

طراحی مدل پیش-بینی مبادلات سهام با استفاده از متغیرهای کلان اقتصادی و شبکه-های اجتماعی

نوع مقاله: پژوهشی

معصومه آقاپور علی شاهی^۱

فرید عسکری^{۲*}

یداله رجائی^۳

عبدالرحیم هاشمی دیزج^۴

تاریخ پذیرش: ۱۴۰۵/۲/۱۶

تاریخ دریافت: ۱۴۰۴/۱۱/۱۷

چکیده

بازارهای مالی، به دلیل ویژگی‌های ذاتی خود نظیر پیچیدگی، پویایی و وابستگی به عوامل مختلف اقتصادی و اجتماعی، نیازمند روش‌های تحلیلی پیشرفته برای درک روندها و پیش-بینی رفتارهای آتی هستند. از این-رو، این پژوهش با هدف ارائه الگویی برای پیش-بینی مبادلات سهام با استفاده از متغیرهای کلان اقتصادی و داده-های شبکه-های اجتماعی، با بهره‌گیری از روش‌های خطی لاسو و فرآیند غیرخطی گوسی، به طراحی پرداخت. در این مطالعه، اطلاعات متغیرهای کلان اقتصادی و شبکه-های اجتماعی در بازه زمانی سال‌های ۱۳۹۶ الی ۱۴۰۲ درباره ۱۵ سهم پذیرفته‌شده در بورس اوراق بهادار تهران مورد بررسی قرار گرفت. نتایج این بررسی نشان داد ترکیب متغیرهای کلان اقتصادی شامل نرخ بهره، نرخ تورم، نرخ ارز، تولید ناخالص داخلی و داده-های شبکه-های اجتماعی، دقت پیش-بینی مبادلات سهام را به‌طور معناداری بهبود می‌بخشند. همچنین مشخص شد قدرت روش‌های «یادگیری ماشین» جهت پیش-بینی مبادلات سهام، بیشتر از روش خطی لاسو است. دانش‌افزایی این پژوهش علاوه بر طراحی مدل پیش-بینی مبادلات سهام و ارزیابی قدرت روش‌های یادگیری ماشین، در حیطه غنای ادبیات این حوزه می‌باشد.

واژه‌های کلیدی: مبادلات سهام، متغیرهای کلان اقتصادی، شبکه-های اجتماعی، مدل-های پیش-بینی.

طبقه-بندی JEL: G17, C53, C45

۱ گروه اقتصاد و مالی، واحد ابهر، دانشگاه آزاد اسلامی، ابهر، ایران
m.aghapouralishahi@gmail.com

۲ گروه اقتصاد و مالی، واحد ابهر، دانشگاه آزاد اسلامی، ابهر، ایران (نویسنده مسئول)
4410247239@iaui.ac.ir

۳ گروه اقتصاد و مالی، واحد زنجان، دانشگاه آزاد اسلامی، زنجان، ایران
dr.yadollah.rajaei@gmail.com

۴ گروه علوم اقتصادی، دانشگاه محقق اردبیلی، اردبیل، ایران
a.hashemi@uma.ac.ir

مقدمه

بازارهای مالی بعنوان یکی از ارکان اصلی نظام اقتصادی، نقش بی‌بدیلی در تخصیص منابع، توزیع ریسک و ارتقای سطح کارایی اقتصادی ایفا می‌کنند. در این میان، بازار سهام به عنوان بازاری پویا و پیچیده بستری مناسب برای تحلیل رفتارهای اقتصادی و اجتماعی به شمار می‌آید. یکی از شاخص‌های کلیدی در حیطه بازارهای سهام، مبادلات آن است که بعنوان نمایه‌ای از وضعیت سلامت بازار و اعتماد سرمایه‌گذاران شناخته می‌شود. پیش‌بینی دقیق مبادلات سهام نه تنها قادر به کاهش ریسک و بهبود تصمیم‌گیری سرمایه‌گذاران است، بلکه به شفافیت و کارایی بازار نیز منجر می‌شود. در این میان، از یک سو متغیرهای کلان اقتصادی نظیر نرخ بهره، نرخ تورم، تولید ناخالص داخلی و نرخ ارز به دلیل تأثیرات گسترده و پیچیده خود بر بازارهای مالی، از دیرباز مورد توجه تحلیل‌گران بوده است. این متغیرها نه تنها به عنوان نشانگرهایی برای سلامت اقتصادی یک کشور عمل می‌کنند، بلکه به‌طور مستقیم و غیرمستقیم بر تصمیم‌گیری‌های سرمایه‌گذاران اثر می‌گذارند. از سویی دیگر، نتایج پژوهش‌های انجام شده طی سالهای اخیر نشان می‌دهد، شبکه‌های اجتماعی قادر است بر مبادلات سهام تأثیر بگذارد. نتایج بررسی بولن و همکاران^۱ (۲۰۱۱) درخصوص احساسات کاربران توئیتر مشخص نمود که تغییرات در حال و هوای عمومی می‌تواند شاخص‌های بازار را پیش‌بینی نماید. نمونه‌ای از تأثیر شبکه‌های اجتماعی «پدیده گیم استاپ»^۲ است که در مطالعه ژلودیو و همکاران^۳ (۲۰۱۴) به آن اشاره شده بود. بررسی این پژوهشگران نشان داد، هماهنگی کاربران «ردیت»^۴ موجب افزایش ۲۰۰ برابری حجم مبادلات این سهم شده است. در این رویداد که سال ۲۰۲۱ به وقوع پیوست، گروهی از کاربران پلتفرم ردیت تصمیم به خرید انبوه سهام شرکت گیم استاپ گرفتند. این حرکت، به دلیل وجود فروش استقرای قابل توجهی بود که بر روی سهام این شرکت، توسط صندوقهای پوشش ریسک^۵ آغاز شد. هدف کاربران ردیت ایجاد یک فشار فروش استقرای^۶ بود که در نتیجه آن سرمایه‌گذاران مجبور شوند برای بازپس‌گیری سهام استقرای خود، سهام بیشتری خریداری نمایند. این استراتژی بی‌سابقه نقطه عطفی در بازارهای مالی به شمار می‌رود؛ زیرا نشان داد شبکه‌های اجتماعی چگونه می‌توانند رفتار بازار را تحت تأثیر قرار دهند. در این رویداد، قیمت سهم از حدود ۱۷ دلار در ابتدای ژانویه ۲۰۲۱ به بیش از ۳۴۷ دلار در پایان همان ماه در سال ۲۰۲۱ رسید و حجم مبادلات روزانه تا دویست برابر میانگین روزانه افزایش یافت. بنابراین، ترکیب داده‌های ساختاریافته، همانند شاخص‌های کلان اقتصادی و غیرساختاریافته همچون داده‌های شبکه‌های اجتماعی، به دلیل ماهیت متفاوت آن‌ها می‌تواند چالش‌برانگیز باشد. علیرغم چالش‌های پیش روی، این تحول، می‌تواند فرصت بینظیری برای پیش‌بینی مبادلات سهام از طریق ادغام متغیرهای کلان اقتصادی و شبکه‌های اجتماعی را فراهم آورد. در راستای رسیدن به این مهم، شاخه‌ای از علوم کامپیوتر و آمار که به طراحی الگوریتم

