فصلنامه مهندسی مالی و مدیریت اوراق بهادار
دوره پانزده، شماره شصت و یک، زمستان ۱۴۰۳
نوع مقاله: علمی پژوهشی
صفحات: ۱۶۷-۱۴۷

پیش‌بینی قیمت سهام توسط رویکرد رگرسیون لاسو در بورس اوراق بهادار تهران

امیر صادقی[۱]
امیر کمالی‌دولت‌آبادی[۲]

تاریخ دریافت مقاله: ۱۴۰۲/۰۸/۲۳ تاریخ پذیرش مقاله: ۱۴۰۳/۰۷/۰۲

چکیده

پیش‌بینی رفتار بازار سرمایه همواره یکی از چالش‌های سرمایه‌گذاران و تحلیل گران بازار است. در طول سالیان متمادی همواره روش‌های پیش‌بینی روندها، تکامل‌یافته و دانش پیش‌بینی رفتار و قیمت سهام همچنان درحال‌توسعه است. نکته حائز اهمیت تحلیل رفتار گذشته قیمت سهام بر اساس روش‌های تکنیکال می‌باشد که این سبک تحلیل غالباً بر تغییرات قیمت، میانگین‌های متحرک و حجم معاملات و غیره متمرکز می‌باشد. درروش‌های پیش‌بینی معمول تاکنون کمتر به نقش بازیگران حقیقی و حقوقی، روان‌شناسی بازار و تابلوخانی پرداخته‌شده است. در این مقاله سعی شده با بررسی نسبت‌های معاملاتی مبتنی بر تابلو خوانی و تحلیل رفتار گذشته و تحرک‌های معامله گران حقیقی در غالب رگرسیون خطی لاسو موردبررسی قرار گیرد. پیش‌بینی قیمت برای شش سهم منتخب از بورس اوراق بهادار تهران به روش مذکور انجام شد و نتایج بسیار دقیقی در مقایسه با رگرسیون خطی به دست آمد.

کلمات کلیدی

تابلو خوانی، درصد تغییر قیمت، پیش‌بینی قیمت سهام، رگرسیون لاسو

۱-استادیار، گروه ریاضی کاربردی، واحد پرند، دانشگاه آزاد اسلامی، تهران، ایران. (نویسنده مسئول) Drsadeghi.iau@gmail.com
۲-استادیار، گروه مهندسی صنایع، واحد پرند، دانشگاه آزاد اسلامی، تهران، ایران. amir.kamali2002@gmail.com

مقدمه

پیش‌بینی قیمت سهام با توجه به تأثیر آن در توسعه استراتژی‌های مؤثر در بازار سرمایه، به شکوفایی تجارت و صنعت کمک می‌کند و درنهایت در اقتصاد کشور مؤثر است. زمانی که شرکتی برای توسعه و سرمایه‌گذاری جدید نیاز به بودجه دارد، دو گزینه دارد. به‌منظور اخذ وام از مؤسسات مالی یا انتشار سهام از طریق بورس اوراق بهادار، یک شرکت می‌تواند سهام خود را که بخشی از مالکیت شرکت است، منتشر کند. برای انتشار سهام برای سرمایه‌گذاری، باید شرکت در بورس اوراق بهادار ثبت شود تا بتواند بودجه موردنیاز تجارت خود را جذب کند (سینر، ۲۰۱۹).

