Finance and Investment, 5*(3), 27-64. [In Persian]

https://doi.org/00.00000/afi.0000.0000000.0000

market and among active companies in an industry creates the necessary motivation for managers to perform more efficiently. On the other hand, competition in the product market reduces prices and controller costs in the market. This issue is a great opportunity for investors and competitors. This issue is a great opportunity for investors and competitors. To measure the highest performance of the company among competitors, the presence of capable management can play a role in increasing the efficiency. Each investment buys shares of a company according to the type of taste and information it acquires. According to the type of taste and the information they get, investors buy shares of companies whose investment policies they consider favorable.

Research Methodology

The variables of the companies are selected based on the RRelief-F method. Then the data is divided into K-Fold validation and the data is divided into test and training data by cross-validation. Then the training data is calculated with four linear and non-linear artificial intelligence algorithms PINSVR and linear and non-linear PLS. In the training phase of linear and non-linear models after learning, the same training-validation data without the dependent variable is provided to them again to determine the value of the fixed asset prediction variable. To check the learning power of the models using the three evaluation criteria errors MAE, MSE and SMAPE in the training phase for each criterion, the errors are reported for the current year and the next year. The statistical population of the current research is all the companies admitted to the Tehran Stock Exchange in the period from 2012 to 2021 and the financial information of 101 companies was used.

Results

The results of the research hypothesis test showed that in terms of regulatory criteria, the size of the board of directors, the independence of the board of directors, the financial expertise of the board of directors, the size of the audit committee, the independence of the audit committee, institutional ownership above 5%, the tenure of the CEO, the existence of an internal auditor, the ratio of expertise Committee members and CEO duality play an important role in predicting firms' fixed investment. Also, linear artificial intelligence algorithms are more efficient in predicting investment in fixed asset of companies than non-linear algorithms.

Discussion and Conclusion

Investing in various matters by companies has always been considered as one of the important ways of developing companies and preventing stagnation and backwardness. In the meantime, the limitation in resources has caused that in addition to the development of investment, increasing the efficiency of investment is of great importance. In general, investment efficiency means accepting projects with real value and investment inefficiency means choosing projects with real and

negative value. Investment efficiency means that there is no deviation from the expected investment level. Almost the majority of economists, regardless of their school and intellectual point of view, put a lot of emphasis on the formation of capital and carrying out strategy in order to increase capital as the most important determining factor of growth and development. In fact, the investment of each commercial unit should be done according to the limited resources and its efficiency; but the main issue is choosing projects and making decisions about investment opportunities and efficiency by the managers of commercial units, which is based on their personal interests. In other words, information asymmetry and conflict of interests prevent an optimal and efficient investment.

The results of this research show that compared to other artificial intelligence methods, the PINSVR method not only achieves comparable regression performance; but also achieves better estimates. The KPLS method also reduces the dimensions of the feature, this method is able to find a large number of hidden variables. For example, it can represent twenty to thousands of features in a subspace of very small dimensions, which may have a large impact on real-time applications. It also has an output vector that can have multiple labels, so that it can solve multiple related problems in general.

Conflict of Interest

The authors of this article declared no conflict of interest regarding the authorship or publication of this article.

Keywords: Artificial Intelligence, Fixed asset invetment, Linear and non-linear models, Regulatory criteria.

JEL Classification: G1, G11.

پیشرفت‌های مالی و سرمایه‌گذاری
سال پنجم، پاییز ۱۴۰۳ – شماره ۳
صفحات ۶۴-۲۶
نوع مقاله: پژوهشی

الگوسازی سرمایه‌گذاری دارایی ثابت با نقش معیارهای نظارتی و رویکرد هوش مصنوعی

فرزانه شمس دوست[۱]، امید محمودی خوشرو[۲]، عطاءالله محمدی ملقرنی[۳]، امیر شیخ‌احمدی[۴]

تاریخ دریافت: ۱۴۰۲/۰۸/۲۱ تاریخ پذیرش: ۱۴۰۳/۰۶/۲۵

چکیده

هدف: سازوکارهای حاکمیت شرکتی و ساختار مالکیت قادر هستند که مستقیماً بر روی انگیزه سرمایه‌گذاران برای وادارکردن مدیریت به استفاده کارا از منابع موجود در سازمان‌ها اثر بگذارند. هدف این پژوهش الگوسازی سرمایه‌گذاری دارایی ثابت با نقش معیارهای نظارتی و رویکرد هوش مصنوعی شرکت‌های پذیرفته‌شده بورس اوراق بهادار تهران می‌باشد.

روش‌شناسی پژوهش: در این پژوهش، نمونه‌ها با استفاده از روش رگرسیون متغیرگزینی ریلیف – اف انتخاب شدند؛ سپس داده‌های پژوهش به روش اعتبارسنجی متقابل ده‌قسمتی به گروه‌های داده‌های آموزش و اعتبارسنجی و داده‌های آزمون تقسیم شدند. سپس با دو الگوریتم خطی و غیرخطی بردار پشتیبان غیرموازی غیرحساس پارامتریک و حداقل مربعات جزئی هسته و الگوریتم حداقل مربعات جزئی هسته غیرخطی مورد آزمون قرار گرفتند. جامعه آماری پژوهش حاضر کلیه شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران در دوره‌های زمانی ۱۳۹۰ الی ۱۳۹۹ است و از اطلاعات مالی ۱۰۱ شرکت استفاده شده است.

یافته‌ها: نتایج آزمون پژوهش نشان داد که متغیرهای اندازه (تعداد اعضای) هیئت‌مدیره، استقلال هیئت‌مدیره، تخصص مالی هیئت‌مدیره، اندازه (تعداد اعضای) کمیته حسابرسی، استقلال کمیته حسابرسی، مالکیت نهادی بالای ۵ درصد، دوره تصدی مدیرعامل، وجود حسابرس داخلی، نسبت تخصص اعضای کمیته و دوگانگی وظایف مدیرعامل بیشترین تأثیر را در پیش‌بینی سرمایه‌گذاری دارایی ثابت شرکت‌ها دارند. همچنین در بررسی الگوریتم‌های هوش مصنوعی خطی و غیرخطی در پیش‌بینی سرمایه‌گذاری دارایی ثابت شرکت‌ها، الگوریتم‌های خطی نسبت به الگوریتم‌های غیرخطی کارایی بیش‌تری نشان دادند.

اصالت / ارزش‌افزوده علمی: باتوجه‌به اینکه ویژگی‌های خاص بازار و عدم شناخت کافی سرمایه‌گذاران نسبت به بازار و تورش‌های رفتاری موجب شده که بازار سرمایه کشور که به منزله قلب پیکره اقتصاد کشور به شمار می‌آید از کارایی و پویایی کافی برخوردار نباشد و سرمایه‌گذاران به علت عدم شناخت کافی از این دام‌ها در اتخاذ تصمیم سرمایه‌گذاری خود دچار اشتباه شوند و همین امر زمینه خروج آن‌ها را از بازار سرمایه فراهم آورده و پیامدهای نامناسبی برای کشور در پی داشته باشد؛ لذا معرفی فرصت‌های سرمایه‌گذاری در این پژوهش موجب ترغیب فعالین بازار به سرمایه‌گذاری و کاهش بحران‌های مالی و اجتناب از بروز زیان‌های بزرگ در بازار سهام خواهد شد.

کلیدواژه‌ها: سرمایه‌گذاری دارایی ثابت، مدل‌های خطی و غیرخطی، معیار نظارتی، هوش مصنوعی.
طبقه‌بندی موضوعی: G1, G11.

۱. گروه مالی، واحد سنندج، دانشگاه آزاد اسلامی، سنندج، ایران.
۲. گروه حسابداری، واحد سنندج، دانشگاه آزاد اسلامی، سنندج، ایران. (نویسنده مسئول). omkhoshro@iausdj.ac.ir
۳. گروه حسابداری، واحد سنندج، دانشگاه آزاد اسلامی، سنندج، ایران.
۴. گروه مهندسی کامپیوتر، واحد سنندج، دانشگاه آزاد اسلامی، سنندج، ایران.

استناد: شمس دوست، فرزانه؛ محمودی خوشرو، امید؛ محمدی ملقرنی، عطاءالله؛ شیخ‌احمدی، امیر. (۱۴۰۳). الگوسازی سرمایه‌گذاری دارایی ثابت با نقش معیارهای نظارتی و رویکرد هوش مصنوعی. *پیشرفت‌های مالی و سرمایه‌گذاری*، ۵(۳). ۶۴-۲۶.

۱- مقدمه

سرمایه‌گذاری در فرآیند ارزش‌آفرینی شرکت یکی از موضوعات مهم تلقی می‌شود و می‌تواند به‌عنوان ابزاری مناسب در جهت بسط و توسعه و همچنین جلوگیری از عقب‌ماندگی و رکود شرکت تلقی گردد؛ اما مسئله‌ای که در این‌گونه سرمایه‌گذاری‌ها از اهمیت بسزایی برخوردار است، میزان سرمایه‌گذاری‌ها، نرخ بازده سرمایه‌گذاری و به‌طورکلی کارابودن چنین سرمایه‌گذاری‌هایی است. به‌عبارت‌دیگر، کارایی سرمایه‌گذاری یا سرمایه‌گذاری در حد بهینه مستلزم آن است که از یک سو از مصرف منابع در فعالیت‌هایی که سرمایه‌گذاری در آن بیش از حد مطلوب انجام شده است جلوگیری شود و از سوی دیگر، منابع به سمت فعالیت‌هایی که نیاز به سرمایه‌گذاری بیش‌تری دارد هدایت شود (Tahmooresi *et al*., 2024). سرمایه‌گذاری یکی از بحث‌های مهم در بازار سرمایه است؛ به‌طوری‌که یک مدیر توانمند می‌تواند بر ارزش شرکت تأثیرگذار باشد و باعث پویایی در بازده‌های آتی سرمایه‌گذاری‌های شرکت گردد؛ کارایی سرمایه‌گذاری مستلزم آن است که از یک سو از مصرف منابع در فعالیت‌هایی که سرمایه‌گذاری در آن بیش‌تر از حد مطلوب انجام شده است ممانعت شود (جلوگیری از بیش سرمایه‌گذاری) و از دیگر سو، منابع به سمت فعالیت‌هایی که نیاز بیش‌تری به سرمایه‌گذاری دارد (جلوگیری از سرمایه‌گذاری کم) هدایت شود. به طور مفهومی، کارایی سرمایه‌گذاری زمانی حاصل می‌شود که شرکت تنها در تمام پروژه‌های با ارزش فعلی مثبت سرمایه‌گذاری کند. البته این سناریو در صورتی سازگار است که بازار کامل باشد و هیچ یک از مسائل بازار ناقص از جمله گزینش نادرست و هزینه‌های نمایندگی وجود نداشته باشد (Feridoni *et al*., 2022). یکی از عوامل مهم جهت حل مسائل اقتصادی کشورها، بسط و توسعه سرمایه‌گذاری است؛ اما این امر به‌تنهایی کافی نیست و باتوجه‌به محدودیت منابع مالی، علاوه بر مسئله توسعه سرمایه‌گذاری، افزایش کارایی سرمایه‌گذاری نیز از جمله مسائل بااهمیت است؛ به طور مفهومی، کارایی سرمایه‌گذاری زمانی حاصل می‌شود که شرکت، فقط در تمامی طرح‌هایی با ارزش فعلی خالص مثبت سرمایه‌گذاری کند. در واقع، برخی از نواقص بازار سرمایه مانند عدم تقارن اطلاعاتی و هزینه‌های نمایندگی، می‌تواند به فرآیند سرمایه‌گذاری بیش‌تر از حد یا سرمایه‌گذاری کمتر از حد منتهی گردد. به این مفهوم که نه طرح‌های با ارزش فعلی خالص مثبت (سرمایه‌گذاری بیش‌تر از حد) و نه طرح‌های با ارزش فعلی خالص منفی (سرمایه‌گذاری کم‌تر از حد) رد نمی‌شود، است (Khoshkar *et al*., 2020). پرداختن به موضوع خطر نشان می‌دهد که واکنش به خطر جنبه مهمی است که در سال‌های اخیر در خصوص افزایش کارایی سرمایه‌گذاری و کاهش ناکارایی سرمایه‌گذاری موردتوجه زیادی قرار گرفته است. مدیری توانمندتر به شمار می‌آید که از جنبه واکنش به خطر نیز توانایی خود را به‌خوبی به نمایش بگذارد. مواجهه با خطرهای متعدد می‌تواند توان مدیریت

را تحت‌تأثیر قرار دهد. از این دید چگونگی واکنش به خطر، میدان آزمایشی برای ارزیابی توان مدیریت در جهت تأثیرگذاری بر کارایی سرمایه‌گذاری است. نااطمینانی حاصل از وجود خطر در عملیات هر شرکتی وجود دارد؛ اما اعتقاد بر این است که نحوه برخورد با خطرها می‌تواند تهدید ناشی از آن خطر را به فرصتی برای رشد و کارایی سرمایه‌گذاری‌های شرکت تبدیل نماید که البته این به مدیریت خطر شرکت بستگی دارد (Sayyadi *et al*., 2019). یکی از نقش‌های بسیار مهم حسابداری، تخصیص کارآمد سرمایه است؛ بنابراین، تعجب‌آور نیست که ادبیات حاضر در این زمینه، بر نقش حسابداری در تصمیم‌های تخصیص سرمایه تمرکز دارند؛ ازاین‌رو، یکی از چالش‌های مشترک این مطالعات چگونگی شناسایی و اندازه‌گیری کارایی سرمایه‌گذاری است. از لحاظ تئوری، تخصیص کارآمد سرمایه به معنای گردش سرمایه برای استفاده بهینه و ارزشمند آن است. با وجود این، مشاهده جریان سرمایه و تمایز ارزش‌های بالا و پایین سرمایه‌گذاری، مشکل است (Naim and Condamin, 2019). رویکرد ترغیب سرمایه‌گذاران به سرمایه‌گذاری در شرکت‌ها از اهمیت زیادی برخوردار است. این امر به‌ویژه زمانی اهمیت بیشتری دارد که سازوکار حاکمیت شرکتی و ساختار مالکیت منابع موجود را بر سرمایه‌گذاری با استفاده از منابع کارا شناسایی نماید. هر چند برخی مطالعات داخلی بر نقش معیارهای نظارتی بر سرمایه‌گذاری در بورس اوراق بهادار تهران موردتوجه قرار گرفته است، درعین‌حال، در این مطالعات بر روی متغیرگزینی برای کاهش میزان خطا و محاسبه روش‌های خطی و غیرخطی در پیش‌بینی سرمایه‌گذاری در شرکت‌های بورسی کم‌تر تمرکز شده است. ازاین‌رو، در راستای پوشش این شکاف مطالعاتی، در پژوهش حاضر تلاش شده است الگوسازی سرمایه‌گذاری دارایی ثابت شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران با نقش معیارهای نظارتی راهبری و رویکرد هوش مصنوعی مورد ارزیابی قرار گیرد و فرضیه‌های پژوهش بررسی شود: ۱) الگوریتم‌های خطی و غیرخطی هوش مصنوعی، توانایی پیش‌بینی سرمایه‌گذاری در دارایی ثابت شرکت‌ها را دارد. ۲) معیارهای نظارتی نظام راهبری شرکتی بر میزان سرمایه‌گذاری دارایی ثابت شرکت‌ها تأثیرگذار هستند و از طرفی دیگر مسئله اصلی پژوهش در قالب سؤالات پژوهش عبارتند از: ۱) تأثیر معیارهای حسابداری نظارتی بر سرمایه‌گذاری شرکت در دارایی ثابت چگونه است؟ ۲) الگوریتم‌های خطی و غیرخطی هوش مصنوعی چه تأثیری بر سرمایه‌گذاری دارایی ثابت دارند؟ شایان‌ذکر است مسئله اساسی که این پژوهش برای حل آن تلاش می‌کند، ارائه یک استراتژی علمیِ بهینه برای پیش‌بینی سرمایه‌گذاری در دارایی‌های ثابت شرکت‌ها با هدف افزایش بازدهی و کاهش ریسک در بازارهای مالی است.