1 Bollen et al

2 GameStop

3 Zheludev et al

4 Reddit

5.Hedge Funds

6 Short Squeeze

می‌پردازد با عنوان «ماشین یادگیری^۱» وجود دارد که بسیار کمک کننده است. این شاخه از علم، روشهای پیشرفتهای را با خود به همراه دارد. همانند روشهای رگرسیون خطی لاسو^۲ و فرآیند گوسی^۳ که در مسائل پیش‌بینی و مدلسازی کاربردی استفاده می‌شود. پژوهش یو و همکاران^۴ (۲۰۱۳) نشان داد، ترکیب مدل‌های یادگیری عمیق نظیر شبکه‌های حافظه بلندمدت^۵ و شبکه‌های عصبی پیچشی^۶، میتواند دقت پیش‌بینی بازار سهام را به میزان قابل توجهی افزایش دهد. علاوه بر این، بررسی ژنگ و همکاران^۷ (۲۰۱۱) نشان داد، داده‌های اقتصادی معمولاً با تأخیر زمانی منتشر می‌شوند، درحالیکه داده‌های شبکه‌های اجتماعی به‌صورت لحظه‌ای تولید می‌شوند. به همین دلیل، استفاده از مدل‌های پیشرفته برای پوشش این چالش ضروری است. باتوجه به ادبیات مطرح شده، این پژوهش به دنبال پاسخ به این سوال است که الگوی مناسب برای پیش‌بینی مبادلات سهام با رویکرد متغیرهای کلان اقتصادی و شبکه‌های اجتماعی چیست؟

۱. مبانی نظری و مروری بر پیشینه پژوهش

۱-۱. متغیرهای کلان اقتصادی و تأثیر آن بر مبادلات بازار سهام

عوامل کلان اقتصادی مانند نرخ بهره، نرخ ارز، نرخ تورم و تولید ناخالص داخلی، تأثیر مستقیمی بر رفتار سرمایه‌گذاران و الگوهای معاملاتی دارند. مطالعات فاما^۸ (۱۹۷۰) نشان داد، افزایش نرخ بهره به کاهش نقدینگی در بازارهای سهام منجر می‌شود. این شاخص اقتصادی نه تنها بر هزینه فرصت سرمایه‌گذاری تأثیر می‌گذارد، بلکه از طریق تغییرات در رفتار سرمایه‌گذاران و نهادهای مالی می‌تواند ارزش و مبادلات سهام را نیز تحت تأثیر قرار دهد (مشکین^۹، ۲۰۱۹). بر اساس نظریه جریان نقد تنزیل شده^{۱۰}، نرخ بهره به عنوان یک عامل کلیدی در تعیین ارزش فعلی سهام و مبادلات آن نقش قابل توجهی دارد. با افزایش نرخ بهره، نرخ تنزیل نیز افزایش می‌یابد که این امر منجر به کاهش ارزش فعلی جریان‌های نقدی آتی شرکت‌ها می‌شود. بنابراین، افزایش نرخ بهره معمولاً باعث کاهش ارزش مبادلات و کاهش مبادلات سهام است (داموداران^{۱۱}، ۲۰۲۰). پژوهش‌های انجام‌شده در بازار سهام ایالات متحده نشان داد، افزایش نرخ بهره فدرال رزرو، منجر به کاهش شاخص‌های بازار و حجم مبادلات سهام می‌شود (برنانک^{۱۲} و کاتنر^{۱۳}، ۲۰۰۵). از این‌رو، تغییرات نرخ بهره به عنوان یکی از ابزارهای سیاست پولی، تأثیرات گسترده‌ای بر بازار سهام دارد. نتایج

1. Machine Learning

2. LASOO

3. Gaussian Process

4. Yu et al

5. Long Short-Term Memory (LSTM)

6. Convolutional Neural Networks (CNN)

7. Zhang et al

8. Fama

9. Mishkin

10. Discounted Cash Flow Theory

11. Damodaran

12. Bernanke

13. Kuttner

پژوهش‌های فوق‌نشان می‌دهد که افزایش نرخ بهره معمولاً با کاهش ارزش و کاهش مبادلات سهام همراه است.

نرخ تورم یکی دیگر از متغیرهای اقتصاد کلان است که قادر است بر تصمیم‌گیری سرمایه‌گذاران اثرگذار باشد. نرخ تورم از طریق عواملی همچون تغییرات در نرخ بازده واقعی، هزینه‌های تولید و انتظارات سرمایه‌گذاران، باعث تغییر در ارزش و مبادلات سهام می‌شود (فاما، ۱۹۷۰). دیگر متغیر اثرگذار در میان مولفه‌های اقتصاد کلان، تولید ناخالص داخلی است. تولید ناخالص داخلی تأثیر عمیقی بر فعالیت‌های اقتصادی و بازارهای مالی دارد. تغییرات در این شاخص می‌تواند بر رفتار سرمایه‌گذاران، ارزش‌گذاری سهام و مبادلات بازار سهام اثرگذار باشد (مانکیو^۱، ۲۰۲۰). براساس نظریه رشد اقتصادی، افزایش تولید ناخالص داخلی نشان‌دهنده رشد اقتصادی است که معمولاً منجر به بهبود عملکرد شرکت‌ها، افزایش سودآوری و در نتیجه افزایش قیمت سهام می‌شود. از سوی دیگر، کاهش تولید ناخالص داخلی اغلب با رکود اقتصادی همراه است که باعث کاهش مبادلات سهام و کاهش قیمت آن می‌شود (رومر^۲، ۲۰۱۹). همچنین براساس نظریه چرخه‌های تجاری، تغییرات در تولید ناخالص داخلی منجر به نوسانات قیمتی و حجم مبادلات بازار سهام است و منجر به افزایش اعتماد سرمایه‌گذاران و افزایش فعالیت‌های تجاری می‌گردد. بالعکس در دوره‌های رکود، کاهش تولید ناخالص داخلی باعث افزایش فروش و کاهش مبادلات سهام می‌شود (لوکاس^۳، ۱۹۷۷).

پژوهش چن و همکاران^۴ (۲۰۱۲) که با تحلیل داده‌های جهانی انجام پذیرفت نشان داد، نوسانات نرخ ارز می‌تواند رفتار سرمایه‌گذاران خارجی را به‌شدت تحت تأثیر قرار دهد، این تغییرات به‌ویژه در بازارهای نوظهور که نرخ ارز بی‌ثبات‌تر است، بیشتر مشاهده می‌شود. تغییرات و نوسانات نرخ ارز، از طریق تأثیرگذاری بر درآمد صادراتی شرکت‌ها، هزینه‌های وارداتی و انتظارات سرمایه‌گذاران، قادر است مبادلات سهام را تحت تأثیر قرار دهد (مادورا^۵، ۲۰۲۱). براساس نظریه ترجیحات سرمایه‌گذاران، نوسانات نرخ ارز می‌تواند سطح ریسک و بازده مورد انتظار سرمایه‌گذاران را تغییر دهد. افزایش نوسانات نرخ ارز معمولاً با افزایش عدم اطمینان در بازار همراه است که ممکن است موجب کاهش ارزش مبادلات سهام می‌گردد (آدلر^۶ و دumas^۷، ۱۹۸۴). این نظریه بیان می‌کند، تغییرات نرخ ارز می‌تواند مزیت رقابتی شرکت‌ها در بازارهای بین‌المللی را تغییر دهد. به عنوان مثال، افزایش ارزش پول ملی، ممکن است درآمد صادراتی شرکت‌ها را کاهش دهد و بر سودآوری آن‌ها تأثیر منفی بگذارد. این موضوع می‌تواند منجر به کاهش ارزش سهام شرکت‌ها و کاهش مبادلات سهام گردد (شاپیرو^۸، ۲۰۲۰).

- 1 Mankiw
- 2 Romer
- 3 Lucas
- 4 Chen et al
- 5 Madura
- 6 Adler
- 7 Dumas
- 8 Shapiro

۱-۲. شبکه های اجتماعی و تاثیر آن بر مبادلات بازار سهام

در سال های اخیر، تحلیل محتوای شبکه های اجتماعی به عنوان یک ابزار جدید و مؤثر در پیش بینی روندهای بازار سهام مطرح شده است. شبکه های اجتماعی به دلیل گستردگی و دسترسی آسان، به منابع عظیمی از اطلاعات و احساسات عمومی تبدیل شده اند که می توانند تأثیرات زیادی بر رفتار بازار سهام داشته باشند (بولن^۱ و همکاران، ۲۰۱۱). بر اساس نظریه رفتار سرمایه گذاران، تصمیمات خرید و فروش سرمایه گذاران تحت تأثیر احساسات و هیجانات بازار است. احساسات منفی در شبکه های اجتماعی می تواند موجب وحشت و فروش سهام و احساسات مثبت موجب خرید سهام شود. این تغییرات احساساتی می توانند به نوسانات کوتاه مدت در قیمت های سهام منجر شود (شیلر^۲، ۲۰۰۰). نتایج برخی تحقیقات نشان می دهد، تغییر احساسات جمعی می تواند به طور قابل توجهی نوسانات بازار را پیش بینی کند و به تحلیل گران امکان می دهد تا روندهای بازار را پیش بینی کنند (تتلاک^۳، ۲۰۰۷). داده های غیررسمی مانند محتوای شبکه های اجتماعی، به عنوان یک منبع اطلاعاتی جدید، می توانند به پیش بینی کنندگان بازار سهام کمک کنند. در این زمینه، تحلیل احساسات کاربران شبکه های اجتماعی و تعاملات آنها می تواند به شناسایی روندهای قیمتی یا نوسانات مبادلات کمک کند (چن^۴ و همکاران، ۲۰۱۴). تحقیقاتی که در بازار سهام ایالات متحده انجام شد، نشان داد که احساسات مثبت در توئیتر و دیگر شبکه های اجتماعی می تواند به افزایش قیمت سهام منجر شود (بولن و همکاران، ۲۰۱۱). همچنین، مطالعاتی که در بازارهای جهانی صورت گرفته نشان می دهد، حجم تعاملات و گفتگوهای آنلاین می تواند پیش بینی کننده تغییرات قیمت ها و مبادلات سهام در کوتاه مدت باشد (دوان^۵ و همکاران، ۲۰۱۹). در زمان هایی که احساسات عمومی منفی است، سرمایه گذاران ممکن است به طور ناخودآگاه به فروش سهام بپردازند، این امر می تواند باعث کاهش قیمت ها شود. برعکس، احساسات مثبت می تواند منجر به خرید بیشتر سهام و افزایش قیمت ها و مبادلات سهام گردد (لیو^۶ و همکاران، ۲۰۱۸). حجم تعاملات و سرعت انتشار اخبار در شبکه های اجتماعی می تواند بر روند مبادلات سهام در بازار تأثیرگذار باشد. در مواردی که حجم تعاملات به طور ناگهانی افزایش می یابد، این تغییر نشان از جابجایی های عمده در نگرش عمومی دارد و می تواند به عنوان یک شاخص برای پیش بینی روندهای بازار در نظر گرفته شود (ژنگ^۷ و همکاران، ۲۰۱۱). تحلیل محتوای شبکه های اجتماعی به عنوان یک ابزار نوین، پتانسیل بالایی در پیش بینی روندهای کوتاه مدت بازار سهام دارد. با استفاده از اطلاعات به دست آمده از احساسات کاربران و حجم تعاملات، می توان سیگنال هایی برای پیش بینی تغییرات مبادلات سهام شناسایی کرد.

1 Bollen

2 Shiller

3 Tetlock

4 Chen

5 Duan

6 Liu

7 Zhang

۱-۳. ترکیب داده‌های اقتصادی و شبکه‌های اجتماعی

پیش‌بینی دقیق روند بازار سهام، یکی از مهمترین چالش‌های سرمایه‌گذاران و تحلیل‌گران به شمار می‌رود. لذا استفاده از داده‌های کلان اقتصادی و اطلاعات موجود در شبکه‌های اجتماعی به عنوان منابع رسمی و غیررسمی برای پیش‌بینی تحولات بازار در این بررسی مورد توجه قرار گرفته است. ترکیب این دو منبع ممکن است به بهبود دقت پیش‌بینی‌های مبادلات بازار سهام کمک کند، زیرا داده‌های کلان اقتصادی در گذشته رخ داده است و به‌طور سنتی شاخص‌هایی از وضعیت کلی اقتصاد و روندهای بلندمدت هستند، در حالی که داده‌های شبکه‌های اجتماعی در حال حاضر است و می‌تواند اطلاعات به‌روز و احساسات کوتاه‌مدت سرمایه‌گذاران را در همین لحظه نشان دهند (ژنگ و همکاران، ۲۰۱۸).

از یک سو در بازارهای نوظهور مانند ایران، استفاده از ترکیب داده‌های کلان اقتصادی و داده‌های شبکه‌های اجتماعی در پیش‌بینی مبادلات بازار می‌تواند دقت پیش‌بینی نوسات بازار را افزایش دهد. از سویی دیگر، بحران‌های اقتصادی به‌عنوان دوره‌های پرتلاطم و ناپایدار اقتصادی، معمولاً نوسانات شدید در بازارهای مالی ایجاد می‌کنند. در چنین شرایطی، پیش‌بینی روند بازار سهام یکی از چالش‌های اساسی تحلیل‌گران مالی و سرمایه‌گذاران به شمار می‌رود. داده‌های کلان اقتصادی که به‌طور سنتی برای تحلیل بازار سهام مورد استفاده قرار می‌گیرند، در زمان بحران‌های اقتصادی ممکن است کارایی کمتری داشته باشند، زیرا رفتار بازار تحت تأثیر عوامل روان‌شناسی و احساسات عمومی قرار می‌گیرد که در مدل‌های اقتصادی معمولی به‌راحتی قابل‌پیش‌بینی نیستند. در این مواقع، داده‌های شبکه‌های اجتماعی به‌عنوان یک منبع سریع و قابل‌اطمینان از احساسات جمعی و اخبار فوری می‌تواند نقش مهم‌تری در پیش‌بینی روندهای بازار ایفا کنند (بولن و همکاران، ۲۰۱۱). داده‌های شبکه‌های اجتماعی با ارائه اطلاعات در زمان واقعی می‌توانند داده‌های اقتصادی را تکمیل کنند. برای مثال، نظرات کاربران درباره قیمت کالاها و خدمات می‌تواند به پیش‌بینی تورم کمک نماید (بولن و همکاران، ۲۰۱۱).

۱-۴. چالش‌های ترکیب داده‌ها

داده‌های کلان اقتصادی و شبکه‌های اجتماعی به دلیل تفاوت در ماهیت و فرکانس زمانی، به ترکیب و تحلیل‌های پیچیده‌ای نیاز دارند. داده‌های کلان اقتصادی معمولاً با تأخیر منتشر می‌شوند، در حالی که داده‌های شبکه‌های اجتماعی به‌صورت لحظه‌ای تولید می‌شوند (ژنگ و همکاران، ۲۰۱۱). ترکیب داده‌های اقتصادی و شبکه‌های اجتماعی، علی‌رغم پتانسیل بالا برای بهبود تحلیل‌ها و پیش‌بینی‌های اقتصادی، با چالش‌های متعددی مواجه است. به عنوان نمونه، داده‌های شبکه‌های اجتماعی اغلب شامل اطلاعات نادرست یا گمراه‌کننده هستند که می‌تواند به تحلیل‌های نادرست منجر شوند. محتوای منتشرشده در این شبکه‌ها ممکن است بازتاب‌دهنده نظر اقلیت باشد و تصویری غیرواقعی از جامعه ارائه دهد (گایو^۱ و اولو^۲، ۲۰۱۲). داده‌های شبکه‌های اجتماعی از نظر فرمت، زبان و نوع محتوا بسیار متنوع هستند

1 Gayo

2 Avello

(متن، تصویر، ویدئو و غیره)، این تنوع فرآیند مسیر جمع‌آوری، پردازش و ارزیابی آنها را پیچیده‌تر می‌کند (چن و همکاران، ۲۰۱۲). کاربران شبکه‌های اجتماعی لزوماً نماینده جمعیت کل جامعه نیستند. برای مثال، برخی گروه‌های سنتی یا طبقات اقتصادی ممکن است کمتر در این شبکه‌ها حضور داشته باشند، که همین مولفه می‌تواند به سوگیری داده‌ها منجر شود (اولتنا^۱، ۲۰۱۹).

تحلیل احساسات داده‌های شبکه‌های اجتماعی با چالش‌هایی مانند چندمعنایی کلمات، کنایه‌ها و زبان غیررسمی مواجه است. این مسائل می‌تواند دقت مدل‌های تحلیل احساسات را کاهش دهد (پنگ^۲ و لی^۳، ۲۰۰۸). داده‌های اقتصادی معمولاً در بازه‌های زمانی مشخص (ماهانه یا سالانه) منتشر می‌شوند، در حالی که داده‌های شبکه‌های اجتماعی به‌صورت آنی تولید می‌شوند. این اختلاف زمانی می‌تواند ترکیب این داده‌ها را دشوار کند (لیو و همکاران^۴، ۲۰۲۰). اطلاعات نادرست یا شایعات منتشرشده در شبکه‌های اجتماعی می‌توانند تأثیر قابل‌توجهی بر رفتار اقتصادی و تحلیل داده‌ها داشته باشند. تشخیص و حذف این اطلاعات یکی دیگر از چالش‌های این مسیر است.