رفتار قیمت سهام، چه برای سرمایه‌گذاران حقیقی و حقوقی چه برای کارگزاران و مدیران بورس بسیار حائز اهمیت می‌باشد. در سالیان اخیر، یک طرز تفکر عمومی وجود دارد که بسته به اعتقاد مردم، به‌عنوان معاملات خریدوفروش سهام در نظر گرفته می‌شد. اکنون برخی از ابزارهای جدید با استفاده از روشی معروف به تحلیل تکنیکال برای پیش‌بینی قیمت‌های آتی از داده‌های قیمت تاریخی توسط سرمایه‌گذاران ابداع‌شده است. تجزیه‌وتحلیل تکنیکال اساساً مبتنی بر اندیکاتورهای ریاضی است. از چنین اندیکاتورهای تکنیکی برای کسب اطلاعاتی درباره این‌که آیا روند فعلی قیمت سهام ادامه خواهد یافت یا اینکه آیا سهام بیش‌ازحد فروخته‌شده است و می‌توان بیش‌ازحد خریداری شود، استفاده می‌شود (یانسن و همکاران، ۲۰۱۹). جدای از تحلیل تکنیکال، روشی معروف به تحلیل بنیادی وجود دارد که مربوط به شرکتی است که سهام را عرضه می‌کند. این روش، عملکرد گذشته یک شرکت و همچنین قابلیت اطمینان حساب‌های شرکت را ارزیابی می‌کند. با این روش، بسیاری از نسبت‌های عملکردی تولیدی و مالی به ارزیابی روند منطقی قیمت سهام کمک می‌کند (چن و همکاران، ۲۰۱۸).

قیمت سهام در آینده هم برای خریداران سهام و هم برای فروشندگان آن از اهمیت بالایی برخوردار است. ازآنجاکه پیش‌بینی قیمت سهام یک موضوع بسیار کلیدی برای سهامداران و سرمایه‌گذاران و کلیه فعالان بازار بورس می‌باشد، ارائه مدل‌های دقیق پیش‌بینی قیمت سهام از اهمیت بسیار بالایی برخوردار است. بر این اساس در تحقیق حاضر یک مدل جدید مبتنی بر رگرسیون لاسو برای پیش‌بینی قیمت سهام ارائه می‌شود. هدف نهایی تحقیق ارائه مدلی است که بتواند در قیاس با روش‌های موجود، قیمت سهام را با دقت بالاتری پیش‌بینی نماید. اگرچه تاکنون مطالعات زیادی در رابطه با پیش‌بینی قیمت سهام انجام‌شده است ولی این تحقیق نخستین مطالعه در داخل کشور است که بر یک روش جدید مبتنی بر رگرسیون خطی لاسو و بر اساس تحلیل رفتار معاملاتی سهامداران و بازیگران سهم ارائه می‌دهد و با تحلیل سناریوها از طرفی ارتباط میزان سرمایه‌گذاری، دفعات معامله و حجم

سرمایه‌گذاری با تغییرات قیمت را تحلیل کرده و از طرف دیگر با بهره‌گیری از مدل رگرسیونی لاسو با بهره‌گیری داده‌ها و اعمال جریمه برای مقادیر پرت و غیر مؤثر، مدلی بهینه برای پیش‌بینی قیمت سهام معرفی می‌کند.

روش‌های تحلیلی پیش‌بینی قیمت سهام عمدتاً مبتنی بر تحلیل رگرسیون خطی هستند. روش رگرسیون لجستیک، نوعی روش رگرسیونی است که هدف آن، به حداقل رساندن اختلاف بین مقادیر پیش‌بینی‌شده و مقادیر واقعی تابلو می‌باشد؛ به همین دلیل این روش از کاربرد زیادی برای پیش‌بینی قیمت سهام در بازارهای مالی برخوردار شده است. بااین‌وجود، فرضیات مورداستفاده در این روش بسیار ساده هستند که این امر ممکن است بر دقت پیش‌بینی آن تأثیرگذار باشد (فنگ ژو و همکاران، ۲۰۱۹). تصویر دیگری از روش‌های پیش‌بینی مبتنی بر رگرسیون، روش رگرسیونی لاسو است. مزیت روش لاسو در این است که راه‌حل‌هایی را پیشنهاد می‌کند که دارای پارمترهای کمتری بوده و بدین ترتیب تعداد متغیرهای موردنیاز برای پیش‌بینی قیمت سهام را نیز کاهش می‌دهد. همچنین بیان‌شده است که روش لاسو از روش‌های مبتنی بر رایانه همچون شبکه عصبی نیز عملکرد بهتری در پیش‌بینی قیمت سهام برخوردارند (روی و همکاران، ۲۰۱۵).