۲- مبانی نظری و پیشینه پژوهش

پیش‌بینی سرمایه‌گذاری بسیار ارزشمند است چرا که انسان همیشه در پی یافتن سازوکاری است که بتواند با یک سرمایه‌گذاری مناسب سود چشمگیری به دست آورد. ازاین‌رو، همواره افراد به دنبال موقعیت‌های اقتصادی بهتر برای سرمایه‌گذاری می‌گردند. در این رهگذر، بازارهای مالی موقعیت‌های جذابی برای کسب درآمد و تولید ثروت به وجود می‌آورند؛ کسب سود چشمگیر، نیازمند پذیرش خطر است. چنین خطری بدون برخورداری از دانش و مهارت کافی قابل‌کنترل نخواهد بود؛ ازاین‌رو، یکی از مسائل اساسی در مدیریت مالی موضوعِ اداره صحیح فرآیند سرمایه‌گذاری است که با عنوان بهینه‌سازی سبد سرمایه‌گذاری یا بهینه‌سازی سبد سرمایه‌گذاری شناخته می‌شود. بازارها مهم‌ترین جایگاه تخصیص منابع مالی به فعالیت‌های اقتصادی در سراسر دنیا به شمار می‌روند. ازاین‌رو، توسعه سازوکارهای اثربخش برای مدیریت بهینه فرآیند سرمایه‌گذاری در این بازارها از اهمیت زیادی برخوردار است؛ در این راستا، شرکت‌های سرمایه‌گذار و مؤسسات مالی نقش بسیار پررنگ‌تری ایفا می‌کنند؛ ازاین‌رو، بهینه‌سازیِ فرآیندِ سرمایه‌گذاری شرکت سرمایه‌گذار در این بازارها باعث گردش بهترِ چرخ‌های تولید ملی و اقتصاد کشور خواهد شد؛ لذا شناسایی و به‌کارگیری شیوه‌های بهینه جهت افزایش بازده و کاهش ریسک ناشی از ورود شرکت سرمایه‌گذار به این قبیل فعالیت‌ها ضروری و اجتناب‌ناپذیر است (Amiri *et al*., 2021). واحدهای تجاری، همواره با فرصت‌های سرمایه‌گذاری زیادی روبه‌رو می‌شوند و نیازمند تصمیم‌گیری منطقی نسبت به یک سرمایه‌گذاری بهینه هستند. در واقع، سرمایه‌گذاری هر واحد تجاری باید باتوجه‌به محدودیت منابع و کارابودن آن صورت بگیرد؛ اما مسئلهٔ اصلی، انتخاب طرح‌ها و تصمیم‌گیری راجع به فرصت‌ها و کارایی سرمایه‌گذاری به وسیلهٔ مدیران واحدهای تجاری است که بر اساس منافع شخصی آن‌ها صورت می‌گیرد. به‌عبارتی‌دیگر، عدم تقارن اطلاعاتی و تضاد منافع، مانع انجام یک سرمایه‌گذاری بهینه و کارا می‌شود (Lara *et al*., 2016). کارایی سرمایه‌گذاری به این معنی است که هیچ انحرافی از سطح سرمایه‌گذاری مورد انتظار وجود نداشته باشد. شرکت‌هایی که بالاتر از حد مطلوب سرمایه‌گذاری می‌کنند (انحراف مثبت از سرمایه‌گذاری مورد انتظار) دچار فرآیند سرمایه‌گذاری بیش از حد می‌شوند. درحالی‌که شرکت‌هایی که همه پروژه‌های سودآور را حفظ نمی‌کنند انحراف منفی از سرمایه‌گذاری مورد انتظار، دچار فرآیند سرمایه‌گذاری کمتر از حد می‌شوند (Samet and Jarboui, 2017). در تعیین کارایی سرمایه‌گذاری، حداقل دو معیار نظری وجود دارد: معیار اول بیان می‌کند که به‌منظور تأمین مالی فرصت‌های سرمایه‌گذاری، نیاز به جمع‌آوری منابع وجود دارد و معیار دوم عنوان می‌کند که در یک بازار کارا، همه پروژه‌های با خالص ارزش فعلی مثبت، باید تأمین مالی شوند. اگر چه، تعداد زیادی از پژوهش‌های موجود در حوزه مالی نشان داده است که محدودیت‌های مالی، توانایی مدیران در تأمین

مالی را محدود می‌سازد (Hubbard, 1998). عمدتاً جنبه‌های یادگیری به گونه‌های مختلف تعریف می‌شوند؛ به‌طوری‌که می‌توان یک ماشین را برای حل انواع مسائل با ویژگی‌هایی مانند تشخیص، استفاده از زبان، شکل‌گیری انتزاعی و مفاهیم کلی شبیه‌سازی کرد. انواع سیستم‌های یادگیری ماشینی به سه دسته تقسیم می‌شوند: ۱) سیستمی که با نظارت انسان آموزش دیده است یا خیر؟ شامل: یادگیری نظارت‌شده، نظارت‌نشده و نیم‌نظارت‌شده. ۲) آیا سیستم در حین فعالیت هم آموزش می‌بیند؟ شامل یادگیری آنلاین در مقابل یادگیری بسته محور (آفلاین). ۳) آیا سیستم فقط داده‌های جدید (ناشناخته) را با داده‌های موجود (معلوم) مقایسه می‌کند یا به‌جای آن در طی فرآیند آموزش الگوها را تشخیص می‌دهد (مانند یک دانشمند) مدلی می‌سازد که بتواند پیش‌بینی انجام دهد؛ مانند یادگیری نمونه محور در مقابل یادگیری مدل محور. از مزایای هوش مصنوعی می‌توان به کاهش خطای انسانی، کمک به انجام کارهای تکراری، دستیار دیجیتالی، تصمیم‌گیری سریع، حل مسائل پیچیده در اختراعات جدید و یادگیری سریع وظایف اشاره کرد (Géron, 2022). یکی از بهبودهایی که در مطالعات جدید روش‌های هوش مصنوعی به کار گرفته شده است، استفاده از انتخاب ویژگی به‌عنوان یک پیش مرحله برای مدل طبقه‌بندی‌کننده اصلی می‌باشد؛ بااین‌حال، طبق پژوهش‌های انجام‌شده، ماشین بردار پشتیبان عملکرد قابل‌قبولی در پیش‌بینی سرمایه‌گذاری دارد؛ اما دقت عملکرد آن به طور قابل‌توجهی تحت‌تأثیر تعداد ویژگی‌های متغیرهای ورودی است؛ بنابراین، تعداد ویژگی‌هایی که باید در یادگیری ماشین استفاده شود، تأثیر بسزایی در افزایش دقت نتایج و کاهش هزینه دارد و باتوجه‌به نتایج بسیاری از مطالعات داخلی و خارجی، برتری روش‌های ماشین بردار پشتیبان نسبت به روش‌هایی مانند شبکه‌های عصبی مصنوعی[1] و همچنین روش‌های آماری به اثبات رسیده است (Zhang *et al*., 2014).

خالقی زاده دهکردی و همکاران (Khaleghi Zadeh Dehkordi *et al*., 2022) کاربرد الگوریتم هوش مصنوعی در پیش‌بینی کارایی سرمایه‌گذاری با تأکید بر نقش معیارهای مدیریت ریسک را مورد مطالعه قرار دادند. بدین منظور از اطلاعات مالی ۱۳۹ شرکت بین سال‌های ۱۳۸۸ الی ۱۳۹۷ با استفاده از روش‌های هوش مصنوعی ژنتیک و شبکه عصبی استفاده شد. نتایج پژوهش حاکی از تأثیر معیارهای مدیریت ریسک بر کارایی سرمایه‌گذاری می‌باشد و همچنین نتایج نشان داد که روش هوش مصنوعی شبکه عصبی نسبت به روش هوش مصنوعی ژنتیک قدرت بالاتری جهت پیش‌بینی کارایی سرمایه‌گذاری شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران را دارد. الوی و شیانگ (Lv and Xiong, 2022) در مطالعه‌ای به بررسی این موضوع پرداختند که آیا فناوری اطلاعات می‌تواند کارایی سرمایه‌گذاری شرکتی را بهبود بخشد؟ این مطالعه با بهره‌برداری از نمونه بزرگی از شرکت‌های

1. Artificial Neural Networks (ANN)

فهرست‌شده چینی و داده‌های تابلویی استانی در بازه زمانی ۲۰۱۱ تا ۲۰۱۸، انجام شد و دریافتند که کارایی سرمایه‌گذاری شرکت به طور مثبت با سطح فناوری اطلاعات مرتبط است. علاوه بر این، هم تنوع مالکیت شرکت‌ها و هم جدایی مالکیت و مدیریت به طور قابل‌توجهی ارتباط مثبت را تقویت می‌کند که نشان‌دهنده یک اثر مکمل بین تأثیر فناوری اطلاعات بر کارایی سرمایه‌گذاری شرکتی و دو مکانیسم حاکمیت شرکتی است. منشاوی و همکاران (Menshawy *et al*., 2023) اهمیت ویژگی‌های هیئت‌مدیره و عدم تقارن اطلاعات را در افزایش کارایی سرمایه‌گذاری در میان شرکت‌های مصری را از سال ۲۰۱۴ تا ۲۰۱۸، مورد بررسی قرار دادند؛ نتایج نشان داد که استقلال هیئت‌مدیره، پاداش هیئت‌مدیره و مدیریت هیئت‌مدیره با کارایی سرمایه‌گذاری و عدم تقارن اطلاعات مرتبط هستند. این یافته‌ها در بخش‌های سرمایه‌گذاری بیش از حد و سرمایه‌گذاری‌های کم در سراسر نمونه سازگار بودند. محمدی و همکاران (Mohammadi *et al*., 2023) به ساختن مدل تصمیم‌گیری حاصل از تشکیل سبد سرمایه‌گذاری بهینه سهام شرکت‌ها که با استفاده از الگوریتم‌های یادگیری ماشین و نظریه فازی که بازدهی بهتر از میانگین شاخص بورس کل اوراق بهادار داشته باشند پرداختند؛ در این پژوهش که در بازه زمانی ۱۳۹۷ تا ۱۴۰۱ انجام شده است ورودی‌های بهینه‌سازی با روش فازی با هدف کمینه‌کردن ریسک با سنجه‌های ریسک ارزش در معرض خطر و ارزش در معرض خطر شرطی مورد بررسی قرار گرفتند نتایج پس از مقایسه الگوریتم‌ها به همراه سنجه‌های ریسک با بازده واقعی بازار بیانگر برتری سنجه ریسک ارزش در معرض خطر و الگوریتم ماشین بردار پشتیبان در انتخاب سبد سرمایه‌گذاری بود. بحرینی و همکاران (Bahreini *et al*., 2023) به ارائه مدلی جامع با استفاده از شبکه مصنوعی حافظه کوتاه‌مدت طولانی[1] برای شناسایی و ارزیابی عوامل مؤثر جهت ارزش‌گذاری معاملات بلوکی از صورت‌های مالی و داده‌های تابلویی روزانه معاملات بلوکی ۶۴ شرکت فعال در بازه زمانی سال‌های ۱۳۹۰ تا ۱۴۰۰ پرداختند؛ در این پژوهش برای ارزش‌گذاری معاملات بلوکی از ۱۵ شاخص به‌عنوان متغیرهای مستقل استفاده شده است و نتایج نشان داد که شبکه عصبی طراحی‌شده توانایی پیش‌بینی ارزش‌گذاری معاملات بلوکی با استفاده از شاخص‌های اندازه بلوک، نوسانات قیمت سهام و بازده صنعت را دارا می‌باشد. تقی‌زاده خانگاه و همکاران (Taghizadeh Khanqah *et al*., 2021) به اعتبارسنجی مدل‌های کارایی سرمایه‌گذاری بر اساس نظریه نمایندگی، عدم تقارن اطلاعاتی، جبهه‌گیری مدیریتی و حداکثرسازی ارزش شرکت اقدام کردند؛ برای این منظور از ۱۸۰ شرکت برای دوره زمانی ۱۳۸۶ الی ۱۳۹۶ استفاده شده است و یافته‌ها نشان داد جریان‌های نقد آزاد و محدودیت مالی به ترتیب بر سرمایه‌گذاری بیش‌تر و کم‌تر از حد تأثیر مثبت دارد؛ همچنین یافته‌ها حاکی از آن بود که کارایی

1. Long Short Term Memory (LSTM)

سرمایه‌گذاری بر ارزش افزوده اقتصادی و ارزش شرکت تأثیر مثبت دارد؛ اما این تأثیر از طریق تمام مدل‌های کارایی سرمایه‌گذاری تأیید نشد و نتایج نشان داد که آزمون تمام فرضیه‌ها بر اساس مدل بومی کارایی سرمایه‌گذاری مورد تأیید واقع شده است. منصورفر و همکاران (Mansourfar *et al.*, 2020) نقش تعدیلگری ابعاد درونی و بیرونی حاکمیت شرکتی بر رابطه بین عدم تقارن اطلاعاتی و کارایی سرمایه‌گذاری را بررسی کردند. برای دستیابی به اهداف پژوهش، ۱۰۶ شرکت برای دوره زمانی ۱۳۸۷ تا ۱۳۹۶ انتخاب شدند. طبق مبانی نظری و یافته‌های پژوهش، متغیر عدم تقارن اطلاعاتی با کارایی سرمایه‌گذاری رابطه منفی و معناداری دارد. هی و همکاران (He *et al.*, 2019) به مطالعه و ارزیابی کارایی سرمایه‌گذاری با رویکرد یادگیری ماشین پیش‌بینی می‌پردازند؛ این مقاله یک روش شبکه عصبی خودکار غیرخطی برای ارزیابی عملکرد سرمایه‌گذاری شرکت‌های دولتی پیشنهاد می‌دهد. این روش با روش سنتی مبتنی بر یادگیری ماشین، مانند رگرسیون خطی، معادله ساختاری، خوشه‌بندی و تحلیل مؤلفه‌های اصلی متفاوت است و از روش پیش‌بینی رگرسیون برای تحلیل کارایی سرمایه‌گذاری استفاده می‌شود؛ در ابتدا رابطه بین تنوع مالکیت، اهرم بدهی شرکت‌ها و کارایی سرمایه‌گذاری شرکت‌های دولتی را تحلیل می‌کنند، در مرحله دوم، مجموعه‌ای از سیستم شاخص ارزیابی کارایی سرمایه‌گذاری برای شرکت‌های دولتی ساخته شد و از یک روش شبکه عصبی خودرگرسیون غیرخطی برای تأیید استفاده شد؛ اطلاعات مربوط به سهام شرکت‌های دولتی در شانگهای و شنژن از سال ۲۰۰۹ تا ۲۰۱۸ است که به‌عنوان نمونه گرفته شده است و نتایج تجربی نشان می‌دهد که مقدار خروجی از شبکه عصبی خودرگرسیون غیرخطی بسیار متناسب با داده‌های واقعی است؛ این پژوهش تأثیر مستقیم تنوع مالکیت بر کارایی سرمایه‌گذاری شرکت‌های دولتی و تأثیر غیرمستقیم بر کارایی سرمایه‌گذاری شرکت‌های دولتی از طریق اهرم بدهی شرکت‌ها را تأیید می‌کند. سیلورستین (Silverstein, 2020) به بررسی این سؤال می‌پردازند که ویژگی‌های مدیریتی چگونه بر کارایی سرمایه‌گذاری شرکت‌ها تأثیر می‌گذارد؟ پژوهش‌های قبلی، کارایی سرمایه‌گذاری شرکت را به‌عنوان تابعی از محیط اطلاعات شرکت و حاکمیت داخلی تعیین می‌کنند؛ اما این پژوهش بررسی می‌کند که چگونه فرصت‌طلبی مدیریتی، یک تعارض نمایندگی است که سیاست سرمایه‌گذاری شرکت‌ها را تحریف می‌کند؛ نتایج نشان می‌دهد که فرصت‌طلبی مدیریتی کارایی سرمایه‌گذاری شرکت را کاهش می‌دهد و تأثیرات منفی بر حسابداری و عملکرد مالی شرکت دارد. احمد و همکاران (Ahmed *et al.*, 2020) در مقاله‌ای به این موضوع می‌پردازند که آیا چرخه عمر شرکت بر کارایی سرمایه‌گذاری شرکت‌ها تأثیر می‌گذارد؟ در این مقاله نقش مراحل چرخه عمر شرکت در تعیین کارایی سرمایه‌گذاری ۳۵۱ شرکت در پاکستان طی ۱۲ سال (۲۰۰۵ الی ۲۰۱۶) بررسی شده است. با استفاده از معیارهای مرحله چرخه عمر شرکت دیکینسون،

دریافتند که کارایی سرمایه‌گذاری شرکت در مراحل معرفی و افول پایین‌تر و در مراحل رشد و بلوغ چرخه عمر شرکت بالاتر است. علاوه بر این، نتایج تجزیه‌وتحلیل رگرسیون نشان داد که شرکت‌های بالغ از بالاترین سطح کارایی سرمایه‌گذاری و به دنبال آن شرکت‌های رشد برخوردار هستند. به‌طورکلی، کارایی سرمایه‌گذاری شرکت در مراحل مختلف چرخه عمر شرکت یک روند U شکل وارونه به نمایش گذاشت. وانگ و همکاران (Wang *et al*., 2021) یک ابزار مدیریت ریسک با استفاده از یادگیری ماشینی را مورد بررسی قرار دادند که در آن دو سیستم مدیریت هوشمند هوش مصنوعی بر اساس دو الگوریتم شناخته‌شده ماشین‌آلات یعنی شبکه‌های عصبی مصنوعی و درخت تصمیم‌گیری، توسعه یافته است. این دو سیستم به پنج جفت ارز اصلی (فارکس) با استفاده از سیگنال‌های به‌دست‌آمده از یک استراتژی تجاری فاکتور فروش فایقی که ارائه شده است، شامل یک دوره هفت ساله (۲۰۱۰ الی ۲۰۱۶) می‌شود. شاخص‌های فنی و سری زمانی از نقاط ورودی گذشته، سیستم مدیریت هوشمند را به‌منظور طبقه‌بندی سیگنال‌های تولیدشده از استراتژی تجاری به دو دسته تقسیم می‌کنند: سودآور و غیرسودآور. عرضه اوراق بهادار جدید با استفاده از سیگنال طبقه‌بندی‌شده تنها به‌عنوان سودآور منجر به افزایش سود بیش از ٪۵۰ در مقایسه با اصلی است. نتایج به‌دست‌آمده نشان داد که دو سیستم مدیریت هوشمند هوش مصنوعی می‌توانند بهبود چشمگیر عملکرد سبدهای سرمایه‌گذاری سودآور را به دست آورند و ثابت کنند که استفاده از یادگیری ماشین برای ساخت ابزار مدیریت ریسک بسیار امیدوارکننده است. ملو – رومن و همکاران (Mello-Román *et al*., 2021) در یک مطالعه رویه‌ای را بررسی کردند که ظرفیت تعمیم مدل‌های چند متغیره حداقل مربعات جزئی هسته[1] را با استفاده از الگوریتم‌های ژنتیک بهینه می‌کند. مقادیر پارامتر تابع هسته و تعداد مؤلفه‌هایی را که مقدار ضریب اعتبار متقابل را حداکثر می‌کند، انتخاب کردند. در آزمایش‌های اولیه پیکربندی الگوریتم‌های ژنتیک را اضافه نمودند و یک معیار همگرایی از نظر پراکندگی در برآوردها تعریف کردند. الگوریتم‌های ژنتیک عملکرد خوبی را در کار بهینه‌سازی حداقل مربعات جزئی هسته با راه‌حل‌های همگرا به سمت بهینه جهانی نشان داده است. یانگ و همکاران (Yang *et al*., 2016) در مطالعه‌ای که مورد بررسی قرار دادند، یک الگوریتم جدید رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک[2] را برای رگرسیون داده‌ها پیشنهاد دادند. بردار پشتیبان غیرموازی غیرحساس پارامتریک به طور غیرمستقیم یک جفت توابع پروگزیمال غیرموازی را با یک جفت توابع پروگزیمال غیرموازی متفاوت با حل دو مسئله برنامه‌نویسی درجه دوم با اندازه کوچک‌تر را پیدا می‌کند. با استفاده از توابع جدید ازدست‌دادن پارامتری غیرحساس، بردار پشتیبان

1. Kernel Partial Least Squares (KPLS)
2. Parametric Insensitive Nonparallel Support Vector Regression (PINSVR)

غیرموازی غیرحساس پارامتریک پیشنهادی به طور خودکار یک منطقه غیرحساس پارامتری انعطاف‌پذیر با شکل دلخواه و حداقل اندازه را تنظیم می‌کند تا متغیرهای داده‌شده را برای گرفتن ساختار نزدیک‌ترین متغیر با دقت بیشتری در بر گیرد. نتایج آزمایش در مقایسه با رگرسیون بردار پشتیبان فازی سلسله مراتبی و رگرسیون بردار پشتیبان فازی دوقلو سلسله‌مراتبی و بردار پشتیبان غیرموازی غیرحساس پارامتریک که نشان می‌دهد نه‌تنها عملکرد رگرسیون قابل‌مقایسه را کسب می‌کند؛ بلکه تخمین‌های بهتری را نیز به دست می‌آورد. هرتا و همکاران (Huerta *et al*., 2013) در یک مطالعه، سودآوری یک استراتژی معاملاتی را بر اساس آموزش مدل برای شناسایی سهام با بازده مورد انتظار بالاتر یا پایین‌تر مورد بحث قرار دادند. در این روش، یک طبقه‌بندی‌کننده روی مجموعه‌های توالی تاریخی آموزش داده می‌شود و روی داده‌های آینده آزمایش می‌شود. طبقه‌بندی‌کننده به دلیل سادگی و کارایی آن به‌عنوان یک ماشین بردار پشتیبانی غیرخطی[1] انتخاب شده است. دامنه داده‌ها از سال ۱۹۸۱ تا ۲۰۱۰ است. ماشین بردار پشتیبانی یک بار در ماه آموزش می‌بیند تا با شرایط متغیر بازار سازگار شود. سبدهای سرمایه‌گذاری با رتبه‌بندی سهام با استفاده از خروجی طبقه‌بندی‌کننده شکل می‌گیرند. سهم‌هایی که بالاترین رتبه را دارند برای موقعیت‌های خرید و سهم‌هایی که دارای رتبه پایین هستند برای فروش‌های کوتاه استفاده می‌شوند. طبقه‌بندی جهانی صنعت استاندارد برای ساخت مدلی برای هر بخش استفاده می‌شود؛ به‌طوری‌که در مجموع ۸ سبد سهام بلندمدت برای انرژی، مواد، صنایع، الکترونیک، کالاهای اساسی مصرف‌کننده، مراقبت‌های بهداشتی، مالی و فناوری اطلاعات تشکیل می‌شود؛ نتایج نشان داد که ماشین بردار پشتیبانی خطی کارایی کم‌تری نسبت به ماشین بردار پشتیبانی غیرخطی دارند. کورانی و همکاران (Kurani *et al*., 2023) در یک مطالعه، مقایسه‌ای جامع بین شبکه‌های عصبی مصنوعی و ماشین‌های بردار پشتیبان در پیش‌بینی سهام انجام دادند. در نتایج آزمایش به طور هم‌زمان، ماشین‌های بردار پشتیبان با موفقیت در پیش‌بینی سهام به کار گرفته شده‌اند که دقتی در حدود ٪۶۰ تا ٪۷۰ را برای ماشین‌های بردار پشتیبان ساده می‌دهد که با ترکیب روش‌هایی مانند جنگل تصادفی، نتایج دقیق‌تر الگوریتم ژنتیک، پیش‌بینی بهبود قابل‌توجهی یافت. رگرسیون حداقل مربعات جزئی[2] یک تکنیک رگرسیون محبوب است و معمولاً برای تنظیم چندمتغیره ابزارها استفاده می‌شود. این تکنیک ثابت کرده در شرایطی که تعداد متغیرهای مشاهده‌شده (N) به طور قابل‌توجهی بیش‌تر باشد بهتر عمل می‌کند (De Vries and Ter Braak, 1995).

1. Support Vector Machine (SVM)
2. Partial Least Squares (PLS)

در اکثر موارد الگوریتم هوش مصنوعی برای تحلیل و پیش‌بینی سرمایه‌گذاری در دارایی‌های ثابت و برای بهینه‌سازی و به‌روزرسانی سبد سرمایه‌گذاری شرکت سرمایه‌گذار و برای لحاظ‌کردن شرایط و ترجیحات ویژه شرکت در فرآیند سرمایه‌گذاری و برای انتخاب یک سبد سرمایه‌گذاری بهینه از میان تعداد بسیار زیاد دارایی‌ها به کار گرفته می‌شود؛ لذا با بررسی جنبه‌های نوآوری پژوهش، مشاهده شد که پژوهشی پیرامون سرمایه‌گذاری در دارایی ثابت با روش هوش مصنوعی در بازه زمانی ده‌ساله انجام نشده است و جامعیت در متغیرهای معیارهای نظارتی به کار گرفته شده و استفاده از روش متغیرگزینی ریلیف – اف[1] برای کاهش خطا به‌عنوان نوآوری‌های پژوهش محسوب می‌شود. همچنین استفاده از روش اعتبارسنجی متقابل[2] با محاسبه چهار الگوریتم روش‌های خطی و غیرخطی هوش مصنوعی در کاهش خطاهای پیش‌بینی از دیگر نوآوری‌های پژوهش به شمار می‌آید.

۳- روش‌شناسی پژوهش

پژوهش حاضر به لحاظ هدف، کاربردی و از نظر جمع‌آوری اطلاعات پس‌رویدادی است چرا که از اطلاعات گذشته شرکت‌های نمونه استفاده می‌گردد و از لحاظ روش گردآوری داده‌ها و استنتاج، توصیفی – همبستگی می‌باشد. از نظر منطق اجرا ترکیبی از قیاس و استقرا است. جامعه آماری این پژوهش شامل کلیه شرکت‌های پذیرفته‌شده در بورس اوراق بهادار تهران می‌باشد. نمونه آماری عبارت است از تعداد محدودی از آحاد جامعه آماری که بیان‌کننده ویژگی‌های اصلی جامعه باشد. در این پژوهش برای اینکه نمونه آماری یک نماینده مناسب از جامعه آماری موردنظر باشد، از روش حذف سیستماتیک استفاده شده است. برای این منظور ۴ معیار در نظر گرفته شده و درصورتی‌که شرکتی کلیه معیارها را احراز کرده باشد به‌عنوان نمونه پژوهش انتخاب شده و مابقی حذف می‌شوند. به لحاظ افزایش قابلیت مقایسه سال مالی شرکت منتهی به ۲۹ اسفند بوده و شرکت طی بازه زمانی ۱۳۹۰ تا ۱۳۹۹ سال مالی و نوع فعالیت خود را تغییر نداده باشد. به لحاظ ساختار گزارشگری جداگانه‌ای که شرکت‌های سرمایه‌گذاری و واسطه‌گری مالی (لیزینگ‌ها، بیمه‌ها، هلدینگ‌ها، بانک‌ها و مؤسسات مالی) دارند از نمونه حذف می‌شوند. اطلاعات مالی آن‌ها در بازه زمانی ۱۳۹۰ تا ۱۳۹۹ در دسترس باشد. اطلاعات مالی شرکت‌ها در دسترس باشد. باتوجه‌به شرایط مذکور، تعداد ۱۰۱ شرکت طی مدت ۱۰ سال جهت آزمون فرضیه پژوهش انتخاب شده است. محاسبات داده‌ها با استفاده از نرم‌افزار متلب[3] و ایویوز[4] پردازش می‌شود. این پژوهش به پیش‌بینی سرمایه‌گذاری در دارایی ثابت شرکت‌ها با استفاده از معیار مهم

1. Relief-F
2. K-Fold Cross Validation
3. MATLAB
4. Eviews

نظارتی می‌پردازد که پس از بررسی، میزان سرمایه‌گذاری آینده شرکت در دارایی ثابت به‌عنوان متغیر وابسته، با درصدی خطا به‌عنوان مدل پیش‌بینی خواهد شد. فرآیند اندازه‌گیری متغیرها و شرح چگونگی بررسی آن‌ها از دیدگاه نظارتی در جدول (۱) و جدول (۲) اندازه‌گیری عملیاتی متغیرهای پژوهش نشان داده شده است (He *et al*., 2019).

جدول (۱) اندازه‌گیری عملیاتی متغیرهای مستقل پژوهش (He *et al*., 2019)

Table (1) Operational measurement of independent research variables (He *et al*., 2019)

ردیف	نام متغیر مستقل اولیه	نحوه محاسبات
۱	اندازه هیئت‌مدیره	تعداد اعضای هیئت‌مدیره
۲	استقلال هیئت‌مدیره	نسبت تعداد اعضای غیرموظف به کل اعضای هیئت‌مدیره
۳	تخصص مالی اعضای هیئت‌مدیره	نسبت تعداد اعضای با تخصص مالی به کل اعضای هیئت‌مدیره
۴	اندازه کمیته حسابرسی	تعداد اعضای کمیته حسابرسی شرکت
۵	استقلال کمیته حسابرسی	نسبت تعداد اعضای غیرموظف به کل اعضای کمیته حسابرسی
۶	درصد مالکان نهادی	مجموع مالکیت بیش از ۵ درصد
۷	تغییر مدیرعامل	اگر مدیرعامل شرکت نسبت به دوره قبل تغییر کرده عدد یک و در غیر این صورت صفر تعلق می‌گیرد.
۸	وجود حسابرس داخلی	اگر شرکت واحد حسابرسی داخلی دارد عدد یک و در غیر این صورت عدد صفر استفاده می‌شود.
۹	تخصص مالی اعضای کمیته حسابرسی	نسبت تعداد اعضای با تخصص مالی به کل اعضای کمیته حسابرسی
۱۰	نقش دوگانه مدیرعامل	اگر مدیر عامل شرکت رئیس یا نایب‌رئیس هیئت‌مدیره باشد متغیر مصنوعی یک و در غیر این صورت صفر تعلق می‌گیرد.

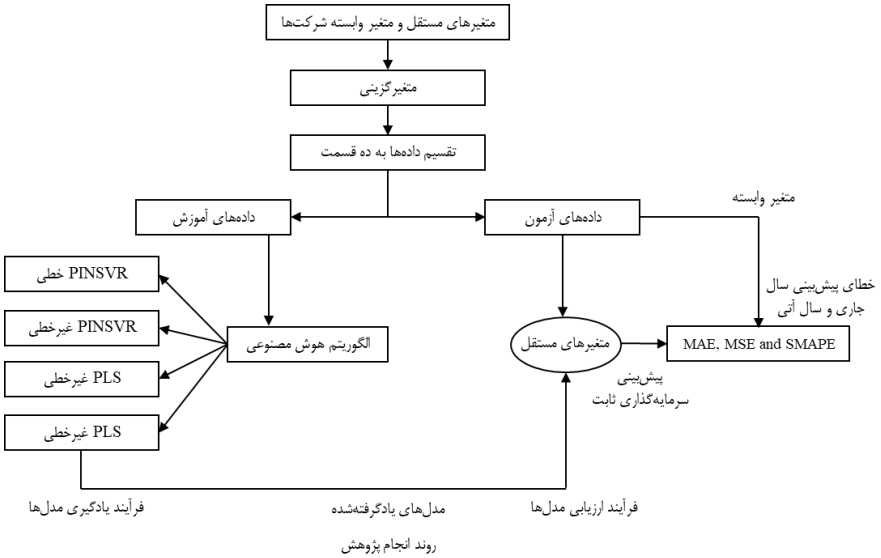
جدول (۲) اندازه‌گیری عملیاتی متغیر وابسته پژوهش (Jahanshad and Khalili, 2013)

Table (2) operational measurement of the dependent variable of the research (Jahanshad and Khalili, 2013)

ردیف	نام متغیر وابسته اولیه	نحوه محاسبات عملیاتی
۱	سرمایه‌گذاری در دارایی‌های ثابت	تفاوت دارایی ثابت اول و آخر دوره تقسیم بر کل دارایی‌ها

متغیرها به روش رگرسیون با استفاده از معیارهای اولیه انتخاب می‌شوند و با استفاده از الگوریتم هوش مصنوعی به پیش‌بینی سرمایه‌گذاری دارایی ثابت پرداخته می‌شود و سپس نرخ میانگین خطاها استخراج و به مقایسه الگوریتم‌ها باتوجه‌به خطاها پرداخته می‌شود و در نهایت روش برتر برای پیشنهاد به استفاده کنندگان معرفی می‌شود. پنج فرآیند (سطح) در روش پیشنهادی، به ترتیب عبارت است از انتخاب داده‌ها، پاکسازی داده‌ها، تقسیم داده‌ها به مجموعه‌های آموزشی و ارزیابی، فرآیند آموزش مدل و ارزیابی مدل آموزش داده‌شده با داده‌های ارزیابی که تا کنون توسط الگوریتم مشاهده نشده است. بعد از حذف داده‌های دورافتاده از مجموعه نمونه‌ها، ویژگی‌های مؤثر در پیش‌بینی سرمایه‌گذاری دارایی ثابت از دیدگاه عملکرد نظارتی بررسی می‌شود و به معرفی دو الگوریتم پیش‌بینی رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک و الگوریتم حداقل مربعات جزئی هسته پرداخته می‌شود. فرآیند پژوهش بر اساس مدل مفهومی شکل (۱) انجام خواهد شد. متغیرهای شرکت‌ها بر اساس روش متغیرگزینی ریلیف – اف انتخاب می‌شوند. پس از انتخاب متغیرها به روش اعتبارسنجی متقابل

ده‌قسمتی، مجموعه نمونه‌های شرکت – سال داده‌ها به گروه داده‌های آموزش و اعتبارسنجی[1] و داده‌های آزمون[2] تقسیم می‌شوند (Alpaydın, 2021). از داده‌های آموزش و اعتبارسنجی به منظور یادگیری پارامترها و ابرپارامترهای مدل‌ها و از داده‌های آزمون جهت ارزیابی پیش‌بینی مدل‌ها استفاده می‌شود. سپس داده‌های آموزشی معیارهای مالی با چهار الگوریتم پیش‌بینی هوش مصنوعی: بردار پشتیبان غیرموازی غیرحساس پارامتریک، خطی و غیرخطی و الگوریتم حداقل مربعات جزئی هسته، خطی و الگوریتم رگرسیون حداقل مربعات جزئی هسته غیرخطی، محاسبه شدند. در مرحله آموزش مدل‌های خطی و غیرخطی، پس از یادگیری، مجدداً همان داده‌های آموزشی بدون متغیر وابسته به آن‌ها داده می‌شود تا مقدار رشد دارایی ثابت را پیش‌بینی کنند؛ سپس قدرت یادگیری و خطای یادگیری این مدل‌ها با استفاده از میانگین ارزیابی خطای درصد قدر مطلق خطای متقارن[3]، میانگین قدر مطلق خطا[4]، میانگین مربعات خطا[5] برای سال جاری و سال آتی گزارش می‌شوند. در تمام این خطاها، هر چه مقدار خطا به صفر نزدیک‌تر باشد، الگوریتم مربوطه قدرتمندتر عمل کرده است



شکل (۱) مدل مفهومی پژوهش

Figure (1) Conceptual model of the research

1. Training and Validation Data
2. Test Data
3. Symmetric Mean Absolute Percentage Error (SMAPE)
4. Mean Absolute Error (MAE)
5. Mean Squared Error (MSE)

روش انتخاب و یافتن اهمیت متغیرهای مستقل مبتنی بر انتخاب متغیرگزینی ریلیف – اف است. این روش از یک راه‌حل آماری برای انتخاب ویژگی استفاده می‌کند. این روش یک الگوریتم مبتنی بر وزن‌دهی به متغیرهای مستقل است که ایده آن از الگوریتم‌های مبتنی بر نمونه الهام گرفته شده است. در الگوریتم این روش مشاهده می‌شود که در ادامه به توضیح آن خواهیم پرداخت. این الگوریتم از میان مجموعه D نمونه آموزشی (به عبارتی مجموعه شرکت – سال به همراه مجموعه متغیر مستقل S که در مجموع تعداد متغیرها N) است، یک زیرمجموعه شرکت انتخاب می‌کند. کاربر تعداد شرکت‌های بدون نمونه در این زیرمجموعه به‌عنوان مقدار از پیش تعریف شده مشخص می‌کند. الگوریتم به‌صورت تصادفی یک شرکت – سال از این زیرمجموعه را به‌عنوان یک نمونه انتخاب می‌کند؛ سپس مبتنی بر ویژگی‌های (متغیرهای مستقل) این نمونه، نزدیک‌ترین برخورد[1] و نزدیک‌ترین شکست[2] را بر اساس تابع ارزیابی فاصله اقلیدسی پیدا می‌کند. نزدیک‌ترین برخورد نمونه‌ای (شرکت – سالی) است که کم‌ترین فاصله اقلیدسی را در میان سایر نمونه‌های هم‌کلاس با نمونه انتخاب‌شده دارد. نزدیک‌ترین شکست نیز شرکت – سالی است که کم‌ترین فاصله اقلیدسی را در میان نمونه‌هایی که هم‌کلاس با نمونه انتخاب‌شده نیستند، دارد. ایده اصلی در این الگوریتم این است که هر چه اختلاف بین اندازه یک ویژگی در شرکت – سال انتخاب شده و نزدیک‌ترین برخورد کم‌تر باشد، این ویژگی بهتر است (Dash and Liu, 1997).

Relief (D, S, No Sample, Threshold)

(1) $T = \emptyset$

(2) *Initialize all weights,* w_i, *to zero.*

(3) *For* $i = 1$ *to NoSample /* Arbitrarily chosen */*

Randomly choose an instance x in D

Finds its nearHit and nearMiss

For $j = 1$ *to N*

$$w_j = w_j - diff(x_j, nearHit_j)^2 + diff(x_j, nearMiss_j)^2$$

(4) *For* j = 1 *to N*

If $w_j \geq$ Threshold

Append feature f_j to T

(5) Return T

شکل (۲) الگوریتم ریلیف (Dash and Liu, 1997)

Figure (2) Relief algorithm (Dash and Liu, 1997)

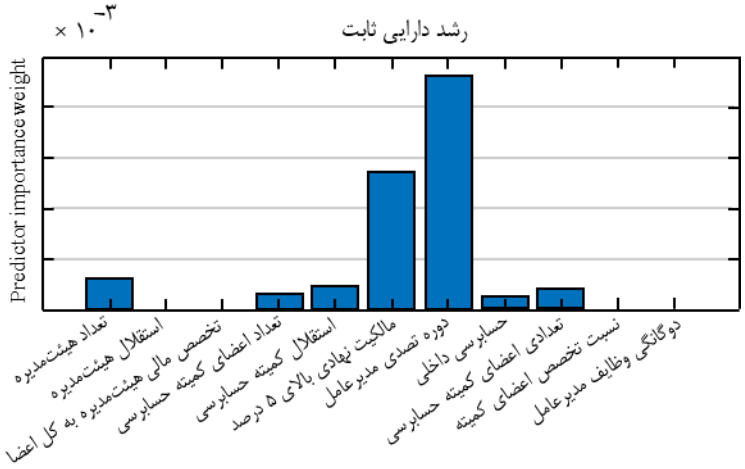
1. Nearest Hit
2. Nearest Miss

در الگوریتم ریلیف هر کدام از متغیرهای مستقل در ابتدا دارای یک وزن W هستند که در شروع الگوریتم مقدار آن برابر صفر است. الگوریتم پس از تعیین نزدیک‌ترین برخورد و نزدیک‌ترین شکست، وزن‌های ویژگی‌ها را به‌روزرسانی می‌کند. این به‌روزرسانی، به این صورت است که مربع اختلاف بین مقدار ویژگی موردنظر در نمونه انتخاب شده و نمونه نزدیک‌ترین برخورد از وزن ویژگی کم می‌شود و مربع اختلاف بین مقدار ویژگی در نمونه انتخاب شده و نزدیک‌ترین شکست، به وزن ویژگی اضافه می‌شود. هر چه مقدار این وزن بزرگ‌تر باشد، ویژگی موردنظر، بهتر می‌تواند شرکت‌های متعلق به یک کلاس را از دیگر شرکت‌ها جدا کند. بعد از تعیین فاصله برای تمام شرکت – سال‌های موجود در مجموعه نمونه‌ها، الگوریتم، ویژگی‌هایی (متغیرهای مستقلی (f)) را که وزن آن‌ها کم‌تر یا مساوی با یک حد آستانه منفی است را حذف می‌کند و سایر ویژگی‌ها را به‌عنوان زیرمجموعه ویژگی جواب (T)، باز می‌گردند. مقدار حد آستانه توسط کاربر تعیین می‌گردد؛ البته ممکن است که به‌صورت اتوماتیک به‌وسیله تابعی از تعداد کل ویژگی‌ها تعیین شود و یا اینکه با سعی و خطا تعیین گردد. ریلیف برای ویژگی‌های نویزی و همبسته خوب عمل می‌کند و پیچیدگی زمانی آن به‌صورت تابعی خطی از تعداد ویژگی‌های داده‌شده و بدون نمونه است. این الگوریتم برای نمونه‌های با ویژگی‌های پیوسته و اسمی هم خوب کار می‌کند. یکی از محدودیت‌های اساسی این الگوریتم این است که ویژگی‌هایی که دارای افزونگی[1] باشند را پیدا نمی‌کند و بنابراین، مجموعه‌های غیربهینه را پیدا می‌کند که دارای افزونگی هستند. این مشکل را می‌توان با یک جستجوی تعیین جامعیت[2] برای زیرمجموعه‌های انتخاب‌شده توسط الگوریتم حل کرد. علاوه بر این، مشکل دیگر الگوریتم این است که با مسائل دوکلاسه خوب کار می‌کند. این محدودیت نیز با الگوریتم ریلیف – اف مرتفع شده است؛ این الگوریتم می‌تواند بر روی مجموعه داده‌ها با مقادیر مفقود و مجموعه داده‌ها با بیش از دو رده داده کار کند. به‌جای انتخاب یکی از نزدیک‌ترین همسایه‌ها که در الگوریتم ریلیف انجام می‌شود، ریلیف – اف تعدادی از همسایگان یک نمونه انتخابی را می‌یابد؛ علاوه بر این، ریلیف – اف از یک تابع متفاوت برای محاسبه وزن ویژگی‌ها استفاده می‌کند. این کار برای مدیریت مجموعه داده‌های ناقص انجام می‌شود که با الگوریتم جدید مشکل داده‌های (نمونه‌های آموزشی غیرکامل) نیز حل شده است (Dash and Liu, 1997).

همچنین، نسخه دیگری از این الگوریتم با نام آرریلیف – اف[3] برای مسائل رگرسیون نیز وجود دارد. بعد از اعمال روش انتخاب ویژگی آرریلیف – اف، این نتایج عملکرد نظارتی به دست آمد. این وزن‌ها صرفاً اهمیت متغیر را نشان می‌دهند و نمی‌توان آن‌ها را به‌عنوان ضرایب رگرسیون خطی در

1. Rdundant
2. Subsequent Exhaustive
3. Regressional Relief-F (RRelief-F)

نظر گرفت. متغیرهای مؤثر در پیش‌بینی سرمایه‌گذاری دارایی ثابت از نقطه نظر معیار نظارتی در شکل (۳) نشان داده شده است. این معیارها عبارتند از: تعداد هیئت‌مدیره، استقلال هیئت‌مدیره، تخصص مالی هیئت‌مدیره به کل اعضا، تعداد اعضای کمیته حسابرسی، استقلال کمیته حسابرسی، مالکیت نهادی بالای ۵ درصد، دوره تصدی مدیرعامل، وجود حسابرس داخلی، نسبت تخصص اعضای کمیته و دوگانگی وظایف مدیرعامل که با روش مذکور انتخاب شده‌اند (Robnik-Šikonja and Kononenko, 1997).



شکل (۳) معیارهای عملکرد نظارتی انتخاب‌شده توسط الگوریتم آرریلیف – اف

Figure (3) Monitoring performance criteria selected by RRelief-F

مدل به الگوریتم بردار پشتیبان غیرموازی غیرحساس پارامتریک و حداقل مربعات جزئی هسته داده می‌شود؛ در ادامه به بررسی این الگوریتم‌ها پرداخته شده است. رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک، برای ارزیابی مدل رگرسیون از طریق تطبیق خودکار ناحیه غیرحساس پارامتریک یک شکل دلخواه و با سایز حداقل باشد؛ به‌طوری‌که دربرگیرنده داده‌های معلوم جهت ذخیره سایز داده و اطلاعات مرزی با دقت بالاتری باشد (Yang *et al.*, 2016).

فرض کنید مجموعه مشاهدات (شرکت – سال) Nتایی به همراه بردار ورودی (متغیرهای مستقل) x موجود باشد که تمام این شرکت – سال‌ها توسط یک ماتریس داده X نشان داده شود؛ به‌طوری‌که nامین ردیف آن را با x_n^T نشان دهیم و بیانگر متغیرهای مستقل شرکت nام باشد و $n = 1, 2, \dots, N$ باشد و y نشان‌دهنده متغیر وابسته یعنی سرمایه‌گذاری دارایی ثابت باشد. الگوریتم رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک به دنبال یافتن توابع خطی پروکسیمال غیرموازی $f_1(x)$ و

$f_2(x)$ به طور هم‌زمان و همچنین، دو تابع خطی پروکسیمال غیرموازی متفاوت $g_1(x)$ و $g_2(x)$ است (Yang *et al.*, 2016)؛ این توابع در رابطه (۱) نشان داده شده‌اند.

$$f_1(x) = w_1^T x + b_1 \text{ و } f_2(x) = w_2^T x + b_2 \quad \text{رابطه (۱)}$$

$$g_1(x) = w_3^T x + b_3 \text{ و } g_{2(x)} = w_4^T x + b_4 \quad \text{رابطه (۲)}$$

در رابطه (۲)، $g_1(x) \geq 0$ و $g_2(x) \geq 0$ است. ابتدا دو تابع ضرر غیرحساس پارامتریک یک‌طرفه[1] به‌صورت رابطه (۳) و رابطه (۴) تعریف می‌شود (Yang *et al.*, 2016).

$$L^{g_1}(x.y.f_1) = \sum_{i=1}^{m} max\{0, -(y_i - f_1(x_i) + g_1(x_i))\} \quad \text{رابطه (۳)}$$

$$L^{g_2}(x.y.f_2) = \sum_{i=1}^{m} max\{0, -(f_2(x_i) - y_i + g_2(x_i))\} \quad \text{رابطه (۴)}$$

بنابراین، ریسک تجربی در این الگوریتم به‌صورت رابطه (۵) و رابطه (۶) تعریف می‌شود.

$$R_{emp}^{g_1}[f_1] = \sum_{i=1}^{m} max\left\{0, (y_i - f_1(x_i))^2\right\} + c_1 \sum_{i=1}^{m} max\{0, -(y_i - f_1(x_i) + g_1(x_i))\} \quad \text{رابطه (۵)}$$

$$R_{emp}^{g_2}[f_2] = \sum_{i=1}^{m} max\{0, (f_2(x_i) - y_i)^2\} + c_2 \sum_{i=1}^{m} max\{0, -(f_2(x_i) - y_i + g_2(x_i))\} \quad \text{رابطه (۶)}$$

در رابطه (۵) و رابطه (۶)، $c_1 > 0$ و $c_2 > 0$ پارامترهای آن هستند. مسئله‌های بهینه‌سازی اولیه رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک به‌صورت رابطه (۷) و رابطه (۸) تعریف می‌شود (Yang *et al.*, 2016).

$$\min_{w_1, w_3, b_1, b_3, \xi} \frac{1}{2} c_3 (w_1^T w_1 + b_1^2 + w_3^T w_3 + b_3^2) + \frac{1}{2} \xi^{*T} \xi^* + c_1 e^T \xi$$

$$s.t. \begin{cases} Y - (Aw_1 + eb_1) = \xi^* \\ Aw_3 + eb_3 \geq 0 \\ Y - (Aw_1 + eb_1) \geq -(Aw_3 + eb_3) - \xi, \xi \geq 0 \end{cases} \quad \text{رابطه (۷)}$$

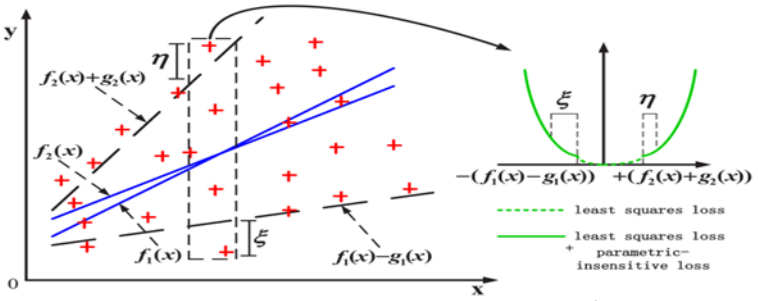
$$\min_{w_2, w_4, b_2, b_4, \eta} \frac{1}{2} c_4 (w_2^T w_2 + b_2^2 + w_4^T w_4 + b_4^2) + \frac{1}{2} \eta^{*T} \eta^* + c_2 e^T \eta \quad \text{رابطه (۸)}$$

1. One Side Parametric Insensitive Loss Function (OSPILF)

$$s.t.\ \begin{cases} (Aw_2 + eb_2) - Y = \eta^* \\ Aw_4 + eb_4 \geq 0 \\ (Aw_2 + eb_2) - Y \geq -(Aw_4 + eb_4) - \eta, \eta \geq 0 \end{cases}$$

در رابطه (۷) و رابطه (۸)، c_1، c_2، c_3 و c_4 پارامترهای ورودی مسئله هستند. اکنون، توضیح هندسی مسئله بهینه سازی را بیان می‌کنیم؛ یک نمونه ساده دوبُعدی از اجرای الگوریتم رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک به همراه تابع ضرر در شکل (۴) نشان داده شده است. ترم اول در تابع هدف حداقل‌کردن عبارت $\frac{1}{2}c_3\left(w_1^T w_1 + b_1^2 + w_3^T w_3 + b_3^2\right)$ است که پیچیدگی مدل $f_1(x)$ و $g_1(x)$ را کنترل می‌کند؛ به‌عبارت‌دیگر، باعث می‌گردد که $f_1(x)$ تا حد ممکن به داده‌ها بچسبد و $g_1(x)$ ثبت‌کننده کران پایین ساختار داده‌ها باشد. علاوه بر این، ریسک ساختاری در رابطه به دلیل وجود ترم تنظیم[1] $\frac{1}{2}c_3\left(w_1^T w_1 + b_1^2\right)$ حداقل می‌شود. ترم دوم در تابع هدف باتوجه‌به قید $Y - (Aw_1 + eb_1) = \xi^*$ نشان‌دهنده تابع ضرر مجذور مربعات خطا[2] بین مقدار تابع تصمیم $f_1(x) = w_1^T x + b_1$ و متغیر وابسته برای نمونه‌های x است که این برچسب با Y نشان داده شده است؛ بنابراین، حداقل‌کردن این ترم به معنی تقریب بهتر متغیر وابسته توسط رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک است. قید اول نشان‌دهنده این است که $g_1(x) \geq 0$ باشد. هدف دومین قید نامساوی تابع هدف این است که دو تابع $f_1(x)$ و $g_1(x)$ در فاصله حداقلی از هم باشند به‌طوری‌که نمونه‌های آموزشی باید از $f_1(x)$، حداقل از $g_1(x)$ بزرگ‌تر باشند. بردار کمکی[3] ξ برای اندازه‌گیری خطا است؛ هر زمان که نمونه‌ها خط $g(x)$ را رد می‌کنند. ترم سوم تابع هدف مجموع متغیرهای ξ را حداقل می‌کند یعنی الگوریتم سعی می‌کند که نمونه‌ها از خط $g_1(x)$ تاحدامکان عبور نکنند و همه یک طرف این خط قرار گیرند. برای مسئله بهینه‌سازی توضیح مشابهی وجود دارد (Yang *et al.*, 2016).

1. Regularization Term
2. Lease Squares
3. Slack Vector



شکل (۴) تفسیر هندسی دوبُعدی و تنظیمات تابع هدف الگوریتم رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک (Yang *et al*., 2016)

Figure (4) Two dimensional geometric interpretation and objective function settings of the parametric insensitive nonparallel support vector regression algorithm (Yang *et al*., 2016)

با استفاده از تابع لاگرانژ و بررسی شرایط کاروش کون تاکر[1] تابع دوگان دو مسئله بهینه به‌صورت مسئله دوگان در رابطه (۹) و رابطه (۱۰) به دست می‌آید.

$$\min_{\alpha,\beta} \frac{1}{2c_3}\alpha^T GG^T\alpha + \frac{1}{2}\beta^T\left(G(G^TG+c_3I_1)^{-1}G^T + \frac{1}{c_3}G^TG\right)\beta + \frac{1}{c_3}\alpha^T GG^T\beta - Y^T(G(G^TG+c_3I_1)^{-1}G^T - I_2)\beta$$

$$s.t.\ \begin{cases}\alpha \geq 0\\ 0 \leq \beta^* \leq c_1 e\end{cases}$$

رابطه (۹)

$$\min_{\alpha^*,\beta^*} \frac{1}{2c_4}\alpha^{*T} GG^T\alpha^* + \frac{1}{2}\beta^{*T}\left(G(G^TG+c_4I_1)^{-1}G^T + \frac{1}{c_4}G^TG\right)\beta^* + \frac{1}{c_4}\alpha^{*T} GG^T\beta^* - Y^T(G(G^TG+c_4I_1)^{-1}G^T - I_2)\beta^*$$

$$s.t.\ \begin{cases}\alpha^* \geq 0\\ 0 \leq \beta^* \leq c_2 e\end{cases}$$

رابطه (۱۰)

در رابطه (۹) و رابطه (۱۰)، $G = [A \quad e]$ است. از حل رابطه (۹)، ضرایب لاگرانژ α و β به دست می‌آید و از قرار دادن آن در رابطه (۱۱) و رابطه (۱۲) مقدار پارامترهای خط $f_1(x)$ و $g_1(x)$ به دست می‌آید.

$$u_1 = (G^TG + c_3I_1)^{-1}G^T(Y - \beta)$$

رابطه (۱۱)

$$u_3 = \frac{1}{c_3}G^T(\alpha + \beta)$$

رابطه (۱۲)

1. Karush Kuhn Tucker (KKT)

در رابطه (۱۱) و رابطه (۱۲)، $u_1 = [w_1^T \; b_1^T]^T$ و $u_3 = [w_3^T \; b_3^T]^T$ است. از حل بهینه‌سازی ضرایب لاگرانژ α^* و β^* به دست می‌آید و از قرار دادن آن در رابطه (۱۳) و رابطه (۱۴) مقدار پارامترهای خط $f_2(x)$ و $g_2(x)$ به دست می‌آید.

$$u_2 = (G^T G + c_4 I_1)^{-1} G^T (Y - \beta^*) \quad \text{رابطه (۱۳)}$$

$$u_4 = \frac{1}{c_4} G^T (\alpha^* + \beta^*) \quad \text{رابطه (۱۴)}$$

حال، تابع تصمیم خطی این الگوریتم به‌صورت رابطه (۱۵) به دست می‌آید.

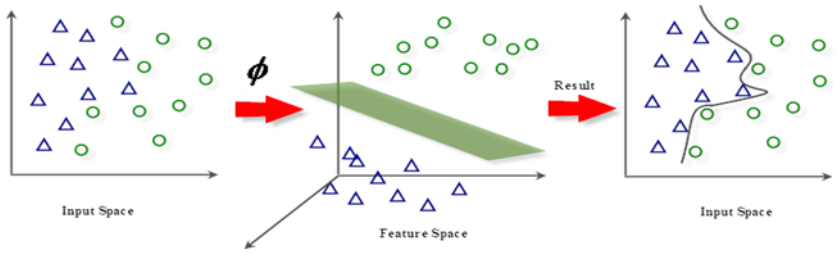
$$f(x) = \frac{1}{2}\big(f_1(x) + f_2(x)\big) = \frac{1}{2}(w_1 + w_2)^T x + \frac{1}{2}(b_1 + b_2) \quad \text{رابطه (۱۵)}$$

کران بالا و پایین مدل رگرسیون به‌صورت رابطه (۱۶) و رابطه (۱۷) محاسبه می‌شود.

$$f_1(x) - g_1(x) = (w_1 - w_3)^T x + b_1 - b_3 \quad \text{رابطه (۱۶)}$$

$$f_2(x) + g_2(x) = (w_2 + w_4)^T x + b_2 + b_4 \quad \text{رابطه (۱۷)}$$

الگوریتم رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک خطی برای حالت غیرخطی با استفاده از حقه هسته بسط داده می‌شود. داده ورودی به فضای ویژگی با ابعاد بالا با استفاده از توابع هسته غیرخطی نگاشت داده می‌شود. در فضای ویژگی یک تابع رگرسیون خطی متناظر با تابع رگرسیونی غیرخطی در فضای ورودی است که در شکل (۵) این موضوع نشان داده شده است (Yang *et al.*, 2016).



شکل (۵) مفهوم فضای ورودی، تابع نگاشت و فضای ویژگی (Yang *et al.*, 2016)

Figure (5) Concept of input space, mapping function and feature space (Yang *et al.*, 2016)

رگرسیون بردار پشتیبانی غیرموازی غیرحساس پارامتریک نیز مشابه حالت خطی، توابع غیرخطی پروکسیمال غیرموازی $f_1(x)$ و $f_2(x)$ و دو تابع مختلف غیرخطی پروکسیمال غیرموازی $g_1(x)$ و $g_2(x)$ به‌صورت رابطه (۱۸) و رابطه (۱۹) در نظر گرفته شده است (Yang *et al.*, 2016).

$$f_1(x) = k(x^T.A^T)w_1 + b_1 \text{ و } f_2(x) = k(x^T.A^T)w_2 + b_2$$ رابطه (۱۸)

$$g_1(x) = k(x^T.A^T)w_3 + b_3 \text{ و } g_{2(x)} = k(x^T.A^T)w_4 + b_4$$ رابطه (۱۹)

در رابطه (۱۸) و رابطه (۱۹)، k تابع هسته است و $g_1(x) \geq 0$ و $g_2(x) \geq 0$ است. مسئله اولیه رگرسیون بردار پشتیبانی غیرموازی غیرحساس پارامتریک غیرخطی به‌صورت توابع مطرح‌شده در بخش رگرسیون بردار پشتیبانی غیرموازی غیرحساس پارامتریک مجدداً تکرار می‌شود.

حداقل مربعات جزئی روشی برای مدل‌سازی رابطه خطی بین مجموعه‌ای از متغیرهای خروجی (پاسخ‌ها) $\{y_i\}_{i=1}^{n} \in R^L$ و مجموعه‌ای از متغیرهای ورودی $\{x_i\}_{i=1}^{n} \in R^N$ (رگرسورها) است. در مرحله اول، حداقل مربعات جزئی متغیرهای نهان غیرهمبسته ایجاد می‌کند که ترکیبی خطی از رگرسورهای اصلی هستند (Rosipal and Trejo, 2001). نکته اساسی این روش این است که وزن‌ها به منظور تعیین ترکیبات خطی رگرسورهای اصلی که متناسب با کوواریانس در بین متغیرهای ورودی و خروجی هستند، استفاده می‌شود. الگوریتم حداقل مربعات در زیرمجموعه متغیرهای پنهان استخراج‌شده انجام می‌شود؛ این منجر به تخمین واریانس مغرضانه اما کمتری از ضرایب رگرسیون در مقایسه با رگرسیون حداقل مربعات اصلی[1] می‌شود. رگرسیون حداقل مربعات جزئی یک فرآیند تکراری است. به‌عنوان‌مثال، پس از استخراج یک مؤلفه، الگوریتم مجدداً با استفاده از ماتریس‌های X و Y دوباره شروع می‌شود؛ بنابراین می‌توانیم ترتیبی از مدل‌ها را تا زمانی که به رتبه X برسیم، به دست آوریم؛ اگر چه، در عمل از روش اعتبارسنجی متقابل معمولاً برای جلوگیری از کم‌برازش[2] یا بیش‌برازش[3] که ناشی از استفاده از مدل‌هایی با ابعاد خیلی کوچک یا خیلی بزرگ است، استفاده می‌شود. پس از استخراج مؤلفه‌های p می‌توان ماتریس‌های T (n × p)، U، W (N × p) و ماتریس C (L × p) که ستون‌های آن از بردارهای $\{t_i\}_{i=1}^{p}$، $\{u_i\}_{i=1}^{p}$، $\{w_i\}_{i=1}^{p}$ و $\{c_i\}_{i=1}^{p}$ ایجاد شده است. مدل رگرسیون حداقل مربعات جزئی را می‌توان به‌صورت ماتریس به‌صورت رابطه (۲۰) نوشت (Rosipal and Trejo, 2001).

$$Y = XB + F$$ رابطه (۲۰)

در رابطه (۲۰)، B یک ماتریس (N × L) از ضرایب رگرسیون و F یک ماتریس باقی‌مانده (n × L) است؛ رابطه (۲۰)، یک رابطه خطی است که در سایر مدل‌های رگرسیون استفاده می‌شود.

1. Ordinary Least Squares (OLS)
2. Underfitting
3. Overfitting

در الگوریتم حداقل مربعات جزئی هسته، فرض می‌شود که یک نگاشت غیرخطی از متغیرهای ورودی $\{x_i\}_{i=1}^{n}$ به یک فضای ویژگی F در نظر گرفته شده؛ به‌عنوان‌مثال، نگاشت ویژگی برداری $\Phi: x_i \in R^N \rightarrow \Phi(x_i) \in F$. هدف ما ساخت یک مدل رگرسیون حداقل مربعات جزئی خطی در فضای F می‌باشد. با اصلاح الگوریتم خطی، الگوریتم غیرخطی به‌صورت رابطه (۲۱) تغییر می‌کند. با به‌کارگیری حقه هسته یعنی $\Phi(x_i)^T\Phi(x_j) = K(x_i, x_j)$ می‌توان مشاهده کرد که $\Phi^T\Phi$ نمایش ماتریس گرام هسته K است. به‌جای استفاده از نگاشت داده‌ها به فضای با ابعاد بالای ویژگی، می‌توان محاسبات را در فضای ورودی انجام داد. با تعریف $\Phi^T\Phi = K$ داریم (Rosipal and Trejo, 2001).

$$K \leftarrow (I - tt^T)K(I - tt^T) = K - tt^TK - Ktt^T + tt^TKtt^T$$ رابطه (۲۱)

در رابطه (۲۱)، I یک ماتریس همانی nبعدی است. مشابه حالت خطی می‌توان ماتریس ضرایب رگرسون B را به دست آورد. فقط لازم به ذکر است که باتوجه‌به ناشناس‌بودن و همچنین بُعد بالای نگاشت ویژگی نمی‌توان عملاً این ماتریس را محاسبه نمود؛ ولی به‌جای آن با به‌کارگیری حقه هسته، می‌توان خروجی سیستم را بدون نیاز به دانستن نگاشت ویژگی به کمک تابع هسته و ماتریس گرام محاسبه نمود (Rosipal and Trejo, 2001).

$$\hat{Y} = \Phi B = KU(T^TKU)^{-1}T^TY = TT^TY$$

$$R = \Phi^TU(T^TKU)^{-1} \text{ و } T = \Phi B$$ رابطه (۲۲)

٤- تجزیه‌وتحلیل داده‌ها

آمار توصیفی صرفاً به توصیف جامعه یا نمونه می‌پردازد و هدف از آن محاسبه پارامترهای جامعه یا نمونه پژوهش است (Azar and Momeni, 2009). به‌منظور بررسی و تجزیه‌وتحلیل اولیه داده‌ها، ابتدا اطلاعات مربوط به آماره‌های توصیفی متغیرهای وابسته و توضیحی موردمطالعه در این پژوهش در جدول (۳) ارائه شده است تا شمایی کلی از داده‌هایی که در این پژوهش مورد تحلیل واقع شده‌اند، به دست آید.

جدول (۳) آمار توصیفی متغیرهای مورد استفاده در مطالعه

Table (3) Descriptive statistics of the variables used in the study

متغیر	میانگین	میانه	حداکثر	حداقل	انحراف معیار	چولگی	کشیدگی
تعداد (اندازه) هیئت‌مدیره	۵/۰۳	۵	۷	۵	۰/۲۶	۷/۵۲	۵۴/۷۱
استقلال هیئت‌مدیره	۰/۶۶	۰/۶۰	۱/۰۰	۰/۰۰	۰/۱۹	−۰/۱۹	۰/۳۱
تخصص مالی هیئت‌مدیره	۰/۱۴	۰/۲۰	۰/۶۰	۰/۰۰	۰/۱۳	۰/۶۶	۰/۳۴
اعضای کمیته حسابرسی	۲/۴۰	۳/۰۰	۵/۰۰	۰/۰۰۰	۱/۳۵	۰/۹۰	۰/۲۳

استقلال کمیته حسابرسی	۰/۵۶	۰/۶۷	۱/۰۰	۰/۰۰	۰/۳۵	-۰/۶۲	-۰/۸۶
مالکیت نهادی	۷۰/۸۹	۷۴/۹۳	۱۰۰/۰۰	۰/۰۰	۱۹/۳۱	-۱/۱۵	۱/۲۲
دوره تصدی مدیرعامل	۳/۴۶	۲/۰۰	۱۹/۰۰	۱/۰۰	۲/۹۹	۱/۸۱	۳/۵۱
حسابرس داخلی	۰/۴۹	۰/۰۰	۱/۰۰	۰/۰۰	۰/۵۰	۰/۰۶	-۲/۰۰
نسبت تخصص اعضای کمیته	۰/۵۵	۰/۶۷	۱/۰۰	۰/۰۰	۰/۳۸	-۰/۳۲	-۱/۳۳
دوگانگی وظایف مدیرعامل	۰/۸۵	۱/۰۰	۱/۰۰	۰/۰۰	۰/۳۶	-۱/۹۳	۱/۷۱
دارایی ثابت	۰/۵۷	۰/۰۶	۴۵/۷۹	-۰/۸۲	۲/۷۳	۱۱/۲۲	۱۵۷/۰۹

در تفسیر نتایج آمار توصیفی مربوط به میانگین ۱۴/۳۷ برای اندازه هیئت‌مدیره می‌توان بیان کرد که تقریباً نمونه آماری شامل شرکت‌های بزرگ و کوچک بوده است و باتوجه‌به میانگین اندازه هیئت‌مدیره، ۵/۰۳ می‌توان بیان کرد بیش‌تر شرکت‌ها دارای ۵ عضو هیئت‌مدیره بوده‌اند و باتوجه‌به اینکه استقلال کمیته حسابرسی و هیئت‌مدیره بیش‌تر از ۰/۵ است می‌توان بیان کرد که اکثر شرکت‌ها از اعضای غیرموظف در ترکیب هیئت‌مدیره و کمیته حسابرسی استفاده کرده‌اند. تنظیم و طبقه‌بندی داده‌ها، نمایش ترسیمی و محاسبه مقادیری از قبیل میانگین، میانه و... می‌باشد که حاکی از مشخصات یکایک اعضای جامعه مورد بحث است. در آمار توصیفی اطلاعات حاصل از یک گروه، همان گروه را توصیف می‌کند و اطلاعات به‌دست‌آمده به دسته‌جات مشابه تعمیم داده نمی‌شود. انحراف معیار نیز به‌عنوان یکی از شاخص‌های پراکندگی نشان می‌دهد به‌طور میانگین داده‌ها چه مقدار از مقدار متوسط فاصله دارند. بیش‌ترین پراکندگی از میانگین مربوط به مالکیت نهادی و کم‌ترین مربوط به تخصص مالی هیئت‌مدیره می‌باشد. آمار توصیفی متغیر کیفی پژوهش به‌صورت توزیع فراوانی به شرح جدول (۴) ارائه می‌گردد.

جدول (۴) توزیع فراوانی در مطالعه

Table (4) Frequency distribution in the study

متغیر	تعداد	فراوانی نسبی	فراوانی تجمعی
دوگانگی مدیرعامل	۸۵۵	۸۴/۶۵	۸۴/۶۵
عدم دوگانگی مدیرعامل	۱۵۵	۱۵/۳۵	۱۰۰/۰۰
دارای حسابرس داخلی	۴۹۰	۴۸/۵۱	۴۸/۵۱
فاقد حسابرس داخلی	۱۵۵	۵۱/۴۹	۱۰۰/۰۰

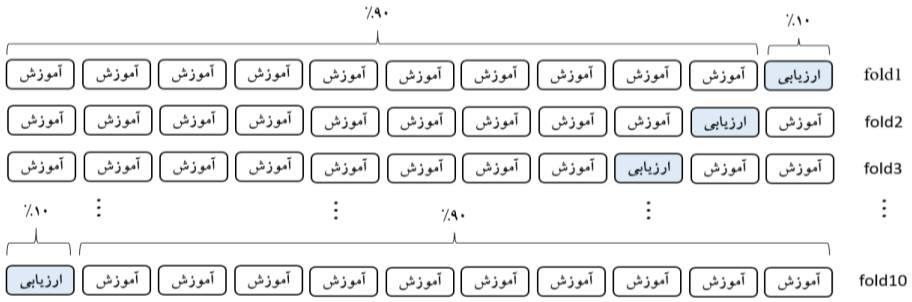
نتایج توزیع فراوانی در جدول (۴) نشان می‌دهد که حدود ۸۴/۶۵ درصد از سال - شرکت‌های موردمطالعه دارای دوگانگی مدیرعامل یعنی مدیرعامل شرکت رئیس یا نایب‌رئیس هیئت‌مدیره نیز می‌باشد و ۱۵/۳۵ درصد عدم دوگانگی مدیرعامل می‌باشد و بدین معناست که مدیرعامل صرفاً سمت

مدیرعاملی شرکت را دارا می‌باشد و همچنین حدود ۴۸/۵۱ درصد از سال – شرکت‌های موردمطالعه دارای حسابرس داخلی و ۵۱/۴۹ درصد فاقد حسابرس داخلی می‌باشند.

در روش تقسیم‌بندی داده‌ها و معیارهای ارزیابی مدل‌ها، برای آموزش مدل‌های خطی و غیرخطی رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک و حداقل مربعات جزئی هسته، در ابتدا باید مجموعه نمونه‌های شرکت – سال باید با روش اعتبارسنجی متقابل به گروه داده‌های آموزش و اعتبارسنجی و داده‌های آزمون تقسیم شوند (Alpaydın, 2021).

از داده‌های آموزش و اعتبارسنجی به منظور یادگیری پارامترها و ابرپارامترهای مدل‌ها و از داده‌های آزمون جهت ارزیابی پیش‌بینی مدل‌ها استفاده می‌شود. به منظور ارزیابی بهتر و کارآمدتر پیش‌بینی مدل‌ها از روش اعتبارسنجی متقابل ۱۰تکه‌ای در این پژوهش استفاده شده است. این نمونه‌ها مطابق شکل (۶) به ۱۰ تکرار جابه‌جایی نمونه‌های ارزیابی تقسیم شدند. این تقسیم‌بندی به این صورت است که از کل نمونه تعداد یک دهم نمونه به‌عنوان داده ارزیابی در تکرار اول انتخاب شد و بقیه داده‌ها در ۹ قسمت به‌عنوان داده‌های آموزش اعتبارسنجی انتخاب شدند. مجموع ۹ قسمت و ۱ قسمت برای ما روش اعتبارسنجی متقابل ۱۰تایی را تشکیل می‌دهد. بدون از دست دادن عمومیت، برای تکرار دوم، مطابق داده‌های قسمت دوم به‌عنوان داده‌های ارزیابی و سایر قسمت‌ها به‌عنوان داده‌های آموزش اعتبارسنجی انتخاب شدند. این تکرارها با همین ترتیب تا ۱۰ مرتبه تکرار می‌شود. به کمک روش اعتبارسنجی متقابل ۱۰تایی و با استفاده از مجموعه داده‌های آموزش اعتبارسنجی برای هر کدام از مدل‌های خطی و غیرخطی رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک و حداقل مربعات جزئی هسته در هر تکرار یک مدل یاد گرفته می‌شود. به عبارت ساده‌تر، در تکرار اول، هر کدام از مدل‌ها به‌صورت مستقل از دیگر به کمک ۹ قسمت آموزش اعتبارسنجی پارامترها و ابرپارامترهای خود را مبتنی بر الگوریتم‌ها یادگیری انجام می‌دهند و مدل خود را می‌سازند. تا اینجا چهار مدل ساخته می‌شود. برای بررسی اینکه چقدر این مدل‌ها داده‌های آموزش اعتبارسنجی را یاد گرفته‌اند، دقیقاً همان ۹ قسمت به مدل‌های آموزش داده‌شده وارد می‌شود تا سرمایه‌گذاری دارایی ثابت را پیش‌بینی کنند. حال با استفاده از معیارهای ارزیابی که در ادامه معرفی می‌شوند مانند میانگین مربعات خطا، میزان خطای پیش‌بینی و مقدار واقعی متغیر دارایی ثابت اندازه‌گیری می‌شوند. این عدد به‌عنوان خطای یادگیری تکرار اول ذخیره می‌شود. این کار برای سایر قسمت‌ها تکرار می‌شود و ۱۰ خطای میانگین مربعات خطا به دست می‌آید. میانگین این خطاها به‌عنوان خطای مرحله یادگیری هر مدل ذخیره می‌شود. معمولاً نرخ خطا روی داده‌های آموزش کم‌تر از نرخ خطا روی داده‌های ارزیابی است که در فرآیند یادگیری دیده نشده‌اند؛ بنابراین، نمی‌توان از خطای یادگیری برای مقایسه دو الگوریتم استفاده

نمود؛ زیرا ممکن است پدیده بیش‌بزارش اتفاق بیفتد. پدیده کم‌بزارش یعنی مدل‌ها، داده‌های آموزش اعتبارسنجی را به‌خوبی پیش‌بینی کنند؛ ولی داده‌های دنیای واقعی که هنوز ندیده‌اند را خیلی بد و با اختلاف زیاد پیش‌بینی کنند؛ بنابراین، علاوه بر مجموعه داده‌های آموزش، مجموعه‌ای از داده‌ها برای آزمون مورد نیاز است. این داده‌ها تا کنون توسط مدل‌ها دیده نشده‌اند و در هر تکرار داده‌ها با تکرارهای دیگر متفاوت هستند. در هر تکرار مدل متناظر آن پیش‌بینی داده‌های ارزیابی را که تا کنون ندیده است انجام می‌دهد و معیارهای ازریابی مقدار پیش‌بینی‌شده و واقعی سرمایه‌گذاری دارایی ثابت محاسبه می‌شود و در نهایت میانگین این خطاها به‌عنوان خطای پیش‌بینی هر مدل ذخیره می‌شود. حال اگر میانگین خطای مرحله یادگیری و خطای مرحله ارزیابی نزدیک هم باشد، یعنی پدیده بیش‌بزارش اتفاق نیفتاده است (Alpaydın, 2021).



شکل (۶) مراحل انتخاب دو مجموعه داده‌های آموزش و آزمون با اعتبارسنجی متقابل ۱۰قسمتی

Figure (6) The selection steps of two sets of training and test data with 10 fold cross validation

پس از تقسیم شرکت سال‌ها به دو گروه داده‌های آموزش اعتبارسنجی و آزمون با استفاده از اعتبارسنجی متقابل ۱۰تایی برای ارزیابی مدل‌های خطی و غیرخطی از سه معیار ارزیابی با نام‌های میانگین درصد قدرمطلق خطای متقارن، میانگین قدر مطلق خطا، میانگین مربعات خطا استفاده شده است که با استفاده از رابطه (۲۳)، رابطه (۲۴) و رابطه (۲۵) محاسبه می‌گردند (Alpaydın, 2021).

$$SMAPE = \frac{1}{n}\frac{\sum_{i=1}^{n}|d_i - y_i|}{\sum_{i=1}^{n}(d_i + y_i)}$$ رابطه (۲۳)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - d_i|$$ رابطه (۲۴)

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - d_i)^2$$ رابطه (۲۵)

در رابطه (۲۳)، رابطه (۲۴) و رابطه (۲۵)، y_i و d_i به ترتیب متغیر وابسته واقعی و متغیر وابسته پیش‌بینی‌شده توسط مدل‌ها برای شرکت – سال iام است و n تعداد شرکت – سال‌ها (در مرحله آموزش یا مرحله آزمون) را نشان می‌دهد. جزئیات خطاهای به‌دست‌آمده در هر تکرار برای هر الگوریتم با معیارهای مختلف نشان داده شده است. در جدول (۵)، جدول (۶)، جدول (۷)، جدول (۸)، جدول (۹) و جدول (۱۰) میانگین خطا و انحراف از معیار هر سطر در انتهای آن نوشته شده است.

جدول (۵) جزئیات میانگین قدر مطلق خطای روش اعتبارسنجی متقابل عملکرد نظارتی در مرحله آموزش

Table (5) Details of the mean absolute error of the mutual validation method of the supervisory performance in the training phase

سال	معیار MSE	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	میانگین انحراف معیار
سال جاری	Linear PINSVR	۰/۰۹۵	۰/۰۹۶	۰/۰۹۴	۰/۰۹۷	۰/۰۹۶	۰/۰۹۶	۰/۰۹۷	۰/۰۹۷	۰/۰۹۷	۰/۰۹۶	۰/۰۰۱ ± ۰/۰۹۶
	Nonlinear PINSVR	۰/۰۸۸	۰/۰۸۹	۰/۰۸۷	۰/۰۹۰	۰/۰۸۹	۰/۰۸۹	۰/۰۹۰	۰/۰۹۰	۰/۰۸۹	۰/۰۸۸	۰/۰۰۱ ± ۰/۰۸۹
	Linear PLS	۰/۰۹۵	۰/۰۹۵	۰/۰۹۴	۰/۰۹۷	۰/۰۹۶	۰/۰۹۷	۰/۰۹۶	۰/۰۹۶	۰/۰۹۸	۰/۰۹۵	۰/۰۰۱ ± ۰/۰۹۶
	Kernel PLS	۰/۰۸۲	۰/۰۸۳	۰/۰۸۲	۰/۰۸۵	۰/۰۸۲	۰/۰۸۴	۰/۰۸۳	۰/۰۸۳	۰/۰۸۴	۰/۰۸۲	۰/۰۰۱ ± ۰/۰۸۳
سال آتی	Linear PINSVR	۰/۰۹۳	۰/۰۹۷	۰/۰۹۳	۰/۰۹۵	۰/۰۹۴	۰/۰۹۴	۰/۰۹۵	۰/۰۹۵	۰/۰۹۴	۰/۰۹۵	۰/۰۰۱ ± ۰/۰۹۴
	Nonlinear PINSVR	۰/۰۸۵	۰/۰۸۸	۰/۰۸۴	۰/۰۸۶	۰/۰۸۵	۰/۰۸۶	۰/۰۸۶	۰/۰۸۶	۰/۰۸۶	۰/۰۸۷	۰/۰۰۱ ± ۰/۰۸۶
	Linear PLS	۰/۰۹۳	۰/۰۹۷	۰/۰۹۲	۰/۰۹۵	۰/۰۹۴	۰/۰۹۴	۰/۰۹۶	۰/۰۹۶	۰/۰۹۴	۰/۰۹۵	۰/۰۰۱ ± ۰/۰۹۴
	Kernel PLS	۰/۰۸۰	۰/۰۸۳	۰/۰۷۹	۰/۰۸۰	۰/۰۸۰	۰/۰۸۰	۰/۰۸۰	۰/۰۸۱	۰/۰۸۰	۰/۰۸۱	۰/۰۰۱ ± ۰/۰۸۱

جدول (۶) جزئیات میانگین مربعات خطای روش اعتبارسنجی متقابل عملکرد نظارتی در مرحله آموزش

Table (6) Details of the mean squared error of the mutual validation method of the supervisory performance in the training phase

سال	معیار MSE	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	میانگین انحراف معیار
سال جاری	Linear PINSVR	۰/۰۱۷	۰/۰۱۷	۰/۰۱۷	۰/۰۱۸	۰/۰۱۷	۰/۰۱۸	۰/۰۱۷	۰/۰۱۸	۰/۰۱۸	۰/۰۱۷	۰ ± ۰/۰۱۷
	Nonlinear PINSVR	۰/۰۱۵	۰/۰۱۵	۰/۰۱۴	۰/۰۱۶	۰/۰۱۵	۰/۰۱۵	۰/۰۱۵	۰/۰۱۶	۰/۰۱۵	۰/۰۱۵	۰ ± ۰/۰۱۵
	Linear PLS	۰/۰۱۷	۰/۰۱۷	۰/۰۱۷	۰/۰۱۸	۰/۰۱۷	۰/۰۱۸	۰/۰۱۷	۰/۰۱۸	۰/۰۱۸	۰/۰۱۷	۰ ± ۰/۰۱۷
	Kernel PLS	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰/۰۱۴	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰/۰۱۳	۰ ± ۰/۰۱۳
سال آتی	Linear PINSVR	۰/۰۱۶	۰/۰۱۷	۰/۰۱۶	۰/۰۱۷	۰/۰۱۶	۰/۰۱۶	۰/۰۱۶	۰/۰۱۷	۰/۰۱۷	۰/۰۱۶	۰ ± ۰/۰۱۶
	Nonlinear PINSVR	۰/۰۱۴	۰/۰۱۵	۰/۰۱۳	۰/۰۱۴	۰/۰۱۴	۰/۰۱۴	۰/۰۱۴	۰/۰۱۴	۰/۰۱۴	۰/۰۱۵	۰ ± ۰/۰۱۴
	Linear PLS	۰/۰۱۶	۰/۰۱۷	۰/۰۱۶	۰/۰۱۷	۰/۰۱۶	۰/۰۱۶	۰/۰۱۶	۰/۰۱۷	۰/۰۱۶	۰/۰۱۶	۰ ± ۰/۰۱۶
	Kernel PLS	۰/۰۱۲	۰/۰۱۳	۰/۰۱۲	۰/۰۱۲	۰/۰۱۲	۰/۰۱۲	۰/۰۱۲	۰/۰۱۲	۰/۰۱۲	۰/۰۱۲	۰ ± ۰/۰۱۲

جدول (۷) جزئیات میانگین ارزیابی خطای درصد قدر مطلق خطای متقارن روش اعتبارسنجی متقابل عملکرد نظارتی در مرحله آموزش

Table (7) Details of the symmetric mean absolute percentage error of the method of mutual validation of supervisory performance in the training phase

سال	معیار MSE	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	میانگین انحراف معیار
سال جاری	Linear PINSVR	۰/۰۹۱	۰/۰۹۱	۰/۰۹۰	۰/۰۹۲	۰/۰۹۲	۰/۰۹۲	۰/۰۹۱	۰/۰۹۲	۰/۰۹۳	۰/۰۹۱	۰/۰۰۱ ± ۰/۰۹۲
	Nonlinear PINSVR	۰/۰۸۴	۰/۰۸۵	۰/۰۸۳	۰/۰۸۶	۰/۰۸۵	۰/۰۸۵	۰/۰۸۴	۰/۰۸۶	۰/۰۸۵	۰/۰۸۵	۰/۰۰۱ ± ۰/۰۸۵
	Linear PLS	۰/۰۹۱	۰/۰۹۱	۰/۰۹۰	۰/۰۹۲	۰/۰۹۲	۰/۰۹۲	۰/۰۹۱	۰/۰۹۲	۰/۰۹۳	۰/۰۹۱	۰/۰۰۱ ± ۰/۰۹۲
	Kernel PLS	۰/۰۷۹	۰/۰۸۰	۰/۰۷۹	۰/۰۸۱	۰/۰۷۹	۰/۰۸۰	۰/۰۷۹	۰/۰۸۰	۰/۰۸۰	۰/۰۷۹	۰/۰۰۱ ± ۰/۰۸۰
سال آتی	Linear PINSVR	۰/۰۸۴	۰/۰۸۷	۰/۰۸۳	۰/۰۸۵	۰/۰۸۴	۰/۰۸۵	۰/۰۸۴	۰/۰۸۶	۰/۰۸۵	۰/۰۸۹	۰/۰۰۱ ± ۰/۰۸۵
	Nonlinear PINSVR	۰/۰۷۶	۰/۰۷۹	۰/۰۷۶	۰/۰۷۷	۰/۰۷۶	۰/۰۷۷	۰/۰۷۶	۰/۰۷۷	۰/۰۷۷	۰/۰۷۹	۰/۰۰۱ ± ۰/۰۷۷
	Linear PLS	۰/۰۸۳	۰/۰۸۷	۰/۰۸۳	۰/۰۸۵	۰/۰۸۴	۰/۰۸۵	۰/۰۸۴	۰/۰۸۶	۰/۰۸۵	۰/۰۸۵	۰/۰۰۱ ± ۰/۰۸۵
	Kernel PLS	۰/۰۷۳	۰/۰۷۵	۰/۰۷۲	۰/۰۷۳	۰/۰۷۲	۰/۰۷۳	۰/۰۷۲	۰/۰۷۳	۰/۰۷۳	۰/۰۷۴	۰/۰۰۱ ± ۰/۰۷۳

جدول (۸) جزئیات میانگین قدر مطلق خطای روش اعتبارسنجی متقابل عملکرد نظارتی در مرحله آزمون

Table (8) Details of the mean absolute error of the mutual validation method of the supervisory performance in the test phase

سال	معیار MSE	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	میانگین انحراف معیار
بازه آزمون	Linear PINSVR	۰/۱۰۳	۰/۰۹۸	۰/۱۱۰	۰/۰۸۷	۰/۰۹۵	۰/۰۹۳	۰/۱۰۳	۰/۰۹۰	۰/۰۸۳	۰/۰۹۸	۰/۰۹۶ ± ۰/۰۰۸
	Nonlinear PINSVR	۰/۰۹۸	۰/۰۸۸	۰/۱۰۵	۰/۰۷۳	۰/۰۸۶	۰/۰۸۶	۰/۰۹۷	۰/۰۸۰	۰/۰۸۱	۰/۰۹۱	۰/۰۸۹ ± ۰/۰۰۱
	Linear PLS	۰/۱۰۴	۰/۰۹۹	۰/۱۱۰	۰/۰۸۸	۰/۰۹۵	۰/۰۹۴	۰/۱۰۳	۰/۰۹۱	۰/۰۸۴	۰/۰۹۹	۰/۰۹۷ ± ۰/۰۰۸
	Kernel PLS	۰/۰۸۹	۰/۰۸۲	۰/۰۹۳	۰/۰۶۵	۰/۰۸۶	۰/۰۷۵	۰/۰۹۴	۰/۰۸۱	۰/۰۷۵	۰/۰۸۸	۰/۰۸۳ ± ۰/۰۰۹
بازه آزمون	Linear PINSVR	۰/۱۰۵	۰/۰۷۱	۰/۱۱۰	۰/۰۹۴	۰/۰۹۹	۰/۰۹۶	۰/۱۰۳	۰/۰۸۵	۰/۰۹۴	۰/۰۸۷	۰/۰۹۴ ± ۰/۰۱۱
	Nonlinear PINSVR	۰/۰۹۱	۰/۰۶۳	۰/۱۰۲	/۰۸۷	۰/۰۹۱	۰/۰۸۶	۰/۰۹۸	۰/۰۸۶	۰/۰۸۴	۰/۰۷۱	۰/۰۸۶ ± ۰/۰۱۱
	Linear PLS	۰/۱۰۶	۰/۰۷۲	۰/۱۱۰	۰/۰۹۴	/۱۰۰	۰/۰۹۷	۰/۱۰۴	۰/۰۸۶	۰/۰۹۴	۰/۰۸۸	۰/۰۹۵ ± ۰/۰۱۱
	Kernel PLS	۰/۰۸۳	۰/۰۶۱	۰/۰۹۲	۰/۰۸۲	۰/۰۸۷	۰/۰۸۱	۰/۰۸۹	۰/۰۷۷	۰/۰۸۱	۰/۰۷۳	۰/۰۸۱ ± ۰/۰۰۹

جدول (۹) جزئیات میانگین مربعات خطای روش اعتبارسنجی متقابل عملکرد نظارتی در مرحله آزمون

Table (9) Details of the mean squared error error of the mutual validation method of the supervisory performance in the test phase

سال	معیار MSE	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	میانگین انحراف معیار
بازه آزمون	Linear PINSVR	۰/۰۱۸	۰/۰۲۰	۰/۰۲۳	۰/۰۱۳	۰/۰۱۷	۰/۰۱۵	۰/۰۲۲	۰/۰۱۵	۰/۰۱۳	۰/۰۱۸	۰/۰۱۷ ± ۰/۰۰۳
	Nonlinear PINSVR	۰/۰۱۷	۰/۰۱۷	۰/۰۲۲	۰/۰۰۹	۰/۰۱۴	۰/۰۱۳	۰/۰۲۰	۰/۰۱۱	۰/۰۱۳	۰/۰۱۶	۰/۰۱۵ ± ۰/۰۰۴
	Linear PLS	۰/۰۱۹	۰/۰۲۰	۰/۰۲۳	۰/۰۱۳	۰/۰۱۸	۰/۰۱۵	۰/۰۲۲	۰/۰۱۵	۰/۰۱۳	۰/۰۱۸	۰/۰۱۸ ± ۰/۰۰۳
	Kernel PLS	۰/۰۱۴	۰/۰۱۸	۰/۰۰۷	۰/۰۱۳	۰/۰۱۳	۰/۰۱۱	۰/۰۱۷	۰/۰۱۱	۰/۰۱۱	۰/۰۱۴	۰/۰۱۳ ± ۰/۰۰۳
بازه آزمون	Linear PINSVR	۰/۰۲۰	۰/۰۰۸	۰/۰۲۴	۰/۰۱۶	۰/۰۱۸	۰/۰۱۶	۰/۰۱۸	۰/۰۱۳	۰/۰۱۷	۰/۰۱۶	۰/۰۱۶ ± ۰/۰۰۴
	Nonlinear PINSVR	۰/۰۱۵	۰/۰۰۷	۰/۰۲۲	۰/۰۱۴	۰/۰۱۷	۰/۰۱۴	۰/۰۱۷	۰/۰۱۳	۰/۰۱۴	۰/۰۱۰	۰/۰۱۴ ± ۰/۰۰۴
	Linear PLS	۰/۰۲۰	۰/۰۰۸	۰/۰۲۴	۰/۰۱۶	۰/۰۱۹	۰/۰۱۶	۰/۰۱۸	۰/۰۱۳	۰/۰۱۷	۰/۰۱۶	۰/۰۱۷ ± ۰/۰۰۴
	Kernel PLS	۰/۰۱۴	۰/۰۰۷	۰/۰۱۷	۰/۰۱۱	۰/۰۱۵	۰/۰۱۲	۰/۰۱۳	۰/۰۱۰	۰/۰۱۳	۰/۰۰۹	۰/۰۱۲ ± ۰/۰۰۳

جدول (۱۰) جزئیات میانگین ارزیابی خطای درصد قدر مطلق خطای متقارن روش اعتبارسنجی متقابل عملکرد نظارتی در مرحله آزمون

Table (10) Details of the symmetric mean absolute percentage error of the mutual validation method of supervisory performance in the test phase

سال	معیار MSE	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	میانگین انحراف معیار
بازه آزمون	Linear PINSVR	۰/۰۹۷	۰/۰۹۴	۰/۱۰۳	۰/۰۸۴	۰/۰۹۰	۰/۰۹۰	۰/۰۹۹	۰/۰۸۵	۰/۰۷۶	۰/۰۹۲	۰/۰۹۱ ± ۰/۰۰۸
	Nonlinear PINSVR	۰/۰۹۳	۰/۰۸۴	۰/۱۰۰	۰/۰۷۱	۰/۰۸۱	۰/۰۸۴	۰/۰۹۳	۰/۰۷۵	۰/۰۷۴	۰/۰۸۵	۰/۰۸۴ ± ۰/۰۰۹
	Linear PLS	۰/۰۹۹	۰/۰۹۴	۰/۱۰۴	۰/۰۸۵	۰/۰۹۰	۰/۰۹۱	۰/۰۹۹	۰/۰۸۷	۰/۰۷۷	۰/۰۹۳	۰/۰۹۲ ± ۰/۰۰۸
	Kernel PLS	۰/۰۸۶	۰/۰۷۸	۰/۰۸۹	۰/۰۶۴	۰/۰۸۲	۰/۰۷۳	۰/۰۹۰	۰/۰۷۷	۰/۰۶۹	۰/۰۸۳	۰/۰۷۹ ± ۰/۰۰۸
بازه آزمون	Linear PINSVR	۰/۰۹۴	۰/۰۶۵	۰/۰۹۸	۰/۰۸۴	۰/۰۸۹	۰/۰۸۶	۰/۰۹۲	۰/۰۷۷	۰/۰۸۳	۰/۰۷۸	۰/۰۸۵ ± ۰/۰۱۰
	Nonlinear PINSVR	۰/۰۸۱	۰/۰۵۸	۰/۰۹۰	۰/۰۷۹	۰/۰۸۳	۰/۰۷۷	۰/۰۸۷	۰/۰۷۸	۰/۰۷۴	۰/۰۶۳	۰/۰۷۷ ± ۰/۰۱۰
	Linear PLS	۰/۰۹۵	۰/۰۶۵	۰/۰۹۸	۰/۰۸۵	۰/۰۹۰	۰/۰۸۷	۰/۰۹۲	۰/۰۷۸	۰/۰۸۳	۰/۰۷۹	۰/۰۸۵ ± ۰/۰۱۰
	Kernel PLS	۰/۰۷۶	۰/۰۵۶	۰/۰۸۳	۰/۰۷۴	۰/۰۸۰	۰/۰۷۴	۰/۰۸۰	۰/۰۷۰	۰/۰۷۲	۰/۰۶۵	۰/۰۷۳ ± ۰/۰۰۸

در مرحله ارزیابی قدرت یادگیری و پیش‌بینی مدل‌ها، با استفاده از روش اعتبارسنجی متقابل ۱۰تایی، مجموعه داده‌های آموزش اعتبارسنجی به چهار الگوریتم رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک خطی و غیرخطی، حداقل مربعات جزئی خطی و غیرخطی که حداقل مربعات جزئی هسته نام دارد داده می‌شود. این الگوریتم‌ها پارامترها و ابرپارامترهای خود را با استفاده از این داده یاد می‌گیرند. مجموعه نمونه‌ها با معیارهای عملکرد نظارتی (انتخاب‌شده توسط الگوریتم آرریلیف – اف به‌عنوان متغیرهای مستقل (ویژگی‌ها) به همراه متغیر وابسته سرمایه‌گذاری دارایی ثابت به چهار الگوریتم داده می‌شوند. در مرحله آموزش مدل‌های خطی و غیرخطی بعد از یادگیری، مجدداً همان

داده‌های آموزشی اعتبارسنجی بدون متغیر وابسته به آن‌ها داده می‌شود تا مقدار متغیر سرمایه‌گذاری دارایی ثابت را پیش‌بینی کنند؛ سپس با محاسبه سه معیار میانگین قدر مطلق خطا، میانگین مربعات خطا و میانگین ارزیابی خطای درصد قدر مطلق خطای متقارن قدرت یادگیری و خطای یادگیری این مدل‌ها مورد محک قرار می‌گیرد. مسئله بعدی که باید بررسی شود قدرت پیش‌بینی مدل‌ها و عدم رخداد پدیده بیش‌برازش در فرآیند مرحله یادگیری است. در مرحله آزمون، به همین علت داده‌های آزمون که در فرآیند اعتبارسنجی متقابل ۱۰تایی کنار گذاشته شده‌اند به مدل‌های یاد گرفته شده وارد می‌شوند تا قدرت پیش‌بینی آن‌ها برای نمونه‌هایی که هنوز مشاهده نکرده‌اند مورد ارزیابی قرار گیرد. انتظار می‌رود که اختلاف خطاهای بین مرحله آموزش و آزمون خیلی فاحش نباشد. ممکن است خطاهای مرحله آزمون کم‌تر یا بیش‌تر از مقدار خطاهای مرحله آموزش باشند؛ این مسئله مهمی نیست؛ چیزی که دارای اهمیت است اختلاف کم خطاهای گزارش‌شده بین این دو مرحله است (Alpaydın, 2021). قدرت یادگیری مدل‌ها و میانگین و ارزیابی جزئیات سه خطا در جدول (۱۱) و جدول (۱۲) نشان داده شده است.

جدول (۱۱) قدرت یادگیری مدل‌ها با استفاده از سه معیار ارزیابی میانگین قدر مطلق خطا، میانگین مربعات خطا و میانگین ارزیابی خطای درصد قدر مطلق خطای متقارن در مرحله آموزش

Table (11) Learning power of models using three error evaluation criteria mean absolute error, mean squared error and symmetric mean absolute percentage error in the training phase

MAE	سال جاری	سال آتی
Linear PINSVR	۰/۰۹۶ ± ۰/۰۰۱	۰/۰۹۴ ± ۰/۰۰۱
Nonlinear PINSVR	۰/۰۸۹ ± ۰/۰۰۱	۰/۰۸۶ ± ۰/۰۰۱
Linear PLS	۰/۰۹۶ ± ۰/۰۰۱	۰/۰۹۴ ± ۰/۰۰۱
Kernel PLS	۰/۰۸۳ ± ۰/۰۰۱	۰/۰۸۱ ± ۰/۰۰۱
MSE	سال جاری	سال آتی
Linear PINSVR	۰/۰۱۷ ± ۰	۰/۰۱۶ ± ۰
Nonlinear PINSVR	۰/۰۱۵ ± ۰	۰/۰۱۴ ± ۰
Linear PLS	۰/۰۱۷ ± ۰	۰/۰۱۶ ± ۰
Kernel PLS	۰/۰۱۳ ± ۰	۰/۰۱۲ ± ۰
SMAPE	سال جاری	سال‌جاری
Linear PINSVR	۰/۰۹۲ ± ۰/۰۰۱	۰/۰۸۵ ± ۰/۰۰۱
Nonlinear PINSVR	۰/۰۸۵ ± ۰/۰۰۱	۰/۰۷۷ ± ۰/۰۰۱
Linear PLS	۰/۰۹۲ ± ۰/۰۰۱	۰/۰۸۵ ± ۰/۰۰۱
Kernel PLS	۰/۰۸۰ ± ۰/۰۰۱	۰/۰۷۳ ± ۰

جدول (۱۲) میانگین و انحراف معیارهای خطا برای قدرت پیش‌بینی مدل‌ها در مرحله آزمون
Table (12) The mean and standard deviation of the error criteria for the predictive power of the models in the test phase

MAE	سال جاری	سال آتی
Linear PINSVR	۰/۰۹۶ ± ۰/۰۰۸	۰/۰۹۴ ± ۰/۰۱۱
Nonlinear PINSVR	۰/۰۸۹ ± ۰/۰۱۰	۰/۰۸۶ ± ۰/۰۱۱
Linear PLS	۰/۰۹۷ ± ۰/۰۰۸	۰/۰۹۵ ± ۰/۰۱۱
Kernel PLS	۰/۰۸۳ ± ۰/۰۰۹	۰/۰۸۱ ± ۰/۰۰۹
MSE	سال جاری	سال آتی
Linear PINSVR	۰/۰۱۷ ± ۰/۰۰۳	۰/۰۱۶ ± ۰/۰۰۴
Nonlinear PINSVR	۰/۰۱۵ ± ۰/۰۰۴	۰/۰۱۴ ± ۰/۰۰۴
Linear PLS	۰/۰۱۸ ± ۰/۰۰۳	۰/۰۱۷ ± ۰/۰۰۴
Kernel PLS	۰/۰۱۳ ± ۰/۰۰۳	۰/۰۱۲ ± ۰/۰۰۳
SMAPE	سال جاری	سال‌جاری
Linear PINSVR	۰/۰۹۱ ± ۰/۰۰۸	۰/۰۸۵ ± ۰/۰۱۰
Nonlinear PINSVR	۰/۰۸۴ ± ۰/۰۰۹	۰/۰۷۷ ± ۰/۰۱۰
Linear PLS	۰/۰۹۲ ± ۰/۰۰۸	۰/۰۸۵ ± ۰/۰۱۰
Kernel PLS	۰/۰۷۹ ± ۰/۰۰۸	۰/۰۷۳ ± ۰/۰۰۸

در پاسخ به فرضیه اول پژوهش (الگوریتم‌های خطی و غیرخطی هوش مصنوعی، توانایی پیش‌بینی سرمایه‌گذاری در دارایی ثابت شرکت‌ها را دارد) که باتوجه‌به عملکرد الگوریتم‌های رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک و حداقل مربعات جزئی هسته که توضیح داده شد، مشاهده شد این مدل هوش مصنوعی که در حوزه داده‌کاوی و شناسایی الگو قرار دارند دارای دو حالت خطی و غیرخطی هستند که برای حالت خطی می‌توان ضرایب مدل خطی آن‌ها را گزارش داد؛ ولی در مورد حالت غیرخطی آن‌ها باتوجه‌به نگاشت داده‌ها از فضای ورودی به فضای ویژگی توسط نگاشت غیرخطی نامعلوم و همچنین بعد بالای فضای ویژگی نمی‌توان پارامترهای مدل را ارائه نمود. همچنین در بررسی الگوریتم‌های هوش مصنوعی خطی و غیرخطی در پیش‌بینی سرمایه‌گذاری دارایی ثابت شرکت‌ها، الگوریتم‌های خطی نسبت به الگوریتم‌های غیرخطی کارایی بیشتری نشان دادند.

در پاسخ به فرضیه دوم پژوهش (معیارهای نظارتی نظام راهبری شرکتی بر میزان سرمایه‌گذاری دارایی ثابت شرکت‌ها تأثیرگذار هستند) می‌توان این‌گونه عنوان کرد که با اعمال روش انتخاب ویژگی آرریلیف – اف بر روی متغیرهای مستقل نظارتی راهبری بیش‌ترین تأثیر را متغیرهای تعداد هیئت‌مدیره، استقلال هیئت‌مدیره، تخصص مالی هیئت‌مدیره به کل اعضا، تعداد اعضای کمیته حسابرسی، استقلال

کمیته حسابرسی، مالکیت نهادی بالای ۵ درصد، دوره تصدی مدیرعامل، وجود حسابرس داخلی، تعداد اعضای کمیته حسابرسی، نسبت تخصص اعضای کمیته و دوگانگی وظایف مدیرعامل بر پیش‌بینی دارایی ثابت نشان دادند (Robnik-Šikonja and Kononenko, 1997)؛ سپس متغیرهای مستقل و متغیر وابسته به چهار الگوریتم داده شدند. در مرحله آموزش مدل‌های خطی و غیرخطی بعد از یادگیری، مجدداً همان داده‌های آموزشی اعتبارسنجی بدون متغیر وابسته به آن‌ها داده شد تا مقدار متغیر سرمایه‌گذاری دارایی ثابت را پیش‌بینی کنند و پس از آن با محاسبه سه معیار میانگین قدر مطلق خطا، میانگین مربعات خطا و میانگین ارزیابی خطای درصد قدر مطلق خطای متقارن قدرت یادگیری و خطای یادگیری این مدل‌ها مورد محک قرار گرفت. در ادامه قدرت پیش‌بینی مدل‌ها و عدم رخداد پدیده بیش‌برازش در فرآیند مرحله یادگیری انجام شد و در مرحله آزمون، متغیرهایی که در فرآیند اعتبارسنجی متقابل ۱۰تایی کنار گذاشته شده‌اند به مدل‌های یاد گرفته شده وارد شدند و قدرت پیش‌بینی آن‌ها برای نمونه‌هایی که هنوز مشاهده نکرده‌اند مورد ارزیابی قرار گرفت.

۵- بحث و نتیجه‌گیری

حاکمیت شرکتی مطلوب نقش مهمی در تصمیمات سرمایه‌گذاری و رشد اقتصادی و درعین‌حال افزایش اعتماد سرمایه‌گذاران دارد. از طرفی دیگر، اعتماد سرمایه‌گذاران نیز نقش مهمی در اقتصاد کشور ایفا می‌کند. آگاهی از این امر که ویژگی‌ها و ساختار هیئت‌مدیره شرکت‌ها می‌تواند بر کیفیت اطلاعات حسابداری اثرگذار باشد و برای سرمایه‌گذارانی که نسبت به خرید سهام شرکت علاقه نشان می‌دهند می‌تواند مفید و اثرگذار باشد. براین‌اساس، کاربرد روش‌هایی ضروری به نظر می‌رسد که امکان انتخاب بهینه در چنین موقعیتی را فراهم آورد (Azad and Pourzamani, 2019). ازاین‌رو، واحدهای تجاری برای سرمایه‌گذاری در طرح‌های مختلف، باید حد یا میزان سرمایه‌گذاری را باتوجه‌به محدودیت منابع، موردتوجه قرار دهند تا بدین‌وسیله حداکثر کارایی را از سرمایه‌گذاری خود کسب نمایند و به یک سرمایه‌گذاری مضاعف دست یابند. افزایش در سرمایه‌گذاری، فعالیتی برای کسب سود است و نااطمینانی و خطر، دوعاملی هستند که بر سود انتظاری و در نتیجه تصمیم‌گیری برای سرمایه‌گذاری مؤثر هستند (Alipour *et al*., 2021). یانگ و همکاران (Yang *et al*., 2016) مطالعه‌ای در زمینه الگوریتم بردار پشتیبان غیرموازی غیرحساس پارامتریک برای رگرسیون داده‌ها، مورد بررسی قرار دادند. نتایج آزمایش در مقایسه با سایر روش‌های هوش مصنوعی نشان می‌دهد که نه‌تنها عملکرد رگرسیون قابل‌مقایسه‌ای را به دست می‌آورد؛ بلکه تخمین‌های بهتری را نیز به دست می‌آورد. روش رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک نیز به طور غیرمستقیم یک جفت توابع پروگزیمال غیرموازی را با یک

جفت توابع پروگزیمال غیرموازی متفاوت با حل دو مسئله برنامه‌ریزی درجه دوم با اندازه کوچک‌تر پیدا می‌کند. با استفاده از توابع جدید ازدست‌دادن پارامتری غیرحساس، پیشنهادی به طور خودکار یک منطقه غیرحساس پارامتری انعطاف‌پذیر با شکل دلخواه و حداقل اندازه را تنظیم می‌کند تا اطلاعات داده‌شده را برای ثبت ساختار داده و اطلاعات مرزی با دقت بیشتری در بر گیرد (Yang *et al*., 2016). روش حداقل مربعات جزئی هسته یا حداقل مربعات جزئی خطی چندین مزیت نسبت به رویکردهای قبلی دارد: ۱) این روش ابعاد ویژگی را کاهش می‌دهد. ۲) این روش قادر است تعداد کمی از متغیرهای پنهان را پیدا کند؛ به‌عنوان‌مثال، هزاران ویژگی را در یک فضای فرعی با ابعاد بسیار پایین نمایش دهد که ممکن است تأثیر زیادی بر برنامه‌های کاربردی زمان واقعی داشته باشد. ۳) همچنین دارای یک بردار خروجی است که می‌تواند چندین برچسب داشته باشد؛ به‌طوری‌که چندین مشکل مرتبط را می‌تواند به‌طورکلی حل کند (Mello-Román *et al*., 2021). در مقایسه با دو روش حداقل مربعات جزئی هسته و رگرسیون بردار پشتیبان غیرموازی غیرحساس پارامتریک، روش دیگری به نام ماشین بردار پشتیبان در حل روش‌های هوش مصنوعی در پیشینه پژوهش وجود دارد. الگوریتم ماشین بردار پشتیبانی همچنین یک مدل یادگیری ماشینی قدرتمند و انعطاف‌پذیر است؛ این الگوریتم نه‌تنها از طبقه‌بندی خطی و غیرخطی بلکه از رگرسیون خطی و غیرخطی رگرسیون بردار پشتیبان[1] نیز پشتیبانی می‌کند. ترفند استفاده از ماشین بردار پشتیبانی برای رگرسیون، معکوس‌کردن هدف است؛ به‌جای تلاش برای جادادن وسیع‌ترین مسیر دوجمله‌ای دوکلاسه که نقطه پرت را به حداقل می‌رساند. در رگرسیون ماشین بردار پشتیبانی، تلاش بر این است که تاحدامکان نمونه‌های بیشتری در این مسیر (لبه) قرار داده شود و درعین‌حال تخلفات لبه به حداقل برسد. الگوریتم ماشین بردار پشتیبانی می‌تواند برای داده‌های پرت نیز استفاده شود؛ اما اشکال این الگوریتم فقط برای داده‌های آموزشی پیچیده کوچک و متوسط خوب است؛ اما با افزایش تعداد نمونه‌های داده‌های آموزشی بسیار کند می‌شود (Géron, 2022). در این پژوهش باتوجه‌به اهمیت تصمیم‌گیری در سرمایه‌گذاری شرکت‌های بورس اوراق بهادار تهران، به الگوسازی سرمایه‌گذاری در دارایی ثابت با رویکرد هوش مصنوعی با تأکید بر نقش معیارهای نظارتی نظام راهبری پرداخته شده است؛ بدین منظور از اطلاعات ۱۰۱ شرکت بورسی طی سال‌های ۱۳۹۰ الی ۱۳۹۹ استفاده شده است و نتایج پژوهش به این صورت می‌باشد:

در این پژوهش متغیرها، پس از غربالگری با روش حذف سیستماتیک انتخاب شدند سپس با استفاده از روش متغیرگزینی آرریلیف – اف متغیرهای مؤثر در پیش‌بینی سرمایه‌گذاری دارایی ثابت شامل متغیرهای اندازه هیئت‌مدیره، استقلال هیئت‌مدیره، تخصص مالی هیئت‌مدیره، اندازه کمیته

1. Support Vector Regression (SVR)

حسابرسی، استقلال کمیته حسابرسی، مالکیت نهادی بالای ۵ درصد، دوره تصدی مدیرعامل، وجود حسابرس داخلی، نسبت تخصص اعضای کمیته و دوگانگی وظایف مدیرعامل انتخاب شدند و پس از آن متغیرها با روش اعتبارسنجی متقابل به داده‌های، آموزش و آزمون تقسیم شدند. سپس متغیرهای آموزش به چهار الگوریتم هوش مصنوعی داده شد و میانگین و قدرت خطای یادگیری مدل‌ها با سه معیار خطا مورد بررسی قرار گرفتند. در تمامی این خطاها، هر مقدار خطا به صفر نزدیک‌تر باشد، الگوریتم متناظر آن قدرت یادگیری بالاتر داشته است و الگوریتم بهتر عمل کرده است. میزان قدرت یادگیری مدل‌های خطی به شدت کم است و هر سه خطا، این مسئله را به‌خوبی نشان می‌دهند؛ بنابراین پدیده بیش‌برازش اتفاق نیفتاده است و کلیه مباحث مطرح‌شده در مرحله آموزش در مرحله آزمون نیز درست است. اختلاف خطای مدل‌های خطی با مدل‌های غیرخطی آن‌چنان زیاد نیست. این نتیجه‌گیری از پیچیدگی فضای ورودی مسئله نتیجه گرفته می‌شود که برای ماهیت مسئله سرمایه‌گذاری دارایی ثابت با معیارهای عملکرد نظارتی مدل‌های خطی از مدل‌های غیرخطی کارآمدتر هستند. در خطاهای گزارش‌شده خطاهای سال جاری و آتی نزدیک به یکدیگر هستند و نشان‌دهنده یادگیری خوب مدل سال جاری و آتی است.

نتایج این پژوهش می‌تواند به‌صورت کاربردی موردتوجه فعالین بازار سرمایه ایران قرار گیرد به‌طوری‌که با پیش‌بینی سرمایه‌گذاری در دارایی ثابت شرکت‌ها و کارکردن بر روی عوامل مؤثر بر آن، قادر هستند نسبت به مدیریت‌کردن جذب سرمایه سهامداران، کاهش بحران‌های مالی، افزایش بهبود حاکمیت شرکتی و کمک به سرمایه‌گذاران جهت اجتناب از زیان‌های بزرگ در بازار سهام، اقدام نمایند.

جهت پژوهش‌های آتی نیز پیشنهاد می‌شود تأثیر معیارهای کیفیت اطلاعات بر تصمیمات سرمایه‌گذاری شرکت‌ها در دارایی ثابت، تأثیر معیارهای سرمایه فکری بر تصمیم به سرمایه‌گذاری فعالان بازار سرمایه در دارایی ثابت و تأثیر معیارهای مدیریت خطر بر تصمیمات سرمایه‌گذاری در دارایی ثابت مورد پژوهش و بررسی قرار گیرند؛ همچنین باتوجه‌به عملکرد بهینه کاربرد هوش مصنوعی که در این پژوهش به آن اشاره شد برای پیش‌بینی سرمایه‌گذاری در دارایی ثابت شرکت‌ها، از ترکیب روش الگوریتم ژنتیک با سایر روش‌های هوش مصنوعی استفاده شود.

در خصوص محدودیت‌های پژوهش می‌توان این‌گونه عنوان کرد که اقلام مندرج در متن صورت‌های مالی به واسطه آثار تورم ممکن است تعـدیل نشوند و ازآنجاکه واحدهای تجاری زمان‌های متفاوت تأسیس دارند و اقلام دارایی‌های خـود را در زمان‌های مختلف تحصیل نموده‌اند؛ لذا کیفیت قابلیت مقایسه‌ای اقلام می‌تواند بر نتایج پژوهش اثر گذاشته و تعمیم نتایج را با محدودیت‌هایی همراه سازد. از سویی دیگر اثرات ناشی از تفاوت در روش‌های حسابداری در اندازه‌گیری گزارش رویدادهای

مالی نیز ممکن است بر نتایج پژوهش اثر بگذارد؛ لذا ضروری است تعدیلی از این بابت صورت پذیرد. ازجمله محدودیت‌هایی که در تفسیر نتایج این پژوهش باید بررسی شود، این است که همانند سایر پژوهش‌هایی که در محدودیت مالی و سرمایه‌گذاری انجام شده است؛ معیارهایی که برای اندازه‌گیری متغیرها به کار گرفته شده است ممکن است در معرض خطای اندازه‌گیری قرار گرفته باشند.

٦- تعارض منافع

هیچ‌گونه تعارض منافع در این پژوهش وجود ندارد.

۷- منابع

Ahmed, B., Akbar, M., Sabahat, T., Ali, S., Hussain, A., Akbar, A., & Hongming, X. (2020). Does firm life cycle impact corporate investment efficiency? *Sustainability*, *13*(1), 197-209.

Alipour, H., Tehrani, R., Alirezaei, A., & Abbaspour Esfndan, G. (2021). Providing a Model of the Effect of Uncertainty about the Inflation Rate and More Trust of Managers on the Increase in Investment (Case Study: Fellowship Affiliated Companies in Tehran Stock Exchange). *Journal of Investment Knowledge*, *10*(38), 223-242. [In Persian]

Alpaydın, E. (2021). Machine learning. *MIT Press*.

Amiri, M., Raeesi Vanani, I., Razavi Hajiagha, S. H., & Jafari, T. (2021). Development of fuzzy Artificial Intelligence and Multi-Objective planning Model to Optimize the Portfolio of Investment Companies. *Business Intelligence Management Studies*, *9*(36), 209-243. [In Persian]

Azad, A. & Pourzamani, Z. (2019). Providing a model for measuring the efficiency of companies with the role of regulatory criteria (cultural approach and Lasso). *Financial Economics*, *14*(53), 65-96. [In Persian]

Azar, A., & Momeni, M. (2009). *Statistics and its application in management volume two: statistical analysis*. Samt. [In Persian]

Bahreini, A., Fard, M. A., & Khoshnood, M. (2023). Developing an LSTM neural network model for predicting blocktrade transaction valuation. *Journal of Advances in Finance and Investment*, *4*(4), 145-176. [In Persian]

Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent data analysis*, *1*(1-4), 131-156.

De Vries, S., & Ter Braak, C. J. (1995). Prediction error in partial least squares regression: a critique on the deviation used in The Unscrambler. *Chemometrics and intelligent laboratory systems*, *30*(2), 239-245.

Feridoni, M., Rezaei, N., Pakmaram, A., & Abdi, R. (2022). Improve Investment Efficiency Based on Cognitive Behavioral Biases CEO by Gray Vikor Analysis. *Journal of Investment Knowledge*, *11*(42), 579-609. [In Persian]

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly.

He, Y., Chen, C., & Hu, Y. (2019). Managerial overconfidence, internal financing, and investment efficiency: Evidence from China. *Research in International Business and Finance*, *47*, 501-510.

Hubbard, R. G. (1998). Capital-market imperfections and investment. *Journal of Economic Literature*, *36*(1), 193-225.

Huerta, R., Corbacho, F., & Elkan, C. (2013). Nonlinear support vector machines can systematically identify stocks with high and low future returns. *Algorithmic Finance*, *2*(1), 45-58.

Jahanshad, A., & Khalili, S. A. (2013). The relationship between stock returns with fixed assets growth a wavelet analysis. *Financial Engineering and Portfolio Management*, *4*(15), 1-15. [In Persian]

Khaleghi Zadeh Dehkordi, M., Sarraf, F., & Najafi Moghadam, A. (2022). Application of artificial intelligence algorithm in predicting investment efficiency Emphasizing the role of risk management criteria. *Journal of Investment Knowledge*, *11*(42), 413-434. [In Persian]

Khoshkar, F., Safari Khah, M., & Ranjbari Share Jini, S. (2020). The effect of ownership structure on the relationship between financial information quality and investment efficiency. *Journal of Accounting and Management Vision*, *3*(33), 162-175. [In Persian]

Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, *10*(1), 183-208.

Lara, J. M. G., Osma, B. G., & Penalva, F. (2016). Accounting conservatism and firm investment efficiency. *Journal of accounting and economics*, *61*(1), 221-238.

Lv, P., & Xiong, H. (2022). Can FinTech improve corporate investment efficiency? Evidence from China. *Research in International Business and Finance*, *60*, 101571.

Mansourfar, G., Joudi, S., & Poursoleiman, E. (2020). Associate Prof., Department of Finance, Faculty of Management and Economics, Urmia University, West Azarbaijan, Urmia, Iran. *Financial Research Journal*, *22*(2), 227-248. [In Persian]

Mello-Román, J. D., Hernández, A., & Mello-Román, J. C. (2021). Improved Predictive Ability of KPLS Regression with Memetic Algorithms. *Mathematics*, *9*(5), 506-518.

Menshawy, I. M., Basiruddin, R., Mohd-Zamil, N. A., & Hussainey, K. (2023). Strive towards investment efficiency among Egyptian companies: Do board characteristics and information asymmetry matter? *International Journal of Finance & Economics*, *28*(3), 2382-2403.

Mohammadi, D., Mohammadi, E., Shokri, N., & Heidari, N. (2023). The use of support vector machine and Naive Bayes algorithms and its combination with risk measure and fuzzy theory in the selection of stock portfolio. *Journal of Advances in Finance and Investment*, *4*(4), 177-206. [In Persian]

Naim, P., & Condamin, L. (2019). *Operational risk modeling in financial services: The exposure, occurrence, impact method*. John Wiley & Sons.

Robnik-Šikonja, M., & Kononenko, I. (1997). An adaptation of Relief for attribute estimation in regression. *In Machine learning: Proceedings of the fourteenth international conference (ICML'97)*, *5*, 296-304.

Rosipal, R., & Trejo, L. J. (2001). Kernel partial least squares regression in reproducing kernel hilbert space. *Journal of machine learning research*, *2*(2001), 97-123.

Samet, M., & Jarboui, A. (2017). How does corporate social responsibility contribute to investment efficiency? *Journal of multinational financial management*, *40*, 33-46.

Sayyadi, M., Dastgir, M., & Aliahmdi, S. (2019). Study on Enterprise Risk Management (ERM) effect on managerial ability in order to increasing investment efficiency. *Financial Management Strategy*, *7*(1), 1-38. [In Persian]

Silverstein, B. (2020). Managerial opportunism and corporate investment efficiency. *SSRN*.

Taghizadeh Khanqah, V., Badavar Nahandi, Y., Mottagi, A., & Taghizadeh, H. (2021). Validation of Investment Efficiency Models Based on Agency Theory, Information Asymmetry, Managerial Entrenchment and Firm value maximization. *Journal of Investment Knowledge*, *10*(38), 287-318. [In Persian]

Tahmooresi, Z., Talebnia, Q., Baradaran Hassanzadeh, R., Mousavi, N. A., & Vakilifard, H. (2024). Modeling The Effective variables on investment efficiency: evidence from Tehran Stock Exchange. *Journal of Investment Knowledge*, *13*(51), 361-384. [In Persian]

Wang, J., Wang, H., & Wang, D. (2021). Equity concentration and investment efficiency of energy companies in China: Evidence based on the shock of deregulation of QFIIs. *Energy Economics*, *93*, 105032.

Yang, Z. M., Hua, X. Y., Shao, Y. H., & Ye, Y. F. (2016). A novel parametric-insensitive nonparallel support vector machine for regression. *Neurocomputing*, *171*, 649-663.

Zhang, X., Hu, Y., Xie, K., Wang, S., Ngai, E. W. T., & Liu, M. (2014). A causal feature selection algorithm for stock prediction modeling. *Neurocomputing*, *142*, 48-59.

COPYRIGHTS
© 2024 by the author. Published by Islamic Azad University, Esfarayen Branch. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) (http://creativecommons.org/licenses/by/4.0/).