استفاده از داده‌های شبکه‌های اجتماعی برای تحلیل‌های اقتصادی می‌تواند مسائل مربوط به حریم خصوصی کاربران را ایجاد کند. این چالش به‌ویژه در کشورهایی با قوانین سخت‌گیرانه حریم خصوصی، بیشتر محسوس است. جمع‌آوری و استفاده از داده‌های شبکه‌های اجتماعی نیازمند رعایت حقوق مالکیت داده‌ها و قوانین مرتبط است. برخی پلتفرم‌ها محدودیت‌هایی برای دسترسی به داده‌های خود اعمال می‌کنند که می‌تواند مانع از انجام پژوهش شود (زیمِر^۵، ۲۰۱۰). داده‌های شبکه‌های اجتماعی معمولاً غیررسمی و غیرقابل‌اعتمادتر از داده‌های اقتصادی رسمی هستند. استفاده از این داده‌ها در سیاست‌گذاری‌های اقتصادی می‌تواند به تصمیمات اشتباه منجر شود (گنزالز^۶، ۲۰۱۳).

در راستای رسیدن به اهداف مطرح شده و با در نظر گرفتن ادبیات پژوهش، سوالات ذیل مطرح می‌شود.

۱. آیا مدل خطی لاسو قادر است متغیرهای کلان اقتصادی و شبکه‌های اجتماعی را به طور معناداری برای پیش‌بینی مبادلات سهام شناسایی کند؟
۲. آیا فرآیند غیرخطی گوسی دقت بالاتری در پیش‌بینی مبادلات سهام نسبت به مدل خطی لاسو ارائه می‌دهد؟
۳. آیا ترکیب متغیرهای کلان اقتصادی و شبکه‌های اجتماعی در مدل غیرخطی گوسی نتایج پیش‌بینی دقیق‌تری نسبت به استفاده صرف از داده‌های کلان اقتصادی در مدل خطی لاسو ارائه می‌دهد؟

1 Olteanu et al

2 Pang

3 Lee

4 Liu et al

5 Zimmer

6 Gonzalez

۲. روش‌شناسی پژوهش

این پژوهش، از نظر هدف، کاربردی و در دسته پژوهشهای کتابخانه‌ای طبقه‌بندی می‌شود. نوع این پژوهش توصیفی بوده و جز تحقیقات پسرودادی محسوب می‌شود، زیرا از داده‌های تاریخی برای بررسی و پیش‌بینی استفاده می‌نماید. این پژوهش از رویکرد کمی استفاده نموده و با ترکیب روشهای یادگیری ماشین برای تحلیل داده‌ها و پیش‌بینی دقیق‌تر رفتار بازار سهام به بررسی پرداخته است. در این پژوهش، داده‌های مرتبط به متغیرهای کلان اقتصادی شامل نرخ بهره، نرخ تورم، نرخ ارز، تولید ناخالص داخلی؛ از گزارشهای بانک مرکزی، داده‌های سامانه نیما و بازار آزاد و گزارش مرکز آمار ایران و سازمان برنامه و بودجه استخراج شده است. داده‌های اجتماعی از طریق پلتفرم سهامیاب و با کمک شاخص تعاملات^۱ و شاخص بیز^۲ اندازه‌گیری شد. در این مطالعه از روش خطی لاسو، الگوریتم ماشین بردار پشتیبانی و فرآیند گوسی برای پیش‌بینی استفاده شده است.

۲-۱. روش‌های اندازه‌گیری

۱-۲-۱. شاخص بیز

همانگونه که اشاره شد، دادهای شبکه‌های اجتماعی با استفاده از شاخص تعاملات و شاخص بیز اندازه‌گیری می‌شود. برای اندازه‌گیری این شاخص، تعداد پست‌ها و نظرات مرتبط با یک سهم که نشانگر سطح توجه عمومی به آن است، در نظر گرفته می‌شود. روش محاسبه چنین است: هر سهم در یک سال بررسی می‌شود و تعداد پستهای امروز و ۷ روز گذشته آن در نظر گرفته می‌شود با تقسیم این دو بر یکدیگر، شاخص حاصل میگردد.

$$Buzz Index = \frac{Number\ Of\ Post\ Today}{Number\ Of\ Post\ (Post\ n\ Day)}$$

آنچه در شاخص بیز اهمیت داد، تعداد کل پست‌ها یا نظرات مربوط به یک سهم خاص یا تعداد دفعات ذکر نام سهم در شبکه‌های اجتماعی است.

۲-۲-۱. شاخص تعاملات

برای سنجش تعاملات کاربران از تاثیرگذاری آنها استفاده می‌شود. به عنوان مثال تعاملات کاربران همچون لایک، کامنت و بانشر نشان دهنده میزان مشارکت و تعامل آنها با سهام مرتبط است. در این زمینه هر یک از فعالیتهای بالا چنین تعبیر می‌شود:

لایک: نشان دهنده تایید ساده و تاثیر کمی در پیش‌بینی رفتار بازار دارند.
کامنت: نشان دهنده تعامل عمیق‌تر هستند و معمولاً حاوی اطلاعات بیشتری هستند. مانند دلایل خرید و فروش سهام.

بازنشر: نشان دهنده تاثیر گسترده یک پست است و معمولاً بیشترین وزن را دارد. زیرا تاثیر زیادی بر دیدگاه عموم دارد.

بنابراین لایک امتیاز ۱، کامنت امتیاز ۲ و بازنشر امتیاز ۳ دریافت می‌کند.

1 Engagement Index

2 Buzz Index

$$Engagement Index = (like \times 1) + (Comments \times 2) + (Shares \times 3)$$

بعنوان نمونه اگر یک پست ۲۰۰ لایک، ۵۰ کامنت و ۳۰ بازنشر دارد. محاسبه شاخص تعامل چنین محاسبه می شود.

$$Engagement Index = (200 \times 1) + (50 \times 2) + (30 \times 3) = 390$$

۲-۲. روشهای ارزیابی

۲-۲-۱. مدل خطی لاسو^۱

این مدل در پاسخ به مشکلات رایج در رگرسیون خطی کلاسیک همانند همخطی و ابعاد بالای داده ها طراحی شده است. ایده اصلی مدل لاسو، افزودن یک جریمه L به تابع هزینه رگرسیون بود که علاوه بر کاهش پیچیدگی مدل به انتخاب ویژگیها نیز کمک مینمود.

$$\beta = \arg \arg \left(\sum_{i=1}^n (M_i - X_i \beta)^2 + \sum_{j=1}^p |\beta_j| \right)$$

در این رابطه:

M_i : مقدار پیش بینی شده مبادله سهام برای مشاهده i است.

X_i : بردار ویژگیهای کلان اقتصادی و شبکه های اجتماعی برای مشاهده i است.

β_j : ضرایب مدل است.

: پارامتر تنظیمی که پیچیدگی مدل را کنترل می کند.

P : تعداد ویژگیها است.

۲-۲-۲. الگوریتم ماشین بردار پشتیبانی^۲

ماشین بردار پشتیبانی که از آن برای طبقه بندی و رگرسیون استفاده می شود. مبنای کار این الگوریتم، دسته بندی خطی داده ها است و در تقسیم خطی داده ها سعی می کند، ابر صفحه های که حاشیه اطمینان بیشتری داشته باشد انتخاب نماید. هدف در این الگوریتم پیدا کردن خطی بهینه برای داده ها به وسیله روشهای برنامه سازی غیرخطی است که در حل مسائل محدودیت دار هستند، صورت می گیرد.

$$f(x) = b + (w, x)$$

در این رابطه:

$f(x)$: پیش بینی مبادلات سهام است.

x : بردار ویژگیها (مجموعه داده های کلان اقتصادی و شبکه های اجتماعی) است.

w : ضرایب مدل است.

این الگوریتم برای افزایش دقت پیش بینی استفاده می شود و قادر است روابط پیچیده تر میان ویژگیها و مبادلات سهام را شبیه سازی نماید.

1 Least Absolute Shrinkage and Selection Operator (LASSO)

2 Support Vector Machine (SVM)

۳-۲-۲. فرآیند گوسی^۱

این مدل به دلیل انعطاف‌پذیری بالا در شبیه‌سازی روابط پیچیده میان متغیرها و پیش‌بینی دقیق‌تر در شرایط داده‌های پراکنده و با نویز زیاد، مورد توجه قرار گرفته است. در پیش‌بینی‌های مالی، فرآیند گوسی می‌تواند روندهای غیرخطی در داده‌های بازار سهام را شبیه‌سازی کند.

$$f(X, y, *X) \sim N(\mu *, \sum *)$$

در این رابطه:

X : مجموعه داده‌های ورودی (ویژگیهای کلان اقتصادی و شبکه‌های اجتماعی) است.

y : مقادیر واقعی مبادلات سهام است.

X : داده‌های جدید که قرار است پیش‌بینی شوند.

μ : میانگین پیش‌بینی است.

Σ : ماتریس کوواریانس پیش‌بینی است.

در این مطالعه، مدل لاسو با استفاده از کتابخانه سایکیت لرن^۲ انجام شد و در اجرای فرآیند گوسی از کتابخانه جی پی^۳ استفاده گردید. در نهایت برای پردازش داده‌های متنی و تحلیل احساسات از ال تی کی^۴ و اسپیس۵ که ابزار پیشرفته پردازش زبان طبیعی برای درک متن است، بکار برده شد.

$$Final\ f = LASSO\alpha f + SAV\beta f + GPR\gamma f$$

در این رابطه:

$LASSO f$: پیش‌بینی حاصل از رگرسیون خطی لاسو است.

$SAV f$: پیش‌بینی حاصل از الگوریتم ماشین بردار پشتیبانی است.

$GPR f$: پیش‌بینی حاصل از فرآیند گوسی است.

α, β, γ : ضرایب ترکیب مدل است.

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

۳. یافته‌های پژوهش

۳-۱. آمار توصیفی

متغیرهای این پژوهش شامل ۱۰۵ سال-شرکت است ($7 \times 15 = 105$). همانطور که جدول شماره ۱ نشان می‌دهد، انحراف معیار زیادی در داده‌ها مشاهده می‌شود. به منظور اجرای صحیح روشهای ارزیابی، داده‌ها نیاز به استاندارد سازی است.

1 Gaussian Process Regression (GPR)

2 Scikit-learn

3 Gpy

4 NLTK

5 Spacy

جدول ۱: آمار توصیفی متغیرهای پژوهش

متغیر	میانگین	میان	حداکثر	حداقل	انحراف معیار
حجم مبادلات سهام	153552380	129000000	382000000	730000000	103537616
نرخ بهره	105	57/18	23	18	26/2
نرخ تورم	5/37	2/41	47.1	9.6	12.54
نرخ ارز	42/224571	250000	420000	42000	127119.17
تولید ناخالص داخلی	153552380	156500000	352000000	93000000	104238312
شاخص بیز	0.85/0	0.071	413/1	0.562-	0.157
شاخص تعاملات	066/0	068/0	625/0	285/1-	171/0

منبع: یافته های پژوهش

۳-۲. مدل های ارزیابی

داده ها در نرم افزار پایتون، استانداردسازی شد. سپس داده ها به دو گروه آموزشی و آزمایشی تقسیم بندی شد. گروه آموزشی سالهای ۱۳۹۶ الی ۱۴۰۰ بود و گروه آزمایشی سالهای ۱۴۰۱ الی ۱۴۰۲ را شامل میشد. پس از این مرحله، مدل لاسو اجرا شد و نتایج زیر بدست آمد.

جدول ۲: وزن ضرایب مولفه های اقتصاد کلان و شبکه های اجتماعی

نرخ بهره	نرخ تورم	نرخ ارز	تولید ناخالص داخلی	شاخص تعاملات	شاخص بیز
702/1-	519/4-	778/2	833/9	135/1-	765/8-
خطای مربعات میانگین			0047/0		
R^2			632/0		

منبع: یافته های پژوهش

مقدار میانگین مربعات خطا بسیار کوچک است لذا تفاوت پیش بینی مدل و مقادیر واقعی در حد کم است. همچنین مقدار R^2 نشان دهنده میزان تطابق مدل با داده ها است. از این رو، مدل میتواند ۶۳ درصد از تغییرات موجود در داده ها را توضیح دهد.

جدول ۳: نتایج حاصل از اجرای الگوریتم ماشین بردار پشتیبانی

621/0	SVM R^2
0054/0	خطای مربعات میانگین

منبع: یافته های پژوهش

جدول ۴: نتایج حاصل از اجرای فرآیند گوسی

641/0	SVM R^2
0033/0	خطای مربعات میانگین

منبع: یافته های پژوهش

وزن منفی متغیر نرخ بهره ($-1/702$) نشان می‌دهد، افزایش نرخ بهره با کاهش حجم مبادلات ارتباط مثبت دارد. در دنیای اقتصاد، نرخ بهره معمولاً بر بازار سرمایه تاثیر می‌گذارد و هنگامی که نرخ بهره افزایش می‌یابد، هزینه‌های استقرار بالاتر می‌رود و در نتیجه منجر به کاهش حجم مبادلات می‌شود.

نتایج بدست آمده نشان می‌دهد، نرخ تورم دارای وزن منفی ($-4/519$) است لذا افزایش نرخ تورم به طور منفی بر حجم مبادلات سهام تاثیر می‌گذارد. متغیر نرخ ارز نیز با وزن ($2/778$) گویای تاثیر مثبت آن بر مبادلات سهام است. بیشترین تاثیر گذاری در این مدل از سوی تولید ناخالص داخلی است، این متغیر با ضریب ($9/833$) نشان می‌دهد رشد تولید ناخالص داخلی با افزایش اعتماد سرمایه گذاران همراه بوده است زیرا منجر به افزایش حجم مبادلات سهام گردیده. همچنین وزن منفی شاخص تعاملات ($-1/135$) نشان می‌دهد، تعاملات بیشتر در شبکه‌های اجتماعی اثر منفی بر مبادلات سهام داشته است و منجر به ایجاد هیجانات کاذب یا رفتارهای ناپایدار در سهام شده است. سرانجام شاخص بیز نشان داد، سرمایه‌گذاران در شبکه‌های اجتماعی نسبت به سهام خود بی‌اعتماد شده‌اند و کمتر به مبادلات سهام پرداخته‌اند.

۴. بحث و نتیجه‌گیری

این پژوهش با هدف طراحی مدل پیش‌بینی مبادلات سهام با استفاده از متغیرهای کلان اقتصادی و داده‌های شبکه‌های اجتماعی به بررسی پرداخت. با توجه به پیچیدگی‌های موجود در پیش‌بینی مبادلات بازار سهام و تاثیرات مختلف مولفه‌های ذکر شده است. یافته‌های بدست آمده برای پاسخ به سوال اول نشان داد، متغیرهای کلان اقتصادی و شاخص‌های شبکه‌های اجتماعی به‌طور معناداری قادر است به پیش‌بینی مبادلات سهام بپردازد. در واقع مدل رگرسیونی لاسو از طریق اعمال جریمه برای ضرایب مدل، توانست تنها متغیرهای کلیدی را که به‌طور مستقیم با پیش‌بینی مبادلات بازار سهام ارتباط دارند شناسایی کند و از پیچیدگی‌های اضافی جلوگیری نماید. به طور اخص متغیرهایی کلان اقتصادی همچون نرخ تورم، نرخ ارز، تولید ناخالص داخلی و نرخ بهره که به‌طور سنتی به‌عنوان عوامل اصلی تاثیرگذار بر بازار سهام شناخته می‌شوند، در این مدل از اجزای بسیار اثرگذار شناخته شدند. افزون بر این، مولفه شبکه‌های اجتماعی شامل شاخص تعاملات و شاخص بیز به‌طور معناداری در پیش‌بینی مبادلات سهام تاثیر دارند. یافته‌های حاصل در این بخش با نتایج پژوهش ژنگ و همکاران (2018) و داموداران (2020) همراستا است.

از دیگر سوالات مطرحه در این پژوهش، مقایسه دقت پیش‌بینی مدل‌های مختلف بود. مدل‌های یادگیری ماشین، به‌ویژه فرآیند گوسی، توانست نسبت به مدل خطی لاسو، دقت بالاتری در پیش‌بینی مبادلات سهام ارائه نماید. این موضوع تایید می‌کند که روابط غیرخطی پیچیده موجود در داده‌های اقتصادی و شبکه‌های اجتماعی بهتر توسط مدل‌های پیشرفته‌تر شبیه‌سازی می‌شود. در حقیقت، قدرت مدل‌های غیرخطی مانند فرآیند گوسی، که قادر به شبیه‌سازی روابط پیچیده میان متغیرها هستند، در پیش‌بینی‌های مالی بسیار برجسته است. فرآیند گوسی، به‌ویژه در شرایطی که داده‌ها روابط پیچیده و غیرخطی دارند، توانست پیش‌بینی‌های به‌مراتب دقیق‌تری ارائه نماید. این فرآیند با استفاده از کرنل‌های مختلف و در نظر گرفتن وابستگی‌های پیچیده میان متغیرهای ورودی و خروجی، به‌خوبی تغییرات بازار سهام را شبیه‌سازی نمود و در عین حال، این فرآیند با قابلیت مدل‌سازی عدم قطعیت و در

نظر گرفتن توزیع‌های احتمال مختلف، توانست نتایج دقیقی حتی در مواجهه با داده‌های پراکنده و پیچیده ارائه دهد.

آخرین یافته‌های این پژوهش نشان داد، ترکیب متغیرهای کلان اقتصادی و شبکه‌های اجتماعی در مدل فرآیند گوسی، نتایج پیش‌بینی دقیق‌تری نسبت به استفاده صرف از داده‌های اقتصادی در مدل لاسو ارائه می‌دهد. به‌طور خاص، داده‌های شبکه‌های اجتماعی می‌تواند عواملی مانند تغییرات احساسات عمومی و روندهای اجتماعی را که تأثیرات مستقیمی بر بازار سهام دارند را مدل‌سازی کنند. لازم به ذکر است، داده‌های شبکه‌های اجتماعی به دلیل امکان دسترسی به احساسات و نظرات عمومی به صورت لحظه‌ای، می‌توانند به عنوان یک منبع پیش‌بینی قدرتمند در کنار داده‌های کلان اقتصادی عمل کنند، اما در کنار مزایای استفاده از داده‌های شبکه‌های اجتماعی، چالش‌هایی نیز وجود دارد. بزرگترین چالش در این زمینه صحت و دقت اطلاعات منتشر شده در شبکه‌های اجتماعی است که ممکن است گمراه کننده و غیرقابل اعتماد باشد و محتوای اشتباهی منتشر نمایند. نتایج این مطالعه به‌وضوح نشان داد، انتخاب مدل مناسب برای پیش‌بینی رفتار بازار سهام بسته به ماهیت داده‌ها و روابط میان متغیرها می‌تواند تأثیر قابل‌توجهی بر دقت پیش‌بینی داشته باشد. از یک سو، مدل رگرسیونی لاسو با استفاده از ویژگی‌های کلان اقتصادی و شبکه‌های اجتماعی به‌خوبی توانست روابط ساده‌تر و خطی میان این متغیرها را شبیه‌سازی کرده و پیش‌بینی‌هایی قابل‌اعتماد ارائه دهد. از سوی دیگر، فرآیند گوسی به دلیل توانایی‌های خود در مدل‌سازی روابط پیچیده و غیرخطی، توانست دقت بالاتری را در پیش‌بینی مبادلات سهام نشان دهد. این مدل به‌ویژه زمانی که ترکیب متغیرهای کلان اقتصادی و داده‌های شبکه‌های اجتماعی در نظر گرفته می‌شود، قادر به تولید پیش‌بینی‌های دقیق‌تر بود.

در نهایت، این پژوهش به این نتیجه رسید که ترکیب مدل‌های آماری سنتی با روش‌های یادگیری ماشین می‌تواند بهبود قابل‌توجهی در پیش‌بینی مبادلات سهام به‌ویژه در بازارهای مالی پیچیده و پویا ایجاد کند. به‌طور خاص، استفاده از مدل‌های غیرخطی مانند فرآیند گوسی در کنار داده‌های کلان اقتصادی و شبکه‌های اجتماعی می‌تواند یک ابزار قدرتمند برای تحلیل‌گران مالی و سرمایه‌گذاران باشد تا درک بهتری از تحولات آتی بازار داشته باشند. این پژوهش به محققان آتی پیشنهاد می‌کند در خصوص عوامل زمینه‌ای و فرهنگی تأثیرگذار بر پیش‌بینی مبادلات سهام بپردازند. زیرا تفاوت‌های فرهنگی و اجتماعی در جوامع مختلف اثرات متفاوتی بر مبادلات سهام می‌تواند داشته باشد. این بررسی نیز مانند هر پژوهش دیگری مستثنی از محدودیت نیست. از جمله محدودیتهای این پژوهش، اندازه کوچک نمونه و عدم بررسی اثر صنعت است.

منابع

1. Adler, M., Dumas, B. (1984). Exposure to currency risk: Definition and measurement. *Financial Management*, 13(2), 41-50.
2. Bernanke, B. S., Kuttner, K. N. (2005). What explains the stock market's reaction to Federal Reserve policy? *Journal of Finance*, 60(3), 1221-1257.
3. Bollen, J., Mao, H., Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8.
4. Chen, H., Chiang, R. H., Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165-1188.
5. Chen, H., De, P., Hu, Y., Hwang, B. H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies*, 27(5), 1367-1403.
6. Duan, J., He, Q., Zhang, Q. (2019). Financial forecasting with social media data. *Journal of Finance and Data Science*, 5(3), 189-202.
7. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417.
8. Gayo-Avello, D. (2012). No, you cannot predict elections with Twitter. *IEEE Internet Computing*, 16(6), 91-94.
9. Gonzalez, S. (2013). Social science in the era of big data. *Policy and Internet*, 5(2), 147-160.
10. Liu, Y., Li, Q., Xu, H. (2018). Predicting stock market movement using Twitter data. *Journal of Financial Markets*, 37, 118-137.
11. Lucas, R. E. (1977). Understanding business cycles. *Carnegie-Rochester Conference Series on Public Policy*, 5, 7-29.
12. Madura, J. (2021). *International Financial Management*. Cengage Learning.
13. Mankiw, N. G. (2020). *Principles of Economics*. Cengage Learning.
14. Mishkin, F. S. (2019). *The Economics of Money, Banking, and Financial Markets*. Pearson.
15. Olteanu, A., Castillo, C., Diaz, F., Kıcıman, E. (2019). Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in Big Data*, 2(13), 1-18.
16. Pang, B., Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
17. Romer, D. (2019). *Advanced Macroeconomics*. McGraw-Hill Education.
18. Shapiro, A. C. (2020). *Multinational Financial Management*. Wiley.
19. Shiller, R. J. (2000). *Irrational exuberance*. Princeton University Press.
20. Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139-1168.
21. Zhang, L., Yang, H. (2011). Social media and stock market prediction: A study on the role of sentiment analysis. *Journal of Financial Economics*, 102(3), 404-420.
22. Zhang, Y., Yang, L. (2018). A hybrid model for stock price prediction using macroeconomic variables and social media data. *Journal of Financial Markets*, 41, 152-166.
23. Zheludev, I., Smith, R., Aste, T. (2014). When can social media lead financial markets? *Scientific Reports*, 4, 4213.
24. Zimmer, M. (2010). "But the data is already public": On the Ethics of research in Facebook. *Ethics and Information Technology*, 12(4), 313-325.