سؤالاتی که در این پژوهش مطرح می‌شود: مدل مبتنی بر تابلو خوانی با رویکرد روش رگرسیون خطی لاسو با چه دقتی می‌تواند تغییرات قیمت سهام را در یک بازه زمانی معین، پیش‌بینی کند؟ استفاده از ورودی‌های مختلف، چه تأثیری بر قدرت پیش‌بینی مدل پیشنهادی دارد؟ همان‌طور که می‌دانیم بسته به هدف، ماهیت موضوع و امکانات اجرایی پژوهش بایستی روش تحقیق مناسب انتخاب گردد. در تحقیق صورت گرفته نظر به ماهیت آن‌که به‌منظور پیش‌بینی قیمت سهام در بورس اوراق بهادار انجام‌گرفته است، هدف از نوع تحقیقات کاربردی و مدل‌سازی می‌باشد. با توجه به هدف پیش‌بینی قیمت سهام توسط رویکرد رگرسیون لاسو در بورس اوراق بهادار تهران می‌باشد.

در این مقاله به‌منظور پیش‌بینی قیمت سهام در بورس اوراق بهادار تهران از رگرسیون لاسو استفاده می‌شود. مزیت عمده روش رگرسیون لاسو این است که علاوه بر ارائه تابع خطی، اولویت‌بندی متغیرها را نیز دربر دارد. دیگر مزیت آن این است که تابع به‌دست‌آمده در برابر داده‌هایی که ایجاد آشفتگی می‌کنند، به نحو بهتری برازش می‌کند. لذا روش رگرسیون لاسو می‌تواند یک انتخاب بهینه برای حل مسأله موردنظر باشد. در بخش دوم مبانی نظری روش تحقیق مشتمل بر رگرسیون خطی، رگرسیون ستیغی (ریج) و رگرسیون لاسو و همچنین نحوه به دست آوردن رابطه محاسباتی روش موردنظر ارائه می‌شود. در این مقاله پیشینه پژوهش در بخش سوم ارائه‌شده است. در بخش چهارم تجربیات عددی با

ترسیم نمودارها و جداول و همچنین مقایسه‌ای از نتایج به‌دست‌آمده با روش‌های کلاسیک انجام می‌شود؛ و دربخش پنجم نتایج به‌دست‌آمده به همراه پیشنهادهای آتی برای پژوهش‌گران آورده شده است.

مبانی نظری

در این قسمت از مقاله به مبانی نظری به کار گرفته‌شده مشتمل بر مدل‌های رگرسیونی و روش‌شناسی تحقیق به‌طور مبسوط خواهیم پرداخت.

مدل‌های رگرسیونی

فرض کنید (x_i, y_i) مجموعه داده‌ها باشند، آنگاه برای $i = 1,\ldots,N$ **مدل رگرسیون خطی ساده** به‌صورت زیر داده می‌شود:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1,\ldots,N \quad (۱)$$

که در آن β_0 عرض از مبدأ و β_1 شیب نامیده می‌شوند و مقدار ε_i همان مقدار خطاست. فرم برداری رابطه (۱) به‌صورت زیر داده می‌شود:

$$\mathbf{y} = \beta_0 + \beta_1 \mathbf{x} + \boldsymbol{\varepsilon}, \quad (۲)$$

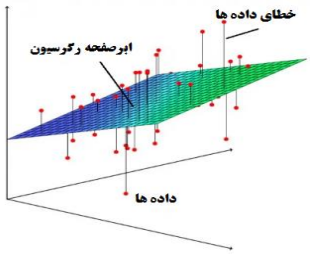
به‌منظور برآورد پارامترهای رگرسیون خطی ساده، کافی است تابع مجموع مربعات خطای یعنی مقدار $\mathrm{SSE} = \sum_{i=1}^{N} e_i^2 = \sum_{i=1}^{N} (y_i - \hat{y}_i)_i^2$ را کمینه کرد؛ بنابراین با مساوی صفر قرار دادن مشتق عبارت فوق نسبت به β_0, β_1 و حل آن پارامترها به فرم زیر حاصل می‌شوند:

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x}, \qquad \hat{\beta}_1 = \frac{\sum_{i=1}^{N} \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sum_{i=1}^{N} \left(x_i - \overline{x}\right)} \quad (۳)$$

که در آن $\overline{x}, \overline{y}$ میانگین x, y هستند. در **رگرسیون خطی چندگانه**، پارامترهای یک مدل خطی توسط یک تابع هدف و مقدارهای متغیرها، برآورد می‌شوند درحالی‌که در مدل در نظر گرفته‌شده، یک رابطه خطی برحسب پارامترهای مدل محاسبه می‌شوند. به‌این‌ترتیب اگر n مشاهده از متغیر مستقل p بعدی x موجود باشد و هدف برقراری یک رابطه خطی با متغیر پاسخ y باشد، می‌توان از مدل رگرسیون خطی زیر استفاده کرد:

$$y_i = \beta_0 .1 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} + \varepsilon_i, \quad i = 1,\ldots,N \quad (۴)$$

ازآنجایی‌که متغیر مستقل x دارای p بعد است، مقدار آن را در هر بعد با یک متغیر مستقل یک‌بعدی جایگزین می‌شود. واضح است که اندیس i شماره مشاهده و ε_i جمله خطای مدل رگرسیونی محسوب می‌شود. توجه شود که رابطه بین متغیر مستقل و وابسته در رگرسیون ساده به‌صورت معادله یک خط بیان می‌شود درحالی‌که در رگرسیون چندگانه، اگر دو متغیر مستقل با یک متغیر وابسته در رابطه خطی باشند، شکل این رابطه به‌صورت یک صفحه بیان خواهد شد. درصورتی‌که بیش از دو متغیر مستقل در مدل رگرسیون خطی به کار رود، مدل به شکل یک ابر صفحه همانند شکل ۱ ظاهر می‌شود.



شکل ۱: ابر صفحه رگرسیون تشکیل‌شده برای چند متغیر مستقل

روش‌شناسی تحقیق

رگرسیون ستیغی (ریج) یکی از نسخه‌های رگرسیون خطی است که در مسائلی با متغیرهای مستقل که دارای همبستگی بالایی می‌باشند به‌منظور تقریب ضرایب استفاده می‌شود و مشکلات هم‌خطی چندگانه را کاهش می‌دهد. در رگرسیون ستیغی، ضریبی از نرم L_2 تعریف‌شده توسط $\|x\|_2 = \left(\sum_{i=1}^{N} x_i^2\right)^{1/2}$ به تابع هزینه اضافه می‌شود و سپس مقدار کمینه آن به‌صورت زیر تعیین می‌شود:

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^n} \left\{ \|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_2^2 \right\} \quad (۵)$$

توجه شود پارامتر λ میزان جریمه روی نرم L_2 را مشخص می‌کند. اضافه کردن ضریبی از نرم L_2 به تابع هزینه معادل ایجاد محدودیتی بر نرم L_2 به‌صورت زیر است (روی و همکاران، ۲۰۱۵):

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^n} \left\{ \|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_2^2 \right\} \quad s.t \quad \|\boldsymbol{\beta}\|_2^2 \le t^2 \quad (۶)$$

که قید داده‌شده در رابطه (۶) درواقع یک n بعدی به فرم $\beta_1^2 + \beta_2^2 + \ldots + \beta_n^2 \le t^2$ می

باشد. با توجه به اینکه $\|\mathbf{x}\|_2^2 = \mathbf{x}^T\mathbf{x}$، فرم بسته ماتریسی تابع هزینه را به‌صورت زیرمی‌توان نوشت:

$$\varphi(\mathbf{X}) = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) + \lambda\boldsymbol{\beta}^T\boldsymbol{\beta} \quad (۷)$$

با به دست آوردن $\frac{\partial\varphi}{\partial\boldsymbol{\beta}} = 0$ خواهیم داشت: $\mathbf{X}^T\mathbf{y} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})\boldsymbol{\beta}$

درنتیجه پارامتر بهینه رگرسیون ستیغی‌خطی به‌صورت زیربدست‌می‌آید (روی و همکاران، ۲۰۱۵):

$$\hat{\boldsymbol{\beta}}_{Ridge} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y} \quad (۸)$$

در **رگرسیون لاسو** ضریبی از نرم L_1 تعریف‌شده به‌صورت $\|x\|_1 = \sum_{i=1}^{N}|x_i|$ به تابع هزینه اضافه می‌شود و سپس مقدار کمینه آن به دست می‌آید یعنی:

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^n}\left\{\frac{1}{N}\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1\right\} \quad (۹)$$

پارامتر λ نیز میزان جریمه روی نرم L_1 را مشخص می‌کند. اضافه کردن ضریبی از نرم L_1 به تابع هزینه معادل ایجاد محدودیتی بر نرم L_1 به‌صورت زیر است (روی و همکاران، ۲۰۱۵):

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^n}\left\{\frac{1}{N}\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1\right\} \quad s.t \quad \|\boldsymbol{\beta}\|_1 \le t \quad (۱۰)$$

که قید داده‌شده در رابطه (۱۰) درواقع یک n بعدی به فرم $|\beta_1| + |\beta_2| + \ldots + |\beta_n| \le t$ می‌باشد. برای به دست آوردن فرم بسته جواب به فرم ماتریسی، تابع هزینه به‌صورت ذیل داده می‌شود:

$$\psi(\mathbf{X}) = \frac{1}{N}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) + \lambda\boldsymbol{\beta}^T\boldsymbol{\beta} \quad (۱۱)$$

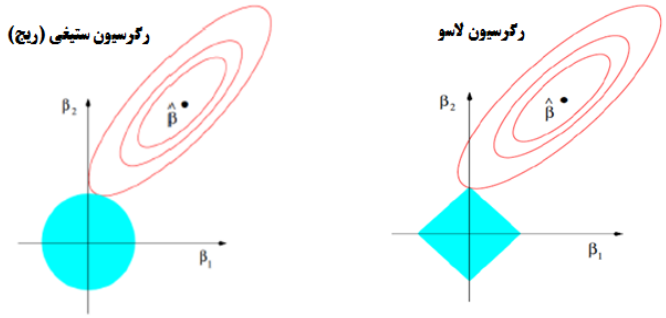
با به دست آوردن $\frac{\partial\psi}{\partial\boldsymbol{\beta}} = 0$ خواهیم داشت: $\frac{2}{N}\mathbf{X}^T(\mathbf{X}\boldsymbol{\beta}-\mathbf{y}) = 0 \Rightarrow \mathbf{X}^T\mathbf{X}\boldsymbol{\beta} - \mathbf{X}^T\mathbf{y} = 0$

درنتیجه داریم (روی و همکاران، ۲۰۱۵):

$$\hat{\boldsymbol{\beta}}_{Lasso} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \quad (۱۲)$$

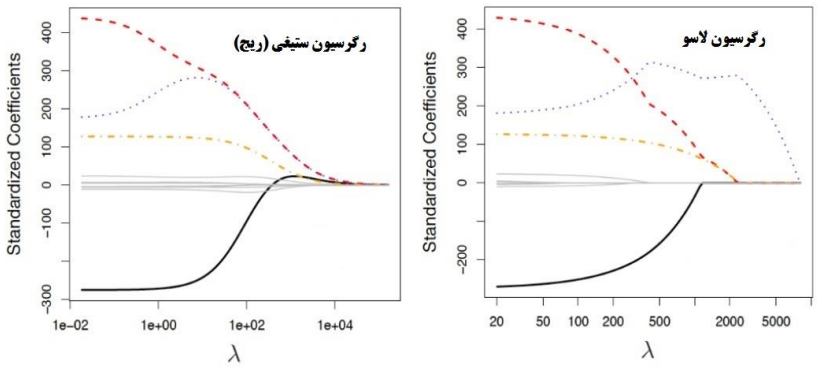
محدودیت‌های اعمال‌شده در دو مدل رگرسیون ستیغی خطی و لاسو با پارامترهای β_1 و β_2 در شکل ۲ نشان داده‌شده است. توجه شود که جمع خطای مربعات، به شکل خطوط تراز بیضوی نمایش داده می‌شود. به‌علاوه در مسئلهٔ رگرسیون ستیغی، ناحیه‌ای که ضرایب مدل را مشخص می‌کند از

رابطه $\beta_1^2 + \beta_2^2 \le t^2$ به دست می‌آید که درواقع دایره‌ای به شعاع t را نمایش می‌دهد؛ اما در رگرسیون لاسو، محدودیت اعمال‌شده روی ضرایب به‌صورت $|\beta_1| + |\beta_2| \le t$ است که در مختصات دوبعدی ناحیه‌ای به شکل لوزی را تشکیل می‌دهد. در هر دو روش اولین محل برخورد این نواحی با خطوط تراز بیضوی به‌عنوان جواب مسأله در نظر گرفته می‌شود.



شکل ۲: برخورد خطوط تراز با محدوده ضرایب رگرسیون ستیغی (ریج) و رگرسیون لاسو

برای یافتن مقدار بهینه برای λ نیز می‌توان از روش‌های مختلف مانند اعتبار سنجی متقابل استفاده کرد. هرچه مقدار λ افزایش پیدا کند، تعداد پارامترها کاهش پیدا خواهد کرد. با کاهش این مقدار نیز برآوردگرها به مقدار مشابه درروش OLS نزدیک می‌شوند. به‌این‌ترتیب بزرگی یا کوچکی مقدار λ می‌تواند روی تعداد پارامترها و مقدار آن‌ها تأثیرگذار باشد. همان‌طور که ملاحظه شد، تعیین مقدار مناسب برای پارامتر λ به‌راحتی میسر نیست و باید با سعی و خطا مقدار آن را جستجو کرد. در اینجا به کمک ترسیم نمودار MSE برحسب مقدار λ به جواب مناسب برای این مقدار رسیده شد که این امر در شکل ۳ نشان داده‌شده است.



شکل ۳: نمودار ضرایب مدل رگرسیون لاسو

همان‌طور که می‌دانیم قدرت پیش‌بینی توانایی یک مدل یا روش در برآورد دقیق روند آتی تغییرات قیمت سهام را نشان می‌دهد (لاهمیری، ۲۰۱۸). در این مقاله دقت پیش‌بینی بر اساس دو شاخص خطای ریشه میانگین مربعات و میانگین درصد خطای مطلق موردسنجش قرار می‌گیرد. برای سنجش شاخص جذر میانگین مربعات خطا از رابطه زیر استفاده می‌شود:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - p_i)^2}{n}} \quad (۱۳)$$

هرچه شاخص جذر میانگین مربعات خطا RMSE کمتر باشد پیش‌بینی ما بهتر خواهد بود. شاخص میانگین درصد خطای مطلق MAPE با استفاده از رابطه زیر اندازه‌گیری می‌شود:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{N}\frac{|y_i - p_i|}{y_i} \times 100\% \quad (۱۴)$$

همچنین، در این تحقیق قدرت پیش‌بینی قیمت سهام در سه سناریوی مختلف بررسی می‌شود:

- سناریوی اول، نسبت قدرت خریدار و نسبت قدرت فروشنده به‌عنوان ورودی‌های مدل در نظر گرفت.
- سناریوی دوم، حجم فروش و حجم خرید هر کد حقیقی به‌عنوان ورودی‌های مدل در نظر گرفت.
- سناریوی سوم، مقدار فروش و نیز خرید هر کد حقیقی به‌عنوان ورودی‌های مدل در نظر گرفت.

پیشینه تحقیق

در پژوهش‌های داخلی: غلامیان و همکاران (۱۳۹۷) در پژوهش خود با استفاده از رویکرد الگوریتم جنگل تصادفی که در زمره روش‌های طبقه‌بندی هوش مصنوعی می‌باشد، به پیش‌بینی روند قیمت در بازار سهام پرداختند. نتیجه پژوهش بر روی داده‌های روزانه شاخص کل بین سال‌های ۱۳۹۳ تا ۱۳۹۵ نشان داد که دقت الگوریتم پیشنهادی در برآورد روند بازار ۶۴ درصد و نسبت به دو روش مقایسه شده رگرسیون لجستیک و روش کاملاً تصادفی از دقت بالاتری برخوردار می‌باشد. موسوی و همکاران (۱۳۹۷) مطالعه‌ای توسط استفاده از الگوریتم ترکیبی عصبی کرم شب‌تاب و روش رگولاسیون بیزین جهت پیش‌بینی قیمت سهام سه شرکت ایران‌خودرو، پتروشیمی شیراز و ذوب‌آهن اصفهان انجام دادند. نتایج حاصل از آنالیز خطای شبکه‌های عصبی تعمیم داده‌شده نشان داد که مدل‌های مذکور با دقت بسیار مناسبی قادر به پیش‌بینی قیمت سهام در روز آینده برای شرکت‌های ذکرشده می‌باشند. مشاری و همکاران (۱۳۹۸) در مقاله‌ای توسط طراحی مدل هوشمند ترکیبی حاوی مدل‌های مبتنی بر

داده‌کاوی و الگوریتم ژنتیک بهینه‌سازی شده، پیش‌بینی نقاط طلایی قیمت سهام صنعت خودرو و ساخت قطعات بین سال‌های ۱۳۸۰ تا ۱۳۹۵ انجام دادند. نتایج حاصل از این پژوهش نشان داد که نقاط طلایی با دقت قابل‌قبول پیش‌بینی پذیرند و بهینه‌سازی در همه مدل‌ها باعث افزایش دقت پیش‌بینی نمی‌گردد اما خطای فاحش را به میزان قابل‌توجهی کاهش می‌دهد. سیف زاده و همکاران (۱۳۹۹)، با ارائه یک سیستم خبره پیش‌بینی قیمت سهام در بورس مبتنی بر شبکه عصبی فازی را بررسی کرده‌اند. نتایج پیش‌بینی نشان داد که تحت تعمیم فازی، قیمت سهام در بورس اوراق بهادار قابل پیش‌بینی بوده و کاملاً با نتایج تحلیل تکنیکال و فاندامنتال همخوانی دارد. فتح علیان و همکاران (۱۳۹۹)، در پژوهشی توسط الگوی بهینه ارزیابی و قیمت‌گذاری عرضه اولیه عمومی سهام با استفاده از تکنیک‌های تصمیم‌گیری چند معیاره فازی برای ۴۲۱ شرکت طی سال‌های ۱۳۸۵ تا ۱۳۹۷ را بررسی کرده‌اند. نتایج تحقیق آن‌ها نشان داد الگوی الگوریتم ژنتیک مدل بهینه‌تری برای پیش‌بینی قیمت‌گذاری سهام عرضه اولیه می‌باشد.

در پژوهش‌های بین‌المللی: مطالعه‌ای توسط داش (۲۰۱۴) توسط یک شبکه ترکیبی تکاملی دینامیک برای تحلیل و پیش‌بینی بازار سهام با استفاده از فیلتر کالمن انجام گردید. در این پژوهش به مقایسه چهار رویکرد یادگیری مختلف از جمله فیلتر کامن و الگوریتم تکامل تفاضلی پرداخته شد و نشان داده شد رویکرد ترکیبی پیشنهادی نسبت به سایر روش‌ها به‌طور معناداری دارای قدرت پیش‌بینی بالاتری می‌باشد. میتنگ و همکاران (۲۰۱۵)، در مطالعه‌ای توسط ترکیب تحلیل مؤلفه غیرخطی و رگرسیون بردار پشتیبان برای پیش‌بینی قیمت سهام به معرفی روشی نوین اقدام کردند. نتایج مقایسه روش پیشنهادی این پژوهش نشان داد که این رویکرد بر کلیه روش‌های ترکیبی رگرسیون بردار پشتیبان برتری داشته است. لاهمیری و همکاران (۲۰۱۸) از روش‌های فرا ابتکاری برای پیش‌بینی قیمت سهام استفاده کردند. این تحقیق مدلی را برای پیش‌بینی قیمت سهام روزانه ارائه می‌دهد که از آنالیز طیف منفرد و رگرسیون بردار پشتیبان همراه با بهینه‌سازی ذرات استفاده می‌کند. عملکرد مدل پیشنهادی با عملکرد چهار مدل که به‌طور گسترده در پیش‌بینی مالی مورداستفاده قرار می‌گیرند من‌جمله تبدیل موجک، رگرسیون و غیره مقایسه شده است. بیاز و همکاران (۲۰۱۸) در پژوهشی به مقایسه مدل‌های مختلف پیش‌بینی قیمت سهام پرداختند. آزمون‌های انجام‌شده بر روی ۱۴۰ شرکت در کشور هند نشان می‌دهد که مدل‌هایی که از شاخص‌های مبتنی بر آنالیز بنیادی استفاده می‌کنند، از آنالیزهای تکنیکال با استفاده از شاخص‌های تحلیل تکنیکی با سطح بازده متفاوت در صنایع متفاوت هستند. علاوه بر این، در بیش از ۹۵ درصد موارد، استفاده از

شاخص‌های ترکیبی منجر به کاهش میانگین ریشه مربعات خطا در مقایسه با استفاده از شاخص‌های اساسی یا فنی به‌تنهایی می‌شود. اسدی (۲۰۱۸) به بررسی استفاده از پارامترهای رگرسیون فازی برای بهینه‌سازی پیش‌بینی پرداخت. این محقق از گسسته سازی اطلاعات ورودی مبتنی بر قواعد فازی ممدانی استفاده کرد. هدف این بود که بر اساس مدل پیشنهادی الگوریتم ریپر با استفاده از هرس افزایشی مکرر برای کاهش خطای پیش‌بینی ارائه دهد. لیو و همکاران (۲۰۱۸)، در مقاله‌ای تحت عنوان پیش‌بینی نوسان قیمت بازار سهام توسط سیستم شبکه عصبی با تابع زمان تصادفی، نوسانات قیمت توسط یک شبکه عصبی را پیش‌بینی کرده‌اند. در مدل پیشنهادی فرض بر این است که سرمایه‌گذاران تصمیم می‌گیرند موقعیت‌های سرمایه‌گذاری خود را با تجزیه‌وتحلیل داده‌های تاریخی بر بازار سهام بسنجند به‌طوری‌که داده‌های تاریخی می‌تواند بر نوسانات بازار سهام تأثیر بگذارد و تابع تصادفی زمان به‌عنوان مدل پیش‌بینی برای دادن وزن به هر داده تاریخی معرفی شود. نتایج حاکی از کارایی بالای شبکه عصبی پیشنهادی می‌باشد. ژانگ و همکاران (۲۰۱۹) در تحقیقی مدلی مبتنی بر تلفیق رگرسیون بردار پشتیبانی و الگوریتم کرم شب‌تاب برای پیش‌بینی قیمت سهام ارائه کردند. در این مقاله یک روش جدید برای پیش‌بینی قیمت سهام با ترکیب با الگوریتم کرم شب‌تاب ارائه‌شده است. نتایج تجربی موارد زیر را نشان می‌دهد: اولاً در مقایسه با الگوریتم‌های دیگر، الگوریتم پیشنهادی دارای عملکرد بهتری است و ثانیاً روش پیش‌بینی پیشنهادی می‌تواند به‌عنوان ابزاری عملی و مؤثر برای پیش‌بینی قیمت سهام در نظر گرفته شود. پدهی و همکاران (۲۰۲۰)، به پیش‌بینی قیمت سهام در روز با استفاده از یک مدل رگرسیون لاسو با هدف ایجاد یک مدل برای پیش‌بینی قیمت سهام روز مانند قیمت پایین، قیمت نزدیک و قیمت بالا را بررسی کرده‌اند. با استفاده از معیارهای عملکرد یعنی خطای کمترین مربعات بهترین الگوریتم مناسب مدل پیداشده است درنتیجه رگرسیون لاسو بهترین عملکرد را در بین همه رگرسیون‌ها دارد.

تجربیات عددی و نتایج

در این فصل پیاده‌سازی به همراه نتایج پیش‌بینی قیمت سهام توسط رویکرد رگرسیون لاسو در بورس اوراق بهادار تهران با اشکال و جدول بحث و تفسیر شده است. در ابتدا همبستگی پیرسون و اسپیرمن موردبررسی قرