

Investigating Factors Affecting the Stock Price Index in Iran: A Machine Learning Approach

Mohammad Mahdi Barghi Osgouei¹ , Reza Rajpour², Saman Hatamerd³,
Maryam Amini⁴, Sadeq Rezaie⁵

1. Professor, Department of Economic Sciences, Faculty of Economics and Management, University of Tabriz, Iran. (Corresponding Author) E-mail: m.barghi@tabrizu.ac.ir
2. Professor, Faculty of Economics and Management, University of Tabriz, Iran. E-mail: reza.ranjpour@gmail.com
3. Assistant Professor, Department of Economic Sciences, Faculty of Humanities, University of Zanjan, Zanjan, Iran. E-mail: samanhatamerad@yahoo.com
4. PhD in Regional & Urban Economics, University of Isfahan, Isfahan, Iran. E-mail: ma.amini@ase.ui.ac.ir
5. Assistant Professor, Entrepreneurship Development Department, Faculty of Entrepreneurship, University of Tehran, Tehran, Iran. E-mail: sadeqrezaie@ut.ac.ir

Abstract

This study rigorously investigates the macroeconomic determinants of the Iranian stock price index over the period 1990–2020. To this end, three feature selection methodologies random forest, regular linear regression, and mutual information were employed to identify the most influential explanatory variables. In addition, the jointness approach was applied to account for potential interdependencies among variables, thereby enhancing the robustness of the empirical findings. The comparative evidence from the random forest models and regular linear regression regression converged, indicating that the exchange rate, financial development, inflation, economic growth, trade openness, and global uncertainty are of paramount significance in explaining fluctuations in Iran's stock price index. Among these determinants, global uncertainty exerted a persistently negative effect, while the remaining macroeconomic variables contributed positively to stock market performance. Furthermore, the application of the mutual information technique allowed for a temporal decomposition of variable importance across different decades. The results revealed that the exchange rate, financial development, and unemployment consistently retained explanatory power throughout the entire sample horizon, underscoring their structural relevance to Iran's stock market dynamics. Conversely, variables such as the interest rate exhibited a more episodic pattern of influence, being highly significant during certain decades but relatively marginal in others. These findings highlight the central role of macro-financial stability in sustaining stock market development. In particular, exchange rate management, deepening of financial markets, and reduction of unemployment emerge as key policy priorities. Moreover, the destabilizing role of global uncertainty underscores the necessity of strengthening resilience mechanisms in order to shield the Iranian stock market from external shocks and systemic vulnerabilities.

Article information

Review History:

Received: May. 13, 2024

Revised: June. 09, 2024

Accepted: July. 10, 2024

Keywords:

Machine learning

Stock market

Stock price

JEL Classification:

G12,G17,C15

Corresponding Author:

m.barghi@tabrizu.ac.ir



Aim and Introduction

Achieving sustained and long-term economic growth necessitates the optimal allocation and utilization of resources at the national level. This goal relies heavily on the existence of efficient financial markets, particularly well-functioning and extensive capital markets. Numerous macroeconomic variables can influence the level of risk associated with shareholder rights, corporate cash flows, and adjusted discount rates. Additionally, changes in economic conditions can alter both the quantity and nature of investment opportunities.

However, establishing a fixed and consistent relationship between macroeconomic variables and stock price indices remains challenging. The complex and dynamic nature of financial markets makes it difficult to identify a method that accurately reflects economic conditions and captures the most critical influencing variables. Therefore, this study employs machine learning models to identify the key macroeconomic factors affecting affecting Iran's stock price index, aiming to provide a comprehensive and systematic understanding of the relationship between these factors and stock market performance.

Methodology

Feature selection is one of the most common and crucial techniques in data preprocessing and serves as an essential component of machine learning. This study employs feature selection models to identify the most relevant predictors of the stock price index. The models utilized include the random forest method and regularized linear regression. To examine the nature of the relationships between variables, the jointness method was applied. Additionally, the mutual information analysis was conducted to assess the influence of key variables over different decades, enabling a deeper understanding of how the impact of macroeconomic factors on stock prices has evolved over time.

Findings

The study analyzed the impact of selected macroeconomic variables on stock price indices, focusing on the Tehran Stock Exchange. The findings from the Random Forest (RF) and Regularized Linear Regression (RLR) models indicate that exchange rates, financial development, inflation, economic growth, trade openness, and global uncertainty significantly influence Iran's stock price index. The results demonstrate that global uncertainty, interest rates, and trade openness exert negative effects on stock prices, whereas the other variables positively influence stock prices.

The jointness method was employed to analyze the relationships between these variables, further confirming their significance. Moreover, the Mutual Information method was used to examine how the influence of these key variables varied across different decades. The findings revealed that the exchange rate, financial development, and unemployment were consistently influential across all three decades. In contrast, other variables, such as interest rates, demonstrated varying degrees of significance depending on the decade, with some variables becoming influential only in particular periods.

Discussion and Conclusion

Among the variables examined, exchange rates, financial development, inflation, economic growth, trade openness, and global uncertainty emerged as the most significant factors influencing Iran's stock price index. This finding is not surprising, given Iran's historical experience with significant exchange rate fluctuations and persistent inflationary pressures. Global uncertainty has consistently influenced domestic markets in Iran due to political and economic instability. Previous research has highlighted the complex relationship between exchange rate fluctuations and stock price indices (Ratanapakorn & Sharma, 2007). Scholars have argued that the relationship between stock prices and exchange rates can significantly affect monetary and fiscal policy, as a recessionary stock market can reduce overall demand and impact broader economic performance.

Extensive research has also investigated the relationship between inflation and stock prices, identifying inflation as a significant factor affecting stock indices (Boudoukh & Richardson, 1993; Fama & Schwert, 1977; Jaffe & Mandelker, 1976). While some studies have reported a positive correlation between inflation and stock prices, others have found a negative relationship.

Moreover, trade openness has been recognized as a key factor influencing stock market fluctuations. Open economies are more vulnerable to external shocks due to increased global risk-sharing among markets. Although some studies have not found conclusive evidence of a direct effect between trade openness and stock prices, trade openness remains one of the influential factors (Nickmansh, 2016). Stock prices reflect the present value of future cash flows, which are subject to two main effects: cash flow changes driven by increased production and interest rates, which serve as a discount factor. Stock prices tend to decline when expected cash flows decrease or interest rates rise. The level of actual economic activity directly influences cash flows, as higher economic activity generally leads to increased cash flow. Among the various indicators used to predict commodity markets, real Gross Domestic Product (GDP) is considered the most comprehensive measure of economic activity (Yuhasin, 2011; Christopher et al., 2006).

Finally, global uncertainty plays a significant role in shaping the internal economic environment of countries, making it an important global macroeconomic variable that influences the performance of publicly traded companies on the stock exchange.

بررسی عوامل مؤثر بر شاخص قیمت سهام در ایران: رویکرد

یادگیری ماشین

محمد مهدی برقی اسگویی^۱ , رضا رنجپور^۲, سامان حاتم راد^۳, مریم امینی^۴, صادق رضایی^۵

۱. استاد، گروه علوم اقتصادی، دانشکده اقتصاد و مدیریت، دانشگاه تبریز، تبریز، ایران (نویسنده مسئول)

m.barghi@tabrizu.ac.ir

۲. استاد، گروه توسعه و برنامه ریزی اقتصادی، دانشکده اقتصاد و مدیریت، دانشگاه تبریز، تبریز، ایران

reza.ranjpour@gmail.com

۳. استادیار، گروه علوم اقتصادی، دانشکده علوم انسانی، دانشگاه زنجان، زنجان، ایران

samanhatamerad@yahoo.com

۴. دکتر در اقتصاد شهری و منطقه ای، دانشکده علوم اداری و اقتصاد، دانشگاه اصفهان، اصفهان، ایران

ma.amini@ase.ui.ac.ir

۵. استادیار، گروه توسعه کارآفرینی، دانشکده کارآفرینی، دانشگاه تهران، تهران، ایران sadeqrezaie@ut.ac.ir

چکیده	اطلاعات مقاله
در این مطالعه، عوامل کلان اقتصادی مؤثر بر شاخص قیمت سهام ایران در طی سال‌های ۱۹۹۰ الی ۲۰۲۰ بررسی می‌شود. در این راستا از سه متد یادگیری ماشین جنگل تصادفی، رگرسیون خطی منظم و اطلاعات متقابل، استفاده شده است. این مدل‌ها به مدل‌های انتخاب ویژگی معروفاند. مضافاً از رویکرد جوبینتسن نیز برای مطالعه ارتباط میان متغیرهای توضیحی با یکدیگر استفاده شده است. نتایج دو مدل جنگل تصادفی و رگرسیون خطی منظم، یکدیگر را تأیید کردند و نشان دادند که از میان متغیرهای مورد مطالعه، متغیرهای نرخ ارز، توسعه مالی، تورم، رشد، بازبودن تجاری و نااطمینانی جهانی، از اهمیت ویژه‌ای برخوردارند. همچنین نااطمینانی، تأثیر منفی و سایر متغیرهای اثرگذار، تأثیر مثبت بر شاخص قیمت سهام ایران داشته است. در این مطالعه، از روش اطلاعات متقابل برای بررسی تأثیرگذاری این متغیرها در دهه‌های مختلف استفاده شد. نتایج نشان داد که در سه دهه مورد بررسی، متغیرهای نرخ ارز، توسعه مالی و بیکاری، بسیار مهم و تأثیرگذار بوده‌اند؛ درحالی که سایر متغیرها همچون نرخ بهره، ممکن است در دو دهه مهم بوده باشد، ولی در یکی از دهه‌ها، تأثیر به مراتب کمتری را داشته، و یا فقط در یک دهه خاص مهم بوده است. این یافته‌ها بر اهمیت ثبات کلان مالی در تقویت توسعه بازار سهام ایران تأکید دارند. مدیریت کارآمد نرخ ارز، تعمیق بازارهای مالی و کاهش نرخ بیکاری به عنوان اولویت‌های اصلی سیاست‌گذاری مطرح می‌شوند. علاوه بر این، نقش بی‌ثبات‌کننده عدم اطمینان جهانی ضرورت تقویت سازوکارهای تاب‌آوری را برای کاهش آسیب‌پذیری بازار سهام در برابر شوک‌های خارجی و ریسک‌های سیستمیک برجسته می‌سازد.	تاریخچه داوری: دریافت: ۱۴۰۳/۰۲/۲۴ بازنگری: ۱۴۰۳/۰۳/۲۰ پذیرش: ۱۴۰۳/۰۴/۱۹ کلمات کلیدی: قیمت سهام یادگیری ماشین بازار سهام طبقه‌بندی JEL: G12, G17, C1 نویسنده مسئول: m.barghi@tabrizu.ac.ir

۱. مقدمه

دستیابی به رشد اقتصادی بلندمدت و مستمر، مستلزم تجهیز و تخصیص بهینه منابع در سطح اقتصاد ملی است؛ که بدون برخورداری از بازارهای مالی، به‌ویژه بازارهای کارآمد و گسترده سرمایه امکان پذیر نیست. در واقع فروپاشی بازارهای سهام همیشه منجر به بحران مالی و کوتاه‌مدت اقتصادی می‌شود و به همین دلیل، بازار سهام یکی از مهم‌ترین بازارهای مالی است (حاتم راد و همکاران ۱۴۰۲). بازار سهام به عنوان یکی از منابع حیاتی تأمین سرمایه و بستری برای گردآوری سرمایه‌های کوچک و بزرگ شناخته می‌شود (صمدی و بیانی، ۱۳۹۰)؛ از این رو تغییرات در این بازار بر متغیرهای کلان اقتصادی مؤثر خواهد بود. در مجموع تغییرات بازار سهام برای ۳ گروه ذی‌نفعان اقتصادی دارای اهمیت است. در شکل (۱) این گروه‌ها طبقه بندی شده‌اند:



شکل ۱: تفاوت ارزش افزوده‌ها در حالت تعدیل شده و نشده

Figure 1: Difference in value added in adjusted and unadjusted form

منبع: عباسی نژاد، محمدی و ابراهیمی، ۱۳۹۴

لذا از طرفی، این بازارها با انتشار سهام، سرمایه‌های مورد نیاز خود را جمع‌آوری می‌کنند و امکان تأمین مالی، گسترش تحقیق و توسعه و سایر ابتکارات را خواهند داشت. همچنین علاوه بر این، بازار سهام، افراد و سرمایه‌گذاران کوچک و بزرگ را قادر می‌سازد تا در خلق ثروت مشارکت کنند و فرصت‌هایی را برای تأثیر بر متغیرهای کلان از طریق تغییر پس‌انداز ایجاد کنند. در واقع این خلق ثروت به این دلیل بروز می‌کند، که بازار سهام با تسهیل خرید و فروش سهام و کشف قیمت، تخصیص کارآمد منابع را تضمین می‌کند. علاوه بر این، بازارهای سهام به سلامت و ثبات کلی اقتصاد ملی کمک می‌کنند. تغییرات در این بازار شاخصی از احساسات جمعی ذی‌نفعان اقتصادی و شرایط کلی اقتصاد است.

بازارهای سهام نقش مهمی در حاکمیت شرکتی ایفا می‌کنند؛ زیرا شرکت‌های سهامی عام مشمول الزامات شفافیت و افشا هستند و باید از پاسخگویی به سهامداران و مردم اطمینان حاصل کنند. این شفافیت، اعتماد سرمایه‌گذاران را تقویت می‌کند، لذا جریان سرمایه به سمت این شرکت‌ها گسیل خواهد شد. این بدان معنا است که عملکرد محیط تجاری بهبود خواهد یافت. در مجموع نظارت و ارزیابی مستمر قیمت سهام و شاخص‌های بازار سهام، بینشی عمیق از وضعیت اقتصادی ارائه می‌کند. از این رو سیاست‌گذاران، اقتصاددانان و تحلیلگران با بهره‌وری بیشتر به تدوین سیاست‌های پولی و مالی مؤثر خواهند پرداخت.

به طور کلی، بازار سهام یکی از اجزای حیاتی اقتصاد ملی است. از طرفی، بروز مسائل اقتصادی ناشی از بازخوردهای یک ساختار سیستمی است؛ لذا همان طور که بازار سهام بر متغیرهای کلان اقتصادی مؤثر است، شرایط کلان اقتصادی نیز بر بازده سهام تأثیر دارد. به گفته مرتون^۱ (۱۹۷۳)، مدل های قیمت گذاری دارایی بر این مفهوم استوارند که هر عاملی که بتواند بر فرصت های سرمایه گذاری آتی تأثیر بگذارد، می تواند به عنوان یک عامل مهم در قیمت گذاری در شرایط تعادلی در نظر گرفته شود. بسیاری از متغیرهای کلان اقتصادی می توانند بر سطوح ریسک مرتبط بر حقوق صاحبان سهام، جریان های نقدی شرکت های متعدد و نرخ تنزیل تعدیل شده تأثیر بگذارند. علاوه بر این، تغییرات شرایط اقتصادی می تواند هم کمیت و هم ماهیت فرصت های سرمایه گذاری را دستخوش تغییر کند.

محققان زیادی تلاش کرده اند تا رابطه های میان متغیرهای کلان اقتصادی و بازده حقوق صاحبان سهام را مورد بررسی قرار دهند. محققان پیشگام در این زمینه عبارتند از نلسون^۲ (۱۹۷۶)، سولنیک^۳ (۱۹۸۳)، چن و همکاران^۴ (۱۹۸۶)، فاما^۵ (۱۹۸۱)، گسک و رول^۶ (۱۹۸۳)، گلتکین^۷ (۱۹۸۳)، و پیرس و رولی^۸ (۱۹۸۳، ۱۹۸۵). مطالعات آنها نشان می دهد که بین بازده سهام و عرضه پول و رشد نرخ تورم، همبستگی منفی وجود دارد. همچنین پارک و راتی^۹ (۲۰۰۸)، ناندا و فاف^{۱۰} (۲۰۰۸)، هامپ و مک میلان^{۱۱} (۲۰۰۹)، ژائو^{۱۲} (۲۰۱۰) و چینزارا^{۱۳} (۲۰۱۱) نیز به بررسی رابطه بین متغیرهای کلان اقتصادی و قیمت سهام پرداخته اند.

در چند دهه اخیر، علاقه پژوهشگران به تحلیل بازار سهام به دلیل در دسترس بودن بهتر داده های آماری افزایش یافته است. علاوه بر این، اقتصادهای در حال توسعه ای همچون چین، هند، برزیل و مالزی رشد اقتصادی سریعی را تجربه کرده اند که از نرخ رشد اقتصادهای توسعه یافته دیگر پیشی گرفته است. از آنجایی که رشد قیمت سهام با رشد بالقوه اقتصادی ارتباط تنگاتنگی دارد، ذی نفعان اقتصادی علاقه زیادی به بازارهای سهام و صندوق های بین المللی نشان داده اند. به عنوان مثال، موری

-
1. Merton (1973)
 2. Nelson (1976)
 3. Solnik (1983)
 4. Chen & et al. (1986)
 5. Fama (1981)
 6. Geske & Roll (1983)
 7. Gultekin
 8. Pearce & Roley (1983 & 1985)
 9. Park & Ratti (2008)
 10. Nandha & Faff (2008)
 11. Humpe & Macmillan (2009)
 12. Zhao (2010)
 13. Chinzara (2011)

و همکاران^۱ (۲۰۲۳)، زاهرا و بنسال^۲ (۲۰۱۸)، دیکسیت و همکاران^۳ (۲۰۲۰)، کومار و همکاران^۴ (۲۰۲۰) و ورما و همکاران^۵ (۲۰۲۱)، به جنبه‌های مختلف بازارهای سهام در اقتصادهای در حال توسعه پرداخته‌اند و از این جهت کمک شایانی به ادبیات موضوعی این حوزه کرده‌اند.

نکته بسیار مهمی که باید در اینجا عنوان شود، این است که همان‌طور که در بخش بعدی دیده خواهد شد، نمی‌توان یک رابطه ثابت و خالص برای اثرگذاری بسیاری از متغیرهای کلان اقتصادی بر شاخص قیمت سهام یافت و این خود باعث می‌شود نتوان روشی متناسب با شرایط اقتصادی برای شناسایی مهم‌ترین متغیرها تعیین کرد. در این راستا در این تحقیق تلاش شده است تا از مدل‌های یادگیری ماشینی برای شناسایی عوامل مهم مؤثر بر شاخص قیمت سهام استفاده شود.

استفاده از یادگیری ماشین برای تخمین تأثیر متغیرهای کلان بر قیمت سهام به چند دلیل مهم است: مدل‌های اقتصادسنجی سنتی اغلب روابط خطی بین متغیرهای کلان اقتصادی و قیمت سهام را فرض می‌کنند. با این حال، بازارهای مالی بسیار پیچیده هستند و الگوهای غیرخطی را نشان می‌دهند. تکنیک‌های ML، مانند شبکه‌های عصبی و درخت‌های تصمیم‌گیری، می‌توانند این روابط غیرخطی را به طور مؤثرتری ثبت، و امکان پیش‌بینی دقیق‌تر را فراهم کنند. از طرفی، بازارهای مالی شامل تعداد زیادی متغیرهای کلان اقتصادی است که می‌تواند بر قیمت سهام تأثیر بگذارد، از جمله نرخ بهره، تورم، رشد تولید ناخالص داخلی، نرخ ارز و رویدادهای ژئوپلیتیکی.

الگوریتم‌های ML در مدیریت داده‌های با ابعاد بالا برتری دارند و آنها را قادر می‌سازد تا طیف وسیع‌تری از عوامل را در نظر بگیرند و پیش‌بینی‌کننده‌های مربوط به حرکت قیمت سهام را شناسایی کنند. از طرفی، بازارهای مالی در واکنش به اطلاعات جدید، احساسات سرمایه‌گذاران و تحولات کلان اقتصادی، پویا، و در معرض تغییرات سریع هستند. مدل‌های ML می‌توانند سریع‌تر از مدل‌های سنتی، خود را با شرایط متغیر بازار تطبیق دهند و به‌روزرسانی‌ها و تنظیمات به‌موقع پیش‌بینی‌ها را امکان‌پذیر می‌سازند.

در واقع، روش یادگیری ماشین بینش‌های مبتنی بر داده است و این تکنیک‌ها می‌توانند الگوها و بینش‌های پنهان را در مجموعه داده‌ها که ممکن است با استفاده از روش‌های آماری سنتی آشکار نباشند، آشکار کند. مدل‌های ML می‌توانند ارزیابی‌های ریسک قوی‌تری را با ترکیب طیف وسیع‌تری از عوامل و گرفتن وابستگی‌های غیرخطی بین متغیرها ارائه کنند و به سرمایه‌گذاران و سیاست‌مداران کمک می‌کنند تا مواجهه خود با نوسانات بازار را بهتر درک و مدیریت کنند.

همان‌طور که بیان شد، هدف از پژوهش حاضر، بررسی ارتباط بین متغیرهای کلان اقتصاد ملی و بازده سهام ایران است. تا به حال هیچ مطالعه دیگری وجود نداشته است که با روش یادگیری

1. Maurya & et al. (2023)
2. Zahera and Bansal (2018)
3. Dixit & et al. (2020)
4. Kumar & et al. (2020)
5. Verma & et al. (2021)

ماشین به بررسی عوامل مختلف بر شاخص قیمت سهام ایران پردازد و همان‌طور که گفته شد، روش تحقیق مورد استفاده در این مطالعه، متمایز از سایر مطالعات پیشین است. نکته بسیار مهم قابل ذکر، انتخاب سال مطالعه بود. سال‌های بعدی پس از این مطالعه، به دلیل انتخاب نشدن. یکی آنکه به دلایلی چون کرونا و حباب‌های موجود در بازار برای جلوگیری از تورش‌دار بودن نتایج از آن سال‌ها صرف نظر شد و دیگر، عدم دسترسی به همه متغیرها برای سال‌های اخیر موجب شد تا تحقیق به محدودیت داده روبرو گردد.

در ادامه مقاله، به شرح زیر، بخش‌های مبانی نظری، داده‌ها و مدل تجربی، روش شناسی و یافته‌های تجربی پرداخته خواهد شد. بخش آخر نیز شامل خلاصه، نتیجه‌گیری، پیامدهای کوتاه سیاست و محدودیت‌های تحقیق است.

۲. مبانی نظری

هدف اصلی این بخش، استخراج حقایقی از ادبیات موضوعی در ارتباط با متغیرهای کلان اقتصادی و اثر آن بر بازار سهام است. نمونه‌هایی از متغیرهای رایج عبارتند از: نرخ بهره، نرخ تورم، نرخ ارز، عرضه پول، تولید صنعتی، قیمت نفت و قیمت طلا (پیرو، ۲۰۱۶؛ و سادورسکی، ۱۹۹۹؛ و ال-شارکاس، ۲۰۰۴؛ و کهو، ۲۰۱۰). شواهد تجربی نشان می‌دهد که نوسانات در بازار دارایی‌های مختلف به شدت دارای همبستگی است. در چنین شرایطی، درک چگونگی روابط بین بازارهای مختلف برای سیاست‌گذاران و سرمایه‌گذاران بسیار مهم است. اتخاذ سیاست‌های مناسب برای توسعه بازارهای مالی و کاهش نوسانات غیرعادی در این بازارها و همچنین تنوع بخشی کارآمد پرتفوی، بدون آگاهی از رفتار و روابط بین دارایی‌های مختلف امکان پذیر نخواهد بود. تحقیقات در مورد رفتار این بازارها و تغییر سایر متغیرهای کلان، می‌تواند بینش روشنی در مورد تغییرات شاخص قیمت سهام ارائه دهد. عدم اطمینان از تغییرات قیمت سهام و نحوه واکنش آنها در شرایط مختلف اقتصادی، یکی از جنبه‌های مهم کل اقتصاد محسوب می‌شود. مهم‌ترین عاملی که سرمایه‌گذاران را برای سرمایه‌گذاری در بازار سهام تشویق می‌کند، شاخص قیمت سهام و اطلاعات مربوط به آن می‌باشد و در نتیجه، طبیعی است که سرمایه‌گذاران به عوامل تأثیرگذار بر آن نیز علاقه مند باشند. به همین دلیل شناخت آن عوامل و نحوه تأثیرگذاری آن برای سیاست‌گذاران اقتصادی از اهمیت بالایی برخوردار است. از لحاظ نظری، ارتباط بین قیمت سهام و متغیرهای کلان اقتصادی را می‌توان با استفاده از مدل‌های مختلفی مانند مدل استاندارد ارزش گذاری سهام، مدل‌های تخصیص پولی و پرتفوی و مدل عرضه و تقاضای کل استاندارد ($AD - AS$) طراحی کرد. در بخش بعدی، مدل استاندارد ارزش گذاری سهام

1. Peiro (2016)
2. Sadorsky (1999)
3. Al-Sharkas (2004)
4. Keho (2010)

برای روشن شدن ارتباط بین اقتصاد کلان و قیمت سهام تشریح خواهد شد. برای درک تأثیر متغیرهای کلان اقتصادی بر شاخص قیمت سهام، بررسی تنزیل جریان‌های نقدی آتی ضروری است. قیمت سهام، هم به جریان پرداخت سود مورد انتظار و هم، به نرخ تنزیل بازار بستگی دارد. در نتیجه، هر متغیر کلان اقتصادی که بر سود سهام یا نرخ تنزیل آتی اثر گذاشت، می‌تواند بر قیمت سهام نیز مؤثر باشد (هامپ و مک میلان، ۲۰۰۹). ارزش فعلی جریان‌های نقدی آتی، یعنی قیمت یک سهم با رابطه (۱) به دست می‌آید.

$$P_t = \frac{E_t(d_{t+1})}{1 + E_tr} + \frac{E_t(P_{t+1})}{1 + E_tr} \quad (۱)$$

در رابطه (۱) ارزش E_t نشان دهنده انتظارات مشروط در مورد تمام اطلاعات موجود در زمان t است. همچنین قیمت واقعی سهام در زمان t ، و $E_t(d_{t+1})$ سود واقعی مورد انتظار هر سهم در پایان سال اول است. از طرفی، $E_t(P_{t+1})$ قیمت مورد انتظار (واقعی) سهم در پایان سال اول، و E_tr نرخ تنزیل (واقعی) یا هزینه سرمایه مورد انتظار (ثابت) تعیین شده توسط بازار است. با نزدیک شدن افق زمانی به بی نهایت، یعنی $T \rightarrow \infty$ از رابطه (۱) به رابطه (۲) خواهیم رسید.

$$P_t = \sum_{i=1}^n \frac{E_t(d_{t+i})}{(1 + E_tr)^i} \quad (۲)$$

بنابراین، قیمت سهام به جریان مورد انتظار پرداخت سود سهام و نرخ تنزیل بازار بستگی دارد. از طرفی، نرخ تنزیل بازار به چندین متغیر کلان اقتصادی حساس است. به عنوان مثال، نرخ تنزیل با سطوح عمومی قیمت‌ها (CPI) و نرخ تورم (درصد تغییر در CPI) مرتبط است. همچنین نرخ ارز نیز می‌تواند بر شاخص قیمت سهام مؤثر باشد. این اثرگذاری، از طریق تغییرات قدرت شرکت‌های موجود در بورس به واسطه تغییر قدرت صادراتی و وارداتی آنان ایجاد می‌شود.

متغیرهای دیگری که به‌طور بالقوه بر قیمت سهام مؤثرند شامل: قیمت طلا و نفت هستند. طلا یک جایگزین سرمایه‌گذاری برای سایر طبقات دارایی از جمله سهام و املاک و مستغلات است. تورم مستمر و بالا، باعث می‌شود تا سرمایه‌گذاری در طلا به عنوان پناهگاه امنی برای حفظ ارزش دارایی انتخاب شود. این موضوع باعث می‌شود که جریان نقدینگی از بازار سهام به سمت طلا گسیل شود که در نتیجه، بر قیمت سهام و کل بازار سهام مؤثر خواهد بود.

از طرفی تورم بالا، باعث می‌شود که شرکت‌ها به دلیل کمبود نقدینگی و هزینه‌های تولیدی بالا، تمایل به استقراض داشته باشند. این بدان معنا است که جریان‌های نقدی آتی کمتر، نرخ‌های تنزیل بالاتر و در نتیجه، کاهش قیمت سهام محتمل‌تر باشد. همچنین قیمت نفت در کشورهای صادرکننده نفت نیز عنصر حیاتی مؤثر بر قیمت سهام است؛ زیرا نفت خام بزرگ‌ترین صادرات و حیاتی‌ترین منبع درآمد صادراتی آنها است.

برای آگاهی در مورد اینکه چه متغیرهایی بر شاخص قیمت سهام اثر دارند و اثرگذاری آنها چگونه است، می‌باید مطالعات فراوانی را مرور کرد. ورما و بانسال^{۱۱} (۲۰۲۱)، خلاصه‌ای جامع از تحقیقات را در مورد بررسی ارتباط بین متغیرهای کلان اقتصادی و عملکرد بازار سهام ارائه کردند. بررسی سیستماتیک آنها بیش از صد مقاله تحقیقاتی را در بر می‌گیرد که نقش متغیرهای کلان اقتصادی در بازارهای سهام را بررسی می‌کند و دوره‌ای تقریباً چهار دهه را پوشش می‌دهد. همان‌طور که بررسی آنها نشان داد، شدت علاقه به این موضوع همچنان بالا است.

چندین مقاله تحقیقاتی کلاسیک وجود دارد که به بررسی ارتباط بین متغیرهای کلان اقتصادی و سهام در اقتصادهای توسعه یافته می‌پردازد.

فلانری و جیمز^۲ (۱۹۸۴) تأثیر تغییرات نرخ بهره اسمی را بر قیمت سهام بررسی کردند. یافته‌های آنها بر اهمیت این ارتباط بسته به ترکیب بلوغ دارایی‌ها و بدهی‌های شرکت‌ها تأکید شد. علاوه بر این، آنها رابطه مثبتی بین اثر و اندازه ترکیب سررسید دارایی‌ها و بدهی‌های اسمی نشان دادند.

به‌طور مشابه، گسک و رول^۳ (۱۹۸۳) و موکرچی و ناکا^۴ (۱۹۹۵)، رابطه منفی بین تغییرات نرخ بهره و بازده سهام در ایالات متحده و ژاپن یافتند. تحقیقات آنها بیشتر از این ایده حمایت می‌کند که نوسانات نرخ بهره بر عملکرد بازار سهام تأثیر می‌گذارد.

چن و همکاران^۵ (۱۹۸۶)، رابطه بین متغیرهای کلان اقتصادی و قیمت سهام را در ایالات متحده مطالعه کردند. تجزیه و تحلیل رگرسیون آنها، از سال ۱۹۵۸ تا ۱۹۸۴، نشان داد که هیچ ارتباط معنی داری بین مخارج مصرف و قیمت نفت با قیمت سهام وجود ندارد. با این حال، آنها یک همبستگی مثبت بین تولید صنعتی ایالات متحده و قیمت سهام، به‌ویژه در دوره‌های نوسانات بازار پیدا کردند. علاوه بر این، این مطالعه رابطه منفی بین نرخ تورم و قیمت سهام را نشان داد.

اخیراً، بولماش و تریولی^۵ (۱۹۹۱)، چن^۶ (۱۹۹۱)، سادورسکی^۷ (۱۹۹۹)، راتاناپاکورن و شارما^۷ (۲۰۰۷)، کیلیان و پارک^۸ (۲۰۰۹)، پارک و راتی^۹ (۲۰۰۸)، رحمان^{۱۰} (۲۰۰۸)، گان و همکاران^{۱۱} (۲۰۰۶)، هامپ و مک میلان^۸ (۲۰۰۹) و پیرو^۹ (۲۰۱۶)، روی این موضوع کار کرده‌اند. روش‌شناسی آنها شامل مدل‌سازی خودرگرسیون برداری (VAR)، آزمون‌های همجمعی، آزمون‌های علیت گرنجر،

11. Verma & Bansal (2021).

2. Flannery & James (1984).

3. Geske & Roll (1983).

4. Mukherjee & Naka (1995).

5. Bulmash & Trivoli (1991)..

6. Chen (1991).

7. Ratanapakorn & Sharma (1999).

8. Kilian & Park (2009).

9. Park & Ratti (2008).

10. Rahman (2008).

11. Gan et al. (2006).

و حالت‌های مختلف *GARCH* بود. خلاصه گسترده یافته‌های آنها نشان می‌دهد که ارتباط بین قیمت یا بازده سهام معمولاً با تولید صنعتی و رشد تولید ناخالص داخلی مثبت و با نرخ‌های بهره، نرخ تورم و رشد عرضه پول منفی است. رابطه با نرخ ارز و بیکاری نامشخص است.

موریتالا و کلیکوم^۱ (۲۰۱۹)، تأثیر قیمت نفت را بر بازارهای سهام آفریقا بررسی کردند. برای این مطالعه، از داده‌های فصلی پنج کشور منتخب تولیدکننده نفت در بازار سهام استفاده شد. تجزیه و تحلیل پانل پویا با ترکیب متغیرهایی مانند بازده سهام، نرخ رشد واقعی تولید ناخالص داخلی، نرخ ارز و قیمت سبد اوپک مورد استفاده قرار گرفت. نتایج نشان داد که قیمت نفت، تأثیر نامطلوبی بر بازارهای سهام آفریقا داشته است که می‌توان آن را به ماهیت پراکنده و توسعه نیافته بازار سرمایه در این منطقه نسبت داد. در همین حال، نرخ رشد واقعی تولید ناخالص داخلی، تأثیر مثبتی بر بازارهای سهام آفریقا داشت که نشان می‌دهد، رشد اقتصادی بر بازده سهام در این بازارها، تأثیر مثبت دارد.

همچنین ارول و آیتکین^۲ (۲۰۱۹)، رابطه بین متغیرهای کلان اقتصادی منتخب و شاخص بورس استانبول (*BIST 100*) را بررسی کردند. متغیرهای کلان اقتصادی که در تحلیل رگرسیونی آنها گنجانده شده، عبارتند از: نرخ تورم، نرخ بهره وام، قیمت طلا، تولید صنعتی و نرخ رشد تولید ناخالص داخلی. در این مطالعه، داده‌های سال‌های ۲۰۰۹ تا ۲۰۱۸ مورد بررسی قرار گرفت. یافته‌های تجربی نشان داد که نرخ بهره وام، تولید صنعتی و نرخ تورم به‌طور آماری بر شاخص *BIST 100* تأثیر می‌گذارد. با این حال، هیچ ارتباط آماری معناداری بین شاخص *BIST 100* و نرخ رشد تولید ناخالص داخلی یا قیمت طلا وجود نداشت.

ابکه و همکاران^۳ (۲۰۲۳)، اثرات شوک‌های مالی قبل و بعد از بحران مالی جهانی را بر رابطه و ریسک سیستماتیک بین سری بازده قیمت نفت اوپک و بازارهای سهام شورای همکاری خلیج فارس بررسی نمودند. محققان از روش‌هایی همچون ارزش در معرض خطر شرطی ناپارامتریک (*NCoVaR*)، انسجام متقاطع چندگانه و مدل‌های اتصال و سرریز دیبولد و ایلماز استفاده کردند. نتایج حاصل از تجزیه و تحلیل نمونه کامل، نشان داد که بحرین تنها بازار سهامی است که سرریز پویا، نوسانات قیمت نفت اوپک را نشان می‌دهد. در مقابل، تمام بازارهای سهام شورای همکاری خلیج فارس، بجز عربستان سعودی، نوسانات پویا را از قیمت نفت اوپک دریافت کردند.

بوید و همکاران^۴ (۲۰۰۱)، استدلال کردند که وقتی نرخ‌های بیکاری اعلام می‌شود، بازارهای سهام واکنش متفاوتی به اخبار نشان می‌دهند. بر اساس گفته او، واکنش بازار، به این بستگی دارد که آیا اقتصاد یک سیر صعودی را طی می‌کند یا در روند نزولی خود به سر می‌برد. به‌طور کلی می‌توان چنین گفت: زمانی که اقتصاد در حال رشد بوده، و سیر صعودی به خود گرفته است، افزایش نرخ‌های

1. Kelikume & Muritala(2019)

2. Erol & Aytekin(2019)

3. Abakah & et al(2023)

4. Boyd & et al(2001)

بیکاری، تأثیر مثبتی بر بازار سهام دارد، اما در دوران رکود اقتصادی، همین اتفاق می‌تواند خبر بد و ناگواری باشد.

همان‌طور که مشخص است، در مطالعات فراوانی، به بررسی رابطه متغیرهای اقتصاد کلان به عنوان ریسک سیستماتیک که قابل حذف نیست و شاخص قیمت سهام پرداخته‌اند؛ اما هیچکدام از پژوهش‌ها از روش‌های یادگیری ماشینی استفاده نکرده‌اند. در نتیجه در این تحقیق، تلاش می‌شود تا از روش یادگیری ماشین برای بررسی عوامل مؤثر بر شاخص قیمت سهام استفاده شود. در عصر تکثیر بی‌سابقه داده‌ها، تلاقی محاسبات پیشرفته، پیچیدگی الگوریتمی و منابع داده‌ای فراوان، پارادایم دگرگون‌کننده‌ای به نام یادگیری ماشینی (ML) ایجاد کرده است.

یادگیری ماشینی، زیرشاخه‌ای از هوش مصنوعی (AI) است که سیستم‌ها را قادر می‌سازد تا الگوها را به‌طور مستقل یاد بگیرند، پیش‌بینی کنند و عملکرد را بدون برنامه‌نویسی صریح بهینه کنند. فرض اساسی آن در استخراج دانش از داده‌ها است، که به الگوریتم‌ها اجازه تعمیم و انطباق با شرایط پیش‌بینی نشده را می‌دهد. این رویکرد در هسته خود، متدولوژی یادگیری ماشینی متنوعی از جمله یادگیری تحت نظارت، یادگیری بدون نظارت، یادگیری تقویتی و یادگیری عمیق را دارا است. این رویکردها ماشین‌ها را برای رمزگشایی الگوهای پیچیده، تشخیص ناهنجاری‌ها و بهینه‌سازی فرایندهای تصمیم‌گیری در طیفی از برنامه‌ها توانمند می‌سازد. یادگیری ماشین قابلیت بررسی تفاوت‌های ظریف انتخاب الگوریتم، آموزش مدل، و تعادل ظریف بین تعصب و واریانس را دارا است. حوزه یادگیری ماشینی، طیف گسترده‌ای از روش‌ها و الگوریتم‌ها را دربر می‌گیرد که هر کدام برای وظایف و اهداف خاصی طراحی شده‌اند. روش‌های مختلف یادگیری ماشینی شامل موارد زیر است:

❖ یادگیری تحت نظارت: رگرسیون خطی، رگرسیون لجستیک^۱

❖ یادگیری بدون نظارت: خوشه بندی $K - Means$ ، خوشه بندی سلسله مراتبی^۲

❖ یادگیری تقویتی: یادگیری Q ، یادگیری تقویتی عمیق^۳

❖ یادگیری عمیق: شبکه‌های عصبی، RNN ، CNN ^۴

❖ یادگیری گروه: جنگل تصادفی، ماشین‌های تقویت کننده گرادیان^۵

❖ کاهش ابعاد: PCA ، $SNE - t$ ^۶

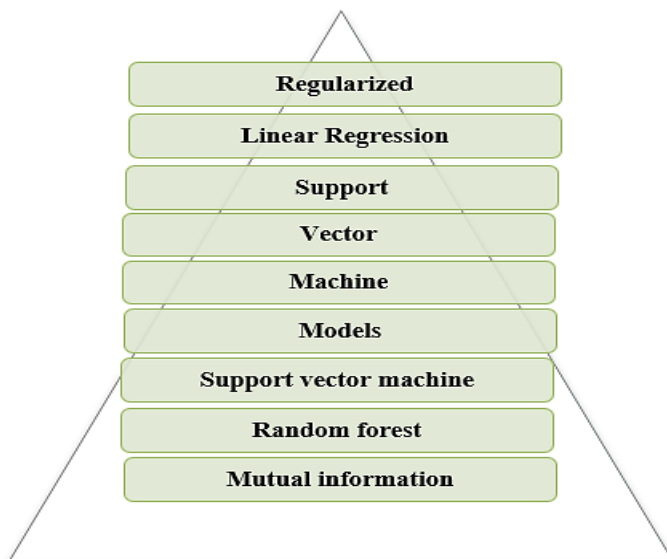
❖ یادگیری مبتنی بر نمونه: $NN - k$ ^۷

-
1. Supervised learning
 2. Unsupervised Learning: K-Means Clustering, Hierarchical Clustering
 3. Reinforcement Learning: Q-Learning, Deep Reinforcement Learning
 4. Deep Learning: Neural Networks, CNNs, RNNs.
 5. Ensemble Learning: Random Forest, Gradient Boosting Machines
 6. Dimensionality Reduction: PCA, t-SNE
 7. Instance-Based Learning: k-NN

- ❖ ماشین‌های بردار پشتیبانی *SVM*: *SVM* خطی، *SVM* هسته^۱
 - ❖ روش‌های بیزی: بیز ساده، شبکه‌های بیزی^۲
 - ❖ نزدیکترین طبقه بندی *Centroid*: طبقه بندی مبتنی بر *Centroid*^۳
 - ❖ الگوریتم‌های تکاملی: الگوریتم‌های ژنتیک^۴
 - ❖ فرا یادگیری: *MAML*^۵
 - ❖ منطق فازی: سیستم‌های فازی^۶
 - ❖ روشهای نزدیک‌ترین همسایه: رگرسیون وزنی موضعی^۷
 - ❖ الگوریتم‌های خوشه بندی: *DBSCAN*^۸
 - ❖ سیستم‌های مبتنی بر قانون: درخت‌های تصمیم، طبقه‌بندی‌کننده‌های مبتنی بر قانون^۹
 - ❖ *AutoML* (یادگیری ماشین خودکار): بهینه سازی فرآپارامتر، مهندسی ویژگی‌های خودکار^{۱۰}
 - ❖ آموزش انتقال: تطبیق دامنه^{۱۱}
- به منظور استفاده مؤثر از روش‌های یادگیری ماشینی، پیش پردازش داده‌ها ضروری است. انتخاب ویژگی، یکی از متداول‌ترین و مهم‌ترین تکنیک‌ها در پیش پردازش داده‌ها است و به یکی از اجزای ضروری فرایند یادگیری ماشین تبدیل شده (کولوسوس و همکاران، ۲۰۰۷) و چندین تکنیک برای کاهش متغیرهای نامربوط و زائد توسعه یافته است. انتخاب ویژگی (حذف متغیر) به درک داده‌ها، کاهش نیاز محاسباتی و بهبود عملکرد پیش بینی کمک می‌کند. این فرایند درواقع شناسایی ویژگی‌های مرتبط و حذف داده‌های نامربوط، اضافی یا نویزدار است. این فرایند الگوریتم‌های داده کاوی را سرعت می‌بخشد، دقت پیش بینی را بهبود می‌بخشد و قابلیت درک را افزایش می‌دهد. ویژگی‌های نامربوط آن‌هایی هستند که هیچ اطلاعات مفیدی ارائه نمی‌دهند و ویژگی‌های اضافی برخلاف قبلی، اطلاعات بیشتری ارائه می‌دهند (مولینا و همکاران ۲۰۰۲).

-
1. Support Vector Machines (SVM): Linear SVM, Kernel SVM
 2. Bayesian Methods: Naive Bayes, Bayesian Networks
 3. Nearest Centroid Classification: Centroid-based Classification
 4. Evolutionary Algorithms: Genetic Algorithms
 5. Meta-Learning: MAML
 6. Fuzzy Logic: Fuzzy Systems
 7. Nearest-Neighbor Methods: Locally Weighted Regression
 8. Clustering Algorithms: DBSCAN
 9. Rule-Based Systems: Decision Trees, Rule-Based Classifiers
 10. Auto ML (Automated Machine Learning): Hyper parameter Optimization, Automated Feature Engineering
 11. Transfer Learning: Domain Adaptation
 12. Kalousis et al. (2007)
 13. Molina et al. (2002)

انتخاب ویژگی یک گام مهم با هدف بهبود عملکرد مدل، تفسیرپذیری و کارایی است. به بیان دیگر، انتخاب ویژگی می‌تواند کارایی یادگیری، دقت پیش بینی را افزایش دهد و پیچیدگی نتایج آموخته شده را کاهش دهد (کولر و سهامی، ۱؛ ۱۹۹۶؛ هال، ۲، ۲۰۰۰). روش‌های مختلفی برای انتخاب ویژگی وجود دارد؛ که در شکل (۲) به آن اشاره شده است:



شکل ۲: روش‌های مختلف انتخاب ویژگی در رویکرد یادگیری ماشینی

Figure 2: Different feature selection methods in machine learning approach

منبع: هونارد، ۱۹۷۰

البته برای تحلیل نتایج هر یک از این روش‌ها، مبانی نظری و دانش آن برای پژوهشگر از ضروریات هر تحقیق است. مبانی نظری به شناسایی مرتبط‌ترین متغیرهای اقتصادی که بر قیمت سهام تأثیر می‌گذارند، مانند نرخ بهره، تورم، رشد تولید ناخالص داخلی، نرخ بیکاری و ... کمک می‌کند. از طرفی، با پایه‌گذاری انتخاب ویژگی در تئوری اقتصادی، می‌توان خطر بیش از حد برآزش را کاهش داد. به بیان دیگر، در این حالت، مدل به جای الگوهای معنادار، سعی می‌کند تا عامل اختلال را یاد گیرد و این یعنی یک تعامل دوطرفه بین تئوری اقتصادی و روش یادگیری ماشین وجود دارد. همچنین تئوری چهارچوبی برای درک اینکه چگونه متغیرهای مختلف احتمالاً با قیمت سهام تعامل دارند، ارائه می‌دهد. این درک می‌تواند به تفسیر خروجی مدل‌های پیچیده یادگیری ماشین کمک کند و مدل‌ها را شفاف‌تر و کاربردی‌تر کند که باعث اعتبار و اعتماد به مدل می‌شود.

از طرفی، مبانی نظری به تعیین انتظارات واقع بینانه برای رفتار مدل کمک می‌کند. به عنوان مثال، در طول یک رکود اقتصادی، می‌توان انتظار داشت که متغیرهای خاصی در شرایط خاصی،

ممکن است تأثیرات متفاوتی بر قیمت سهام داشته باشند که این بینش‌های نظری، امکان تحلیل سناریو و تست استرس بهتر مدل‌ها را در شرایط مختلف اقتصادی فراهم می‌کند و استحکام را بهبود می‌بخشد. در نتیجه، به‌طور خلاصه می‌توان گفت که گنجاندن مبانی نظری در مدل‌های یادگیری ماشین برای پیش‌بینی قیمت سهام برای اطمینان از ارتباط، بهبود تفسیرپذیری، افزایش عملکرد و ارائه شهود اقتصادی ضروری است؛ که شکاف بین بینش‌های مبتنی بر داده و منطق اقتصادی را پر می‌کند و منجر به مدل‌های پیش‌بینی قوی‌تر و معتبرتر می‌شود.

۳. روش‌شناسی پژوهش

براساس مبانی نظری ذکر شده در بخش (۲)، هدف از این پژوهش، بررسی عوامل مؤثر بر شاخص قیمت سهام به کمک رویکرد یادگیری ماشینی است. همان‌طور که در بخش قبل گفته شد، روش‌های مختلفی برای انتخاب ویژگی در رویکرد ML وجود دارد که به برخی از آنان در شکل (۲) اشاره شده است. در این مطالعه، از روش‌های رگرسیون خطی منظم^۱ و جنگل تصادفی^۲ استفاده به عمل آمده است. همچنین پس از آنکه مهم‌ترین متغیرها از طریق این دو روش کشف شد، از طریق معیار میزان وابستگی بین رگرسیون^۳ برای کشف اینکه ارتباط متغیرها با یکدیگر چگونه است، استفاده خواهد شد و در انتها، از روش اطلاعات متقابل نتایج، به تفکیک هر دهه برای بازار بورس ایران مورد بررسی قرار خواهد گرفت. اما در ابتدا نیاز است تا مراحل یافتن متغیرها تشریح، و در انتها روابط بین آنان بررسی شود.

❖ رگرسیون خطی منظم

در واقع در قرن جدید، حجم عظیمی از داده در حال تولید است. بررسی و تجزیه و تحلیل چنین حجم وسیعی از مشاهدات و متغیرها، زمان و هزینه زیادی را می‌طلبد. بنابراین، ساده سازی مسائل به ظاهر پیچیده در تحلیل داده‌ها بسیار ضروری به نظر می‌رسد. تشخیص اینکه کدام پیش‌بینی‌کننده‌ها به یک واکنش خاص مرتبط هستند، آسان نیست. روش RLR تکنیکی است که برای رفع مشکل اضافه برآزش در مدل‌های رگرسیون خطی از آن استفاده می‌شود. برآزش بیش از حد، زمانی اتفاق می‌افتد که یک مدل، بیش از حد پیچیده می‌شود. این موضوع می‌تواند منجر به عملکرد ضعیف مدل گردد.

رویکرد RLR به ایجاد تعادل بین برآزش داده‌های آموزشی و اجتناب از برآزش بیش از حد کمک می‌کند. در واقع با کنترل پارامتر منظم‌سازی، تناسب مدل برقرار خواهد شد و پیچیدگی مدل، کنترل می‌شود. این امر، باعث می‌گردد که پیش‌بینی‌ها قوی‌تر و قابل اعتمادتر شوند. مبانی فکری روش RLR این است که یک متغیر کنترل و تنظیم کننده را به عنوان یک جریمه به تابع ضرر اضافه کند. این جریمه، مدل را از تخصیص وزن‌های بیش از حد بزرگ به ویژگی‌ها منصرف می‌کند و به‌طور مؤثر ضرایب را به سمت صفر کاهش می‌دهد. با انجام این کار، مدل نسبت به نویز در داده‌های آموزشی،

حساسیت کمتری پیدا می‌کند و بهتر می‌تواند به داده‌های جدید تعمیم داده شود. دو نوع از متداول ترین منظم‌سازی در رگرسیون خطی به قرار زیر هستند:

الف) رگرسیون *Ridge* (هول و کنارد، ۱۹۷۰)؛

ب) رگرسیون *Lasso* (پارک و کازلا، ۲۰۰۸).

جریمه ای که *Lasso* استفاده می‌کند، باعث می‌شود، برخی از تخمین‌های ضریب به صفر محدود شود. از طرفی، این روش ابعاد متغیرها را کاهش می‌دهد و باعث می‌شود تا مجموع مربعات اصلاح، و حداقل شود (تبشیرانی، ۱۹۹۶).

در این پژوهش به دلایل ذکر شده، از روش *Lasso* استفاده شده است. بنابراین، تعداد پارامترها با استفاده از تابع جریمه بر روی مجموع مطلق ضرایب مدل رگرسیون کنترل خواهد شد. مجموع مربعات خطای رگرسیون *Lasso* به کمک رابطه (۳) به دست می‌آید.

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p x_{ij}\beta_j)^2 - \lambda \sum_j |\beta_j| \quad (3)$$

در رابطه (۳)، مقدار λ همان پارامتر تنظیم است. اگر $\lambda = 0$ باشد، معادله (۳) به مجموع مجذور خطاها در حداقل مربعات معمولی تبدیل خواهد شد و همه متغیرها در مدل وجود خواهند داشت. با افزایش λ ، تعداد متغیرهای مستقل در مدل کاهش می‌یابد. بنابراین، با انتخاب بی نهایت (∞) برای λ ، عملاً هیچ متغیری در مدل وجود ندارد. مقدار این پارامتر معمولاً با استفاده از روش اعتبارسنجی متقاطع تعیین می‌شود. ۳ گام اصلی برای یافتن نتیجه مطلوب به قرار زیر است:

❖ الگوریتم جنگل تصادفی

الگوریتم جنگل تصادفی (RF)، یک مدل یادگیری ماشینی است که برای وظایفی نظیر دسته بندی^۱ داده‌ها و رگرسیون، از تعداد زیادی درخت تصمیم^۲ استفاده می‌کند. این الگوریتم، اولین بار توسط تین کم هوه در سال ۱۹۹۵ میلادی مورد استفاده قرار گرفت. بریمان^۶ (۲۰۰۱) و کاتلر^۷ و همکاران (۲۰۰۷)، بهبودهایی بر الگوریتم پیشنهادی کم هو (۱۹۹۵) پیشنهاد دادند که باعث شهرت جهانی این الگوریتم شد. اگر در این الگوریتم هدف، دسته بندی یک سطر از ماتریس داده‌ها باشد، دسته آن سطر با توجه به رأی اکثریت در درخت‌های تصمیم انتخاب می‌شود و اگر هدف رگرسیون باشد، نتیجه نهایی به صورت میانگینی از جواب‌های درخت‌های تصمیم مشخص می‌گردد. در این صورت، جواب‌های بهینه با در نظر گرفتن تمام مدل‌های ممکن، مشخص خواهند شد. از آنجایی که جنگل تصادفی از

1. penalty
2. Tibshirani (1996).
3. Classification
4. Decision Tree
5. Tin Kam Ho (1995).
6. Breiman (2001).
7. Cutler et al. (2007).

تعداد زیادی درخت تصمیم برای نتیجه‌گیری استفاده می‌کند، مشکل برازش بیش از حد روی داده‌های مجموعه را حل می‌کند.

❖ اطلاعات متقابل^۲

همان‌طور که ذکر شد، برای برآورد درجه اهمیت هریک از متغیرهای تعیین شده از ادبیات تحقیق بر شاخص سهام، در ابتدا از دو رویکرد بالا استفاده می‌شود؛ اما برای برآورد یک تخمین دقیق‌تر و قابل‌استنادتر، از رویکرد اطلاعات متقابل استفاده می‌شود. همچنین برای اندازه‌گیری وابستگی آماری بین دو متغیر تصادفی، از اطلاعات متقابل (MI) استفاده خواهد شد.

گیو و نیکسون^۳ (۲۰۰۸)، مطرح می‌کنند که اساس اطلاعات متقابل، روش آنتروپی است. این روش در زیر تشریح شده است:

فرض کنید، X یک متغیر تصادفی در فضای (گسسته) Z ، و x یک عنصر از Z باشد. یک بردار تصادفی d بعدی به صورت $(X_1, \dots, X_d) \in Z$ می‌باشد، x هر عدد صحیح مثبت Z می‌باشد. درجه عدم قطعیت در یک متغیر تصادفی در طول یک آزمایش با آنتروپی روی یک فضای گسسته Z به شرح زیر اندازه‌گیری می‌شود.

$$H[X] = - \sum_{x \in Z} pr[X = x] \cdot \log(pr[X = x]) \quad (۴)$$

که در این حالت، عدم قطعیت در مورد ترکیب دو متغیر تصادفی (X, Y) با آنتروپی مشترک آنها بیان می‌شود:

$$H[X, Y] = - \sum_{x \in Z, y \in K} pr[X = x, Y = y] \cdot \log(pr[X = x, Y = y]) \quad (۵)$$

اگر فقط Y تابع قطعی X باشد، آنتروپی ترکیبی همیشه بزرگ‌تر یا مساوی با هر یک از متغیرها است. تساوی در صورتی رخ می‌دهد که دو متغیر مستقل باشند.

$$\text{Max}(H[X], H[Y]) \leq H[X, Y] \leq H[X] + H[Y] \quad (۶)$$

در نهایت، آنتروپی شرطی یک متغیر تصادفی X با توجه به متغیر Y دیگر، عدم قطعیت در X را پس از شناخته شدن Y محاسبه می‌کند.

$$H(X|Y) = - \sum_{x \in Z, y \in K} pr[X = x, Y = y] \cdot \log(pr[X = x, Y = y]) \quad (۷)$$

برخلاف آنتروپی گسسته، آنتروپی دیفرانسیل می‌تواند منفی باشد. مقدار $\text{Pr}[X = x]$ تابع چگالی احتمال متغیر پیوسته X را نشان می‌دهیم، تا بتوانیم سناریوهای ترکیبی را با ترکیب متغیرهای گسسته و پیوسته به راحتی مدیریت کنیم. اندازه‌گیری کلی بین دو متغیر تصادفی، همان اطلاعات متقابل (MI)

است. اطلاعات متقابل برابر با مقدار دانشی است که شخص در مورد X از طریق مشاهده Y به دست می آورد. اطلاعات متقابل دو متغیر تصادفی X و Y در یک حوزه گسسته، به صورت زیر تعریف می شود:

$$I(X, Y) = \sum_{X \in Z, Y \in K} pr[X = x, Y = y] \cdot \log \left(\frac{pr[X = x, Y = y]}{pr[X = x] \cdot pr[Y = y]} \right) \quad (۸)$$

وقتی X و Y مستقل از یکدیگر هستند، اطلاعات متقابل همیشه بزرگ تر یا مساوی صفر است و در نتیجه انتروپی مجموع آن ها از انتروپی هر یک از متغیرها کمتر است (برابری این انتروپی ها تنها زمانی رخ می دهد که یکی از متغیرها تابع قطعی دیگری باشد). هر چه اطلاعات متقابل بیشتری وجود داشته باشد، به همان اندازه آماره MI بزرگ تر و در نتیجه، عدد وابستگی بین X و Y قوی تر است.

❖ معیار میزان وابستگی بین رگرسیون

همان طور که گفته شد، در ابتدا درجه اهمیت هر یک از متغیرهای مهم مدل شناسایی خواهند شد. در دو گام قبل، تلاش شد که با بهره گیری از پایه های آماری مختلف و دو رویکرد رگرسیون خطی منظم و جنگل تصادفی، متغیرهای مهم شناسایی شوند و بعد به کمک روش اطلاعات متقابل، ضرایب قابل استنادتری ارائه شد. در این مرحله، تلاش می شود تا هرگونه ارتباطی که بین متغیرهای مستقل وجود دارد، مورد توجه قرار گیرد. یکی از روش هایی که در سال های اخیر برای این مهم ایجاد شده است، استفاده از معیار میزان وابستگی بین رگرسیون ها است.

در این مقاله، از روش دوپلهوفر و ویکس^۱ (۲۰۰۹)، پیروی می شود. همچنین در این تحقیق، فقط روابط مشترک بین جفت متغیرها مورد بررسی قرار خواهد گرفت؛ هر چند می توان روابط بین سه و بیشتر متغیرها را جست؛ اما این خود نیاز به تعداد زیادی از مشاهدات دارد که این تحقیق به دلیل ماهیت داده ها، فاقد آن است. برای مدل M_j ، احتمالات پسین باید به صورت زیر تعریف شوند:

$$(M_j | y) = p(\varphi_1 = w_1, \varphi_2 = w_2, \dots, \varphi_K = w_K | y, M_j) \quad (۹)$$

در صورتی که متغیر در مدل وجود داشته باشد، مقدار w_K را می توان ۱ و در غیر این صورت ۰ در نظر گرفت. احتمال ترکیبی پسین برای افزودن دو متغیر مدل، X_i و X_h برابر است با:

$$p\left(i \cap h | y\right) = \sum_{j=1}^{2^k} 1(\varphi_i = 1 \cap \varphi_h = 1 | y, M_j) \cdot p(M_j | y) \quad (۱۰)$$

در رابطه بالا، مقدار $p(i \cap h | y)$ ، مجموع احتمال پیشین مدلهایی است که شامل متغیرهایی X_i و X_h بوده اند. همچنین به کمک معیار میزان وابستگی بین رگرسیونی، می توان رابطه مکمل یا جایگزین بودن متغیرها را مورد بررسی قرار داد. دوپلهوفر و ویکس (۲۰۰۹)، در جدول (۱)، نقاط جرم احتمالی را تعریف می کند.

جدول ۱: نقاط جرم احتمالی تعریف شده در فضای $\{0,1\}^2$ برای توزیع یکنواخت

$$P(\varphi_i, \varphi_j | y)$$

Table 1: Probable mass points defined in the space $\{0,1\}^2$ for a uniform distribution

$$P(\varphi_i, \varphi_j | y)$$

$P(\varphi_i, \varphi_j y)$	$\varphi_h = 1$	$\varphi_h = 0$	SUM
$\varphi_i = 1$	$p(i \cap h y)$	$p(i \cap h' y)$	$p(i y)$
$\varphi_i = 0$	$p(i' \cap h y)$	$p(i' \cap h' y)$	$p(i' y)$
SUM	$p(h y)$	$p(h' y)$	1

منبع: دوپلهوفر و ویکس، ۲۰۰۹

با توجه به جدول (۱)، احتمال پسین ترکیبی، افزودن دو متغیر مدل به صورت رابطه (۱۱) است.

$$j_{ih} = \ln \left[\frac{P(i \cap h | y)}{P(i \cap h' | y)} \cdot \frac{P(i' \cap h' | y)}{P(i' \cap h | y)} \right] = \ln [O_{i|h} \cdot PO_{i|h'}] \quad (11)$$

در رابطه (۱۱)، دو مقدار $PO_{i|h}$ و $PO_{i|h'}$ به ترتیب، نشان دهنده احتمال پسین وجود x_i مشروط به درج x_h و احتمال پسین حذف x_i مشروط بر حذف x_h هستند. در جدول (۲)، خلاصه بازه به دست آمده از این معادله بیان شده است:

جدول ۲: طبقه بندی رابطه بین متغیرها

Table 2: Classification of the relationship between variables

$> +2j_{ih}$	مکمل‌های قوی
$+1 < j_{ih} < +2$	مکمل‌های معنی دار
$-1 < j_{ih} < +1$	بدون رابطه
$-2 < j_{ih} < -1$	جایگزین‌های معنا دار
$< -2j_{ih}$	جایگزین‌های قوی

منبع: همان، ۲۰۰۹

همان‌طور که گفته شد، در این مطالعه از روش یادگیری ماشینی برای بررسی تعاملات بین متغیرهای مختلف کلان اقتصادی، از جمله تورم، رشد اقتصادی، نرخ ارز، صادرات، واردات، قیمت طلا، قیمت نفت اوپک و رشد پول گسترده (MG) در رابطه با شاخص قیمت سهام کشور ایران استفاده شده است. سه متغیر اول ((CPI)، تورم و رشد اقتصادی) شاخص‌های رایجی هستند که وضعیت اقتصاد واقعی را نشان می‌دهند. گنجاندن نرخ ارز، صادرات و واردات در تحلیل بسیار مهم است؛ زیرا کشورهای منتخب به شدت به نفت متکی هستند و کل صادرات و ثبات اقتصادی آنها به تجارت بین‌المللی وابسته است. همچنین MG به عنوان نماینده بخش پولی در این مطالعه دیده می‌شود.

علاوه بر این، قیمت نفت (اوپک) و طلا، نشان دهنده متغیرهایی خارج از کنترل هر کشور به صورت جداگانه است. نوسانات قیمت این دو کالا می‌تواند به طور قابل توجهی بر وضعیت اقتصادی

کشورهای صادرکننده نفت تأثیر بگذارد. معادله (۲) مبنای آزمون‌های تجربی انجام شده در این تحقیق است. بر اساس یافته‌های ادبیات، هیچ مجموعه قطعی یا به اتفاق آرا از متغیرهای کلان اقتصادی که بر قیمت‌ها یا شاخص‌های سهام اثرگذار باشد، وجود ندارد. با این حال، اکثر مطالعات بررسی شده به ترکیبی از متغیرهای موجود در معادله (۱۲) اشاره کردند.

جدول (۳) برگرفته از مطالعه ورمان و بانسل (۲۰۲۱) می‌باشد که خلاصه ای عالی از متغیرهای کلان اقتصادی در بیش از ۱۰۰ مقاله، اهمیت آماری آنها و جهت ارتباط با قیمت سهام یا شاخص‌ها را ارائه می‌دهد که در این مطالعه، از آنها استفاده شده است.

$$SPI = f(INF, GP, World un, UNE, R, FD, Open, MG, EG, OPEC, IMP, EXP) \quad (12)$$

متغیرهای مدل در زیر تعریف شده‌اند. علائم مثبت یا منفی داخل پرانتز به جهت ارتباط یک متغیر با قیمت سهام یا شاخص اشاره دارد.

جدول ۳: تعریف متغیرهای رابطه (۱۲)

Table 3: Definition of variables in equation (12)

علامت گذاری در متن	تعریف	جهت علامت
<i>SPI</i>	شاخص سهام	-
<i>INF</i>	تورم	منفی یا گاه مثبت
<i>GP</i>	قیمت طلا	منفی
<i>World un</i>	نااطمینانی جهانی	منفی
<i>UNE</i>	بیکاری	مثبت و منفی
<i>R</i>	تفاوت نرخ سود تسهیلات و سود سپرده که جایگزین نرخ بهره در ایران است.	منفی و مثبت
<i>FD</i>	توسعه مالی	مثبت و منفی
<i>Open</i>	باز بودن تجاری	مثبت و منفی
<i>MG</i>	نرخ رشد پول گسترده	مثبت و منفی
<i>EG</i>	نرخ رشد اقتصادی	مثبت
<i>OPEC</i>	قیمت نفت اوپک	برای کشورهای صادرکننده نفت اثر مثبت و برای کشورهای واردکننده نفت اثر منفی دارد.
<i>IMP</i>	شاخص قیمت واردات براساس سال پایه ۲۰۱۰	منفی یا بسیار ناچیز
<i>EXP</i>	شاخص قیمت صادرات براساس سال پایه ۲۰۱۰	مثبت

منبع: ورمان و بانسل، ۲۰۲۱

چنانچه پیش‌تر گفته شد، ادبیات زیادی در مورد رابطه بین قیمت سهام و متغیرهای کلان اقتصادی صورت گرفته است، اما هیچ‌کدام به این مهم که کدام یک اندازه اثر بیشتری دارند و نیز کدام مدل‌ها با حضور چه متغیرهایی بیشترین قدرت توضیح دهنده‌گی دارند را اشاره نکرده‌اند. در این راستا در این پژوهش، تأثیر ۱۲ متغیر کلان اقتصادی بر شاخص قیمت سهام بورس اوراق بهادار ایران مورد ارزیابی قرار گرفته است.

داده‌های این مطالعه، دوره ۱۹۹۰ تا ۲۰۲۰ را پوشش می‌دهد. منابع داده شامل بانک جهانی، صندوق بین‌المللی پول و اطلاعات بانک مرکزی ایران است. نکته مهم در این پژوهش، سال‌های انتخاب داده‌ها بودند که به دلیل محدودیت در داده‌ها محققان نمی‌توانستند به برخی از داده‌های سالیان قبل از ۱۹۹۰ و بعد از ۲۰۲۰ دسترسی پیدا کنند.

شایان ذکر است که در این تحقیق، تفاوت نرخ سود تسهیلات و سود سپرده به عنوان پراکسی برای نرخ بهره آورده شده است نه خود نرخ بهره؛ به این دلیل که برای سرمایه‌گذاران فقط نرخ بهره سپرده‌گذاری مهم نیست، بلکه همان‌طور که مشخص است، یکی از دیگر از عوامل کلیدی در سرمایه‌گذاری دریافت تسهیلات بانکی است. در نتیجه، تفاوت بین نرخ بهره سپرده و نرخ بهره تسهیلات می‌تواند بر رفتار پس‌اندازکنندگان و سرمایه‌گذاران تأثیر بگذارد. اگر تفاوت این دو نرخ ناچیز باشد یا نرخ بهره سپرده بالاتر باشد، مردم ممکن است ترجیح دهند که پول خود را (اعم از دریافتی به عنوان تسهیلات از بانک و پول‌های شخصی) در بانک‌ها پس‌انداز نمایند تا اینکه در بازارهای مالی همچون بازار اوراق بهادار سرمایه‌گذاری کنند. این امر خود می‌تواند تقاضا برای اوراق بهادار را کاهش دهد و بر قیمت سهام تأثیر منفی بگذارد.

از طرفی، تفاوت بین نرخ بهره سپرده و نرخ بهره تسهیلات بانکی می‌تواند انتظارات تورمی را تحت تأثیر قرار دهد؛ به این معنا که اگر نرخ بهره تسهیلات بسیار بالاتر از بهره سپرده‌ها باشد، ممکن است نشانه‌ای از انتظارات تورمی بالاتر باشد که می‌تواند بازار سهام را تحت تأثیر قرار دهد. از دیگر نکات مهم تفاوت نرخ بهره سپرده و نرخ بهره تسهیلات، این است که می‌تواند بر درآمد خالص بانک‌ها تأثیر گذارد. اختلاف بین این دو نرخ، مارجین بهره‌ای خالص بانک‌ها را تشکیل می‌دهد که درآمد اصلی آن‌ها از تفاوت بین هزینه پرداخت بهره به سپرده‌گذاران و درآمد بهره از وام‌گیرندگان است. درآمد بالاتر بانک‌ها می‌تواند منجر به بهبود ارزش سهام آن‌ها شود. با این تفاسیر، بسیار مهم است که فقط به نرخ بهره سپرده به تنهایی نگاه نشود و باید این نرخ‌ها هردو، در تأثیرگذاری بر بازارهای مالی مورد توجه واقع شوند.

۴. یافته‌های پژوهش

براساس بخش (۳) در همین مقاله، در ابتدا از روش‌های الگوریتم جنگل تصادفی و رگرسیون خطی منظم برای بررسی درجه اهمیت متغیرهای مؤثر بر شاخص قیمت سهام استفاده خواهد شد و سپس از معیار میزان وابستگی بین رگرسیونی، رابطه دو به دو بین متغیرهای مستقل مورد بررسی قرار می‌گیرد و در نهایت، از روش اطلاعات متقابل (MI) برای دقیق شدن در تخمین ضرایب برای هر دهه استفاده می‌شود. در جدول (۴)، نتایج حاصل از روش جنگل تصادفی ملاحظه می‌شود.

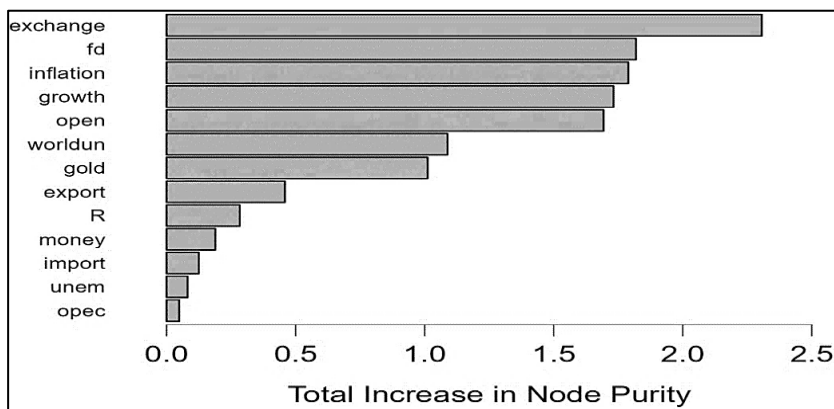
جدول ۴: اهمیت متغیرها براساس روش جنگل تصادفی

Table 4: Importance of variables based on the random forest method

درجه اهمیت	متغیر
۲/۳۰۶	EX
۱/۸۱۹	FD
۱/۷۸۹	INF
۱/۷۳۲	EG
۱/۶۹۳	Open
۱/۰۸۹	World un
۱/۰۱۱	GP
۰/۴۵۹	EXP
۰/۲۸۳	R
۰/۱۸۹	MG
۰/۱۲۵	IMP
۰/۰۸۱	UNE
۰/۰۴۹	OPEC

مأخذ: یافته‌های پژوهش، ۱۴۰۳

مقدار MSE در برآورد درجه اهمیت متغیرها به کمک رویکرد جنگل تصادفی، برابر ۰/۰۵۲ و مقدار R^2 هم برابر ۰/۹۰۶ است. جدول (۴) بیانگر اهمیت اثرگذاری متغیرهای مورد مطالعه بر شاخص قیمت سهام است. برای تصمیم‌گیری در مورد اندازه اهمیت هر یک از این متغیرها، بهتر است به شکل (۳) مراجعه شود. براساس شکل (۳)، هر متغیری که تأثیرگذاری آن بیشتر از نصف مقیاس باشد، از اهمیت بالایی برخوردار است.



شکل ۳: اهمیت تأثیرگذاری متغیرهای مورد بررسی

Figure 3: The importance of the influence of the variables under study

مأخذ: یافته‌های پژوهش، ۱۴۰۳

براساس شکل (۳)، متغیرهای نرخ ارز، توسعه مالی، تورم، رشد، بازبودن تجاری و نااطمینانی جهانی، از اهمیت ویژه‌ای برخوردارند. این توالی اصلاً تعجب برانگیز نیست؛ به آن دلیل که برای مثال، ایران در دوره‌های مختلف در اقتصاد، شاهد تغییرات نرخ ارزهای بالا و تورم‌های فزاینده بوده، و از طرفی، به دلایل سیاسی، همواره نااطمینانی‌های جهانی، تأثیر بالایی بر بازارهای داخلی ایران داشته است. به‌طور دقیق‌تر، مطالعات مختلف نشان داده‌اند که میان تغییرات نرخ ارز و شاخص قیمت سهام، روابط مختلفی وجود دارد (راتاناکورن و شارما، ۲۰۰۷).

محققان و اقتصاددانان استدلال کرده‌اند که ارتباط بین قیمت سهام و نرخ ارز به چند دلیل قابل توجه است. به طور خاص، گاوین^۲ (۱۹۸۹)، نشان داد که این پیوند ممکن است بر سیاست‌های پولی و مالی تأثیر بگذارد، زیرا زمانی که بازار سهام در حال رکود است، بر تقاضای کل تأثیر مثبت می‌گذارد. در این مطالعه، بیان شد که رابطه بین قیمت سهام و تورم به‌طور گسترده در پژوهش‌های مختلفی بررسی شده است و آن را از متغیرهای مهم تأثیرگذار بر شاخص سهام معرفی نموده‌اند؛ مانند جاف و ماندکلر^۳ (۱۹۷۶)؛ فاما و شورت (۱۹۷۷) و بودوخ و ریچاردسون^۴ (۱۹۹۳).

هر چند در برخی از این پژوهش‌ها، به رابطه مثبت میان تورم و قیمت سهام اشاره شده است و در برخی دیگر، رابطه منفی میان این دو مورد تأکید بوده است. از طرفی پذیرفته شده است که یک اقتصاد باز به دلیل افزایش اشتراک ریسک بین‌المللی در بین بازارها، با تعداد بیشتری از شوک‌های نامطلوب مواجه می‌شود و بنابراین، باز بودن تجارت هنگام انتقال نوسانات بین کشورها مهم است و نقش مهمی در پیش‌بینی نوسانات بازار سهام دارد. این در حالی است که در برخی مطالعات، اثر معین و مشخصی بین بازبودن تجاری و قیمت سهام یافت نشده، اما در میان عوامل تأثیرگذار بوده است (نیکمنش^۵، ۲۰۱۶).

از طرفی، قیمت سهام برابر با ارزش جریان‌های نقدی آتی است که در معرض دو اثر متفاوت است، یکی، جریان نقدی که تحت تأثیر افزایش تولید است و دیگری، نرخ بهره که عامل تنزیل است. زمانی که جریان نقدی مورد انتظار کاهش یابد یا نرخ بهره افزایش یابد، قیمت سهام کاهش می‌یابد. سطح فعالیت واقعی اقتصادی مستقیماً بر جریان نقدی تأثیر می‌گذارد و فعالیت اقتصادی بالاتر به معنای جریان نقدی بالاتر است. متغیرهای زیادی برای پیش‌بینی بازار کالا وجود دارند. در این میان، تولید ناخالص داخلی واقعی به عنوان جامع‌ترین معیار موجود برای فعالیت اقتصادی است (یوهاسین^۶، ۲۰۱۱؛ کریستوفرگان و همکاران^۷، ۲۰۰۶).

1. Ratanapakorn & Sharma (2007).

2. Gavin (1989).

3. Jaffe & Mandelker (1976).

4. Boudoukh & Richardson (1993).

5. Nikmanesh

6. Yu Hsing (1011).

7. Gan et al. (2006).

نااطمینانی جهانی نیز به نوبه خود، تأثیر مهمی بر وضعیت داخلی کشورها می‌گذارد و به همین دلیل، به‌عنوان یک متغیر کلان جهانی، نقش مهمی را برای شرکت‌های پذیرفته شده در بورس ایفا می‌کند. در جدول (۴)، ضرایب برای هریک از این متغیرها با استفاده از روش رگرسیون خطی منظم به دست آمده است. همان‌طور که گفته شد، روش‌های یادگیری ماشینی، این امکان را به محقق می‌دهد تا از خطای بیش برآزش، رگرسیون‌های کاذب یا تمامی مسائل مربوط به مشکلات اقتصادسنجی کلاسیک رهایی یابد.

جدول ۵: ضرایب برآوردی در رگرسیون خطی منظم

Tabel 5: Estimated coefficients in regular linear regression

متغیرها	ضرایب
INF	۰/۰۰۹
EX	۰/۹۲۱
EXP	۰/۰۰۰
IMP	۰/۰۰۰
EG	۰/۰۳۱
MG	۰/۰۰۰
OPEC	۰/۰۰۰
GP	۰/۰۰۰
World un	-۰/۰۳۶
UNE	۰/۰۰۰
Open	-۰/۰۰۳
R	-۰/۰۳۰
FD	۰/۱۶۸

مأخذ: یافته‌های پژوهش، ۱۴۰۳

همچنین مقدار آماره MSE برابر ۰/۰۵ است. از طرفی، مقدار R^2 برابر ۰/۹۴ است. همان‌طور که مشخص است، براساس جدول (۵)، نااطمینانی جهانی، نرخ بهره و باز بودن تجاری، اثر منفی بر شاخص قیمت بورس اوراق بهادار تهران دارد. همچنین متغیرهایی که فاقد معنی‌داری‌اند، شامل صادرات، واردات، رشد پول، قیمت نفت اوپک و قیمت طلا است. مابقی متغیرها دارای اثری مثبت بر شاخص قیمت سهام ایران می‌باشند. این یافته نیز تا حد بالایی، روش قبل را تأیید، و از طرفی، ضرایب و نوع اثرگذاری را مشخص می‌کند.

در ادامه، بااستفاده از روش JW رابطه بین متغیرها مورد بررسی قرار گرفته است. همان‌طور که گفته شد، این روش راهی برای شناسایی رفتار متغیرهای وابسته در قبال یکدیگرند. نحوه شناسایی این ارتباطات براساس جدول (۲) مشخص می‌شوند. خلاصه جدول (۶) در زیر آن آمده، و ارتباط متغیرهای مستقل از جهت مکمل یا جانشین بودن بیان شده است. این روش می‌تواند الگویی مکمل برای متدهای انتخاب ویژگی باشد؛ زیرا می‌توان با استفاده از این روش، متغیرهای تحت تحقیق را کاهش داد. در این مطالعه، برای مثال طبق یافته‌ها، نرخ بهره و نرخ ارز می‌توانند جایگزین هم شوند،

و یا توسعه مالی و صادرات کاملاً مکمل هم هستند؛ به این معنا هر جا در مدلی از صادرات استفاده شد، می‌باید توسعه مالی نیز در مدل لحاظ گردد.

جدول ۶: نتایج حاصل از روش JW

Tabel 6: Results from the JW method

MG	GE	IMP	EXP	EX	INF	
-۰/۱۳۳	۰/۲۳۲	-۰/۲۳۹	-۰/۲۵۲	-۰/۴۰۹	۱	INF
-۰/۲۶۸	-۰/۹۲۹	-۰/۰۸۳	۰/۵۵۳	۱	-۰/۴۰۹	EX
-۰/۵۹۹	-۰/۰۲۴	-۰/۷۱۹	۱	۰/۵۵۳	-۰/۲۵۲	EXP
-۰/۲۲۸	-۰/۰۷۳	۱	-۱/۷۱۹	-۰/۰۸۳	-۰/۳۳۹	IMP
۰/۰۵۰	۱	-۰/۰۷۲	-۰/۰۲۴	-۰/۹۲۹	۰/۲۳۲	GE
۱	۰/۰۵۰	-۰/۲۲۸	-۰/۵۹۹	-۰/۲۶۸	-۰/۱۳۳	MG
-۰/۳۵۳	۰/۰۰۱	-۱/۳۵۸	-۰/۶۲۴	۰/۵۳۲	۰/۲۷۷	OPEC
۰/۰۷۴	-۰/۰۳۴	-۰/۷۴۴	-۰/۶۶۲	۰/۱۴۶	۰/۰۴۴	GP
۰/۰۹۷	۰/۱۰۷	-۰/۱۴۵	-۰/۲۲۶	-۱/۰۳۹	-۰/۰۳۵	World up
-۰/۰۰۶	-۰/۰۵۲	-۰/۱۵۵	۰/۰۳۹	-۰/۵۳۰	-۰/۰۱۶	UNE
۰/۲۴۵	-۰/۰۳۱	-۰/۹۶۱	-۱/۱۴۳	-۰/۱۹۵	-۰/۶۹۷	Open
-۰/۳۵۲	۰/۱۲۱	-۱/۲۳۷	-۱/۴۹۶	-۱/۹۲۴	۰/۶۲۶	R
FD	R	Open	UNE	World up	GP	
۰/۲۲۱	۰/۶۲۶	-۰/۶۹۷	-۰/۰۱۶	-۰/۰۳۵	۰/۰۴۴	INF
۱/۸۸۱	-۱/۹۲۴	-۰/۱۹۵	-۰/۵۳۰	-۱/۰۳۹	۰/۱۴۶	EX
۱/۳۹۷	-۱/۴۹۳	-۱/۱۴۳	۰/۰۳۹	-۰/۲۲۶	-۰/۶۶۲	EXP
۰/۷۹۶	-۱/۲۳۷	-۰/۹۶۱	-۰/۱۵۵	-۰/۱۴۵	-۰/۷۴۴	IMP
۰/۰۹۲	۰/۱۲۱	-۰/۰۳۱	-۰/۰۵۲	۰/۱۰۷	-۰/۰۳۴	GE
۱/۰۰۱	-۰/۳۵۲	۰/۲۴۵	-۰/۰۰۶	۰/۰۹۷	۰/۰۷۴	MG
۱/۲۰۵	-۱/۴۵۴	-۱/۱۰۲	۰/۰۳۰	-۰/۱۳۳	-۰/۵۱۹	OPEC
-۰/۵۴۳	-۰/۵۳۲	-۰/۰۴۱	۰/۰۰۷	-۰/۱۲۰	۱	GP
۱/۲۲۵	-۰/۱۱۹	۰/۰۴۹	-۰/۰۶۹	۱	-۰/۱۲۰	World up
۰/۲۹۷	۰/۱۲۸	۰/۱۰۳۹	۱	-۰/۰۶۹	۰/۰۰۷	UNE
۰/۴۲۷	-۰/۰۷۱۵	۱	۰/۰۳۹	۰/۰۴۹	-۰/۰۴۱	Open
۱/۲۹۹	۱	-۰/۷۱۵	۰/۱۲۸	-۰/۱۱۹	-۰/۵۳۲	R

مأخذ: یافته‌های پژوهش، ۱۴۰۳

جدول ۷: جایگزینی معنادار

Tabel 7: s significance replacement

جایگزینی معنادار	متغیرها
-۱/۹۲۴	EX, R
-۱/۸۸۱	EX, FD
-۱/۷۱۹	EX, IMP
-۱/۶۲۴	EXP, OPEC
-۱/۴۹۶	EXP, R
-۱/۴۵۴	OPEC, R
-۱/۳۵۸	IMP, OPEC
-۱/۲۳۷	IMP, R
-۱/۲۲۵	World un, FD

مأخذ: یافته‌های پژوهش، ۱۴۰۳

جدول ۸: مکملی معنادار

Tabel 8: significance Complementary

متغیرها	جایگزینی معنادار
EX , Open	-۱/۱۴۳
Open, Open	-۱/۱۰۲
EX, World un	-۱/۰۳۹
ME, FD	-۱/۰۰۱
OPEC, FD	۱/۲۰۵
R, FD	۱/۲۲۹
EXP, FD	۱/۳۹۷

مأخذ: یافته‌های پژوهش، ۱۴۰۳

در ادامه، روش MI برای شناسایی ارتباط بین متغیرهای مستقل با متغیر وابسته به تفکیک دهه‌های مختلف مورد استفاده قرار گرفت. نتایج این روش، نشان می‌دهد که در دهه ۷۰ متغیرهای کمتری بر شاخص قیمت سهام تأثیرگذار، و نسبت به سایر دهه‌های دیگر از طرفی تأثیرگذاری برخی متغیرها در دهه ۹۰ بسیار قابل توجه است. برای مثال، ناپایداری جهانی و رشد پول در این دهه، تأثیر بسیار پررنگی را نشان می‌دهد. همچنین توسعه مالی نیز به بیشترین درجه اهمیت نسبت به سایر دهه‌ها رسیده است. از طرف دیگر، تأثیر واردات در حال از دست رفتن می‌باشد. همچنین همین اتفاق برای صادرات در حال رخ دادن است و دیگر صادرات توانایی تأثیر بر شاخص سهام را ندارد. به بیان دیگر، هرچقدر اقتصاد از ثبات برخوردار باشد همچون دهه ۷۰، متغیرهای کمتری بر شاخص سهام تأثیرگذار است. اما با تلاطم‌های اقتصادی، ترکیبات متغیرها پیچیده‌تر و عمیق‌تر می‌شود. برای مثال در دهه ۷۰ از میان متغیرهای اثرگذار تنها دو متغیر نرخ بهره و نرخ ارز، بیشترین تأثیر را بر شاخص قیمت سهام دارد، که این خود نشان دهنده تأثیر عمیق بخش پولی بر بخش مالی کشور بوده است. در حالی که در دهه ۸۰ صادرات و بیکاری و نرخ ارز، این نقش را داشتند؛ زیرا در آن دوره، دولت، شروع صادرات و اتصال به سمت بازارهای جهانی را در دستور کار خود قرار داده بود.

جدول ۹: نتایج حاصل از روش MI

Tabel 9: Results from the MI method

متغیر	درجه اهمیت 1368-1378	درجه اهمیت 1378-1388	درجه اهمیت 1389-1399
EX	0.51	0.45	0.34
FD	0.35	0.21	0.48
INF	0.0	0.0	0.0
EG	0.0	0.0	0.0
Open	0.0	0.01	0.02
World un	0.0	0.10	0.16
GP	0.05	0.16	0.04
EXP	۰.15	0.29	0.0
R	0.53	0.07	0.16
MG	۰.0	0.0	0.27
IMP	0.14	0.43	0.06
UNE	0.35	0.52	0.22
OPEC	0.0	0.36	0.13

مأخذ: یافته‌های پژوهش، ۱۴۰۳

در نتیجه به‌طور خلاصه می‌توان چنین گفت که سیاست‌گذاران اقتصادی، نه تنها می‌باید به وضعیت کلی ارتباط متغیرهای مورد مطالعه با شاخص بورس توجه نمایند بلکه باید برنامه‌ای مدون و دقیق برای پیشبرد هرچه بهتر سیاست‌هایشان تدوین کنند؛ زیرا سیاست‌های انتخابی آنها در هر دهه می‌تواند باعث تأثیر عمیق یکی یا چند شاخص مهم اقتصادی بر بازار بورس شود که با شناسایی و در نظر گرفتن آنها در برنامه خود، به نتایج هرچه بهتر و مطلوب مد نظر خود دست یابند.

همان‌طور که در نمودار ۱ مشخص است، روند حرکتی متغیرهای مهم یافت شده در سه دهه به تفکیک نشان داده شده‌اند. مشخص است که روند حرکتی بازبودن تجاری، نرخ ارز و نااطمینانی، بسیار شبیه روند حرکتی شاخص قیمت سهام بوده است.

۶. بحث و نتیجه‌گیری

استفاده از تکنیک‌های یادگیری ماشین برای پیش‌بینی تأثیر متغیرهای کلان اقتصادی بر قیمت سهام، نشان‌دهنده پیشرفت قابل‌توجهی در تحلیل مالی و تصمیم‌گیری است. از طریق استفاده از الگوریتم‌های مختلف یادگیری ماشین مانند رگرسیون، درخت‌های تصمیم، جنگل‌های تصادفی و شبکه‌های عصبی، محققان و سرمایه‌گذاران می‌توانند الگوها، روندها و همبستگی‌هایی را که ممکن است از طریق روش‌های سنتی به آسانی قابل مشاهده نباشند، کشف کنند. این امکان پیش‌بینی دقیق‌تر و به‌موقع حرکت قیمت سهام را بر اساس تغییرات عوامل اقتصاد کلان مانند نرخ بهره، تورم، رشد تولید ناخالص داخلی و نرخ ارز فراهم می‌کند. از طرفی، این مدل‌ها در بازارهای مالی که بازارهای به‌شدت پویا هستند، بسیار سریع‌تر و بهتر از مدل‌های سنتی اقتصاد سنجی کلاسیکی، خود را تطبیق می‌دهند و به‌روزرسانی می‌کنند.

در این راستا در این پژوهش، سعی شد تا عوامل مؤثر کلان بر شاخص قیمت سهام ایران به وسیله متدهای یادگیری ماشین همانند رگرسیون خطی منظم، جنگل تصادفی و اطلاعات متقابل برای سال‌های ۱۹۹۰ الی ۲۰۲۰ مورد سنجش قرار گیرد. انتخاب این سال‌ها به این دلیل بود که در سال‌های اخیر، همه‌گیری کرونا و حباب بازار، امکان منحرف کردن نتایج را داشتند. در این پژوهش، ابتدا به وسیله RF و RLR متغیرهای کلیدی شناسایی شدند و در ادامه به وسیله JW ارتباطات متغیرهای مستقل شناسایی شد. در نهایت برای هر دهه به وسیله MI تأثیر متغیرها بر قیمت سهام به تفکیک مورد بحث قرار داده شد.

نتایج پژوهش حاضر، نشان داد که دو روش RLR و RF هردو تأیید می‌کنند که از میان متغیرهای مورد مطالعه، تنها متغیرهای نرخ ارز، توسعه مالی، تورم، رشد، بازبودن تجاری و نااطمینانی جهانی، در تأثیرگذاری بر شاخص قیمت سهام، از اهمیت ویژه‌ای برخوردارند؛ به‌گونه‌ای که نااطمینانی جهانی، نرخ بهره و بازبودن تجاری، اثر منفی بر شاخص قیمت بورس اوراق بهادار تهران دارند و مابقی متغیرهای یاد شده، اثر مثبت دارند. در ادامه با استفاده از روش JW رابطه بین متغیرها مورد بررسی قرار گرفته، که نتایج آن به تفصیل در جدول ۶ بیان شده است.

از دیگر نتایج مهمی که در این مطالعه به آن اشاره شد، استفاده از روش MI برای بررسی تأثیرگذاری متغیرهای مهم بر شاخص قیمت سهام به تفکیک هر دهه بود. نتایج نشان داد که در دهه‌های متفاوت، متغیرهای مختلفی بر شاخص قیمت سهام تأثیرگذار هستند.

سپاسگزاری: -

تأییدیه‌های اخلاقی: موردی توسط نویسنده گزارش نشده است.

تعارض منافع: نویسندگان هیچ‌گونه تعارض منافع در تحقیق ندارند.

سهام نویسندگان در مقاله: در فرایند این پژوهش این تحقیق، تمامی نویسندگان مشارکت داشتند.

منابع مالی / حمایت‌ها: موردی توسط نویسنده گزارش نشده است.

References

- Abbasinejad, H., Mohammadi, S., & Ebrahimi, S. (2017). Dynamics of the relation between macroeconomic variables and stock market index. *Journal of Asset Management and Financing*, 5(1), 61-82.
doi: 10.22108/amf.2017.21153. [In Persian]
- Hatamerad, Saman, Adrangi, Bahram, Asgharpur, Hossein, & Haghighat, Jafar. (2023). Investigating the Impact of Macroeconomic Factors on the Iranian Stock Price Index by Using Averaging Methods. *Iranian Journal of Economic Research*, 28(95), 193-236. magiran.com/p2631396. [In Persian]
- Hatamerad, S., Asgharpur, H., Adrangi, B. et al. (2024). Stock price index analysis of four OPEC members: A Bayesian approach. *Financ Innov.*, 10, 134.
- Samadi, Saeed, & Bayani, Ozra. (2011). The relationship between macroeconomic variables and stock returns in the Tehran stock exchange. *Journal of Financial Economics (Financial Economics and Development)*, 5(16). [In Persian]
- Abakah, E. J. A., Tiwari, A. K., Alagidede, I. P., & Hammoudeh, S. (2023). Nonlinearity in the causality and systemic risk spillover between the OPEC oil and GCC equity markets: A pre-and post-financial crisis analysis. *Empirical Economics*, 1-77. <https://ideas.repec.org/a/spr/empeco/v65y2023i3d10.1007>
- Al-Sharkas, A. (2004). The dynamic relationship between macroeconomic factors and the Jordanian stock market. *International Journal of Applied Econometrics and Quantitative Studies*, 1, 1.
- Boyd J.H., Jagannathan R., & Hu, J. (2001). The stock market's reaction to unemployment news: Why bad news is usually good for stocks? *NBER Working Paper* 9182.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Bulmash, S. B., & Trivoli, G. W. (1991). Time-lagged interactions between stocks prices and selected economic variables. *The Journal of Portfolio Management*, 17(4), 61-67.
- Chen, N. F. (1991). Financial investment opportunities and the macroeconomy. *The Journal of Finance*, 46(2), 529-554.
- Chen, N. F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business*, 383-403. <https://www.jstor.org/stable/2352710>
- Chinzara, Z. (2011). Macroeconomic uncertainty and conditional stock market volatility in South Africa". *South African Journal of Economics*, 79(1), 27-49.
- Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783-2792.
- Dixit, A., Mani, A. & Bansal, R. (2020). Depsosvm: Variant of differential evolution based on PSO for image and text data classification. *International Journal of Intelligent Computing and Cybernetics*, Vol. 13, No. 2, 223- 238. <https://doi.org/10.1108/IJICC-01-2020-0004>

- Doppelhofer, G., & Weeks, M. (2009). Jointness of growth determinants. *Journal of Applied Econometrics*, 24(2), 209-244.
- Erol, A. F., & Aytekin, S. (2019). The role of macroeconomic indicators on stock index evidence from Turkey. *Global Journal of Economics and Business Studies*, 8(15), 61-68.
- Fama, E. F. (1981). Stock returns, real activity, inflation, and money. *The American Economic Review*, 71(4), 545-565.
- Flannery, M. J., & James, C. M. (1984). The effect of interest rate changes on the common stock returns of financial institutions. *The Journal of Finance*, 39(4), 1141-1153.
- Gan, C., Lee, M., Yong, H. H., and Zhang, J. (2006). Macroeconomic variables and stock market interactions: New Zealand evidence. *Investment Management and Financial Innovations*, 3(4), 89-101
- Geske, R., & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *The Journal of Finance*, 38(1), 1-33.
- Geske, R., & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *The Journal of Finance*, 38(1), 1-33.
- Gultekin, N. B. (1983). Stock market returns and inflation. Evidence from other countries. *The Journal of Finance*, 38, 49-65.
- Guo, B., & Nixon, M. S. (2008). Gait feature subset selection by mutual information. *IEEE Transactions on Systems, MAN, and Cybernetics-part a: Systems and Humans*, 39(1), 36-46.
- Hall, M. (2000). *Correlation-based Feature Selection of Discrete and Numeric Class Machine Learning*. in Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000), Stanford University, Stanford, CA, USA, June 29 - July 2, 2000.
- Humpe, A., & Macmillan, P. (2009). Can macroeconomic variables explain longterm stock market movements? A comparison of the US and Japan. *Applied Financial Economics*, 19(2), 111-119.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer.
- Kalousis, A., Prados, J., & Hilario, M. (2007). Stability of feature selection algorithms: a study on high-dimensional spaces. *Knowledge and information systems*, 12, 95-116.
- Keho, Y. (2010). Budget deficit and economic growth: Causality evidence and policy implications for WAEMU countries. *European Journal of Economics, Finance and Administrative Sciences*, 18, 99-104.
- Kelikume, I., & Muritala, O. (2019). The impact of changes in oil price on stock market: Evidence from Africa. *International Journal of Management, Economics and Social Sciences (IJMESS)*, 8(3), 169-194.
- Kilian, L. & C. Park. (2009). The impact of oil price shocks on the US stock market. *International Economic Review*, 50(4), 1267-1287.

- koller, D., and Sahami, M. (1996). Toward Optimal Feature Selection. in *International Conference on Machine Learning*. Stanford, CA: Stanford InfoLab.
- Kumar D., Kumar S., and Bansal R.(2020). Multi-objective multi-join query optimisation using modified grey wolf optimization. 67-79 <https://doi.org/10.1504/IJAIP.2020.108760>.
- Maurya, P.K., Bansal, R. and Mishra, A.K. (2023).Russia-Ukraine conflict and its impact on global inflation: An event study-based approach. *Journal of Economic Studies*, Vol. ahead-of-print No. ahead-of print. <https://doi.org/10.1108/JES-01-2023-0003>
- Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 41, 867-887.
- Molina, L. C., Belanche, L., & Nebot, À. (2002, December). Feature selection algorithms: A survey and experimental evaluation. In *2002 IEEE International Conference on Data Mining, 2002. Proceedings.* (pp. 306-313). IEEE.
- Mukherjee, T. K., & Naka, A. (1995). Dynamic relations between macroeconomic variables and the Japanese stock market: An application of a vector error correction model. *Journal of Financial Research*, 18(2), 223-237.
- Nandha, M., & Faff, R. (2008). Does oil move equity prices? A global view. *Energy economics*, 30(3), 986-997.
- Nelson, C. R. (1976). Inflation and rates of return on common stocks. *Journal of Finance*, 31(2): 471-483
- Nikmanesh, L. (2016). Trade openness and stock market volatility in the ASEAN-5 countries: New evidence using SUR. *Investment Analysts Journal*, 45(3), 194-211.
- Park, J. W., & Ratti, R. A. (2008). Oil price shocks and stock markets in the U.S. and 13 European countries. *Energy Economics*, 30(5): 2587-2608.
- Pearce, D. K., & Roley, V. V. (1983). The reaction of stock prices to unanticipated changes in money: A note. *The Journal of Finance*, 38(4), 1323-1333.
- Pearce, D. K., & Roley, V. V. (1985). Stock prices and economics news. *Journal of Business*, 58, 49-67.
- Peiro, A. (2016). Stock prices and macroeconomic factors: Some European evidence. *International Review of Economics & Finance*, 41, 287-294.
- Rahman, M. (2008). Influences of money supply and oil price on US stock market. *North American Journal of Finance and Banking Research*, 2(2).
- Ratanapakorn, O., & Sharma, S. C. (2007). Dynamic analysis between the US stock returns and the macroeconomic variables. *Applied Financial Economics*, 17(5), 369-377.
- Sadorsky P. (1999). Oil price shocks and stock market activity. *Energy Economics*, 2, 449-469.
- Solnik, B. (1983). International arbitrage pricing theory. *The Journal of Finance*, 38(2), 449-457.

- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series*. No. 1, 267-288.
- Verma R. K. , A. Kumar A. and Bansal R. (2021). Impact of COVID-19 on Different Sectors of the Economy Using Event Study Method: An Indian Perspective. pp. 109-120 | Published online: <https://doi.org/10.1080/10599231.2021.1905492>.
- Verma, R.K. & Bansal, R. (2021). Impact of macroeconomic variables on the performance of stock exchange: A systematic review. *International Journal of Emerging Markets*, Vol. 16 No. 7, 1291-1329. <https://doi.org/10.1108/IJOEM11-2019-0993>
- Zahera, S.A. & Bansal, R. (2018). Do investors exhibit behavioral biases in investment decision making? A systematic review. *Qualitative Research in Financial Markets*, Vol. 10 No. 2, 210-251. <https://doi.org/10.1108/QRFM-04-2017-0028>
- Zhao, H. (2010). Dynamic relationship between exchange rate and stock price: Evidence from China. *Research in International Business and Finance*, 24(2), 103-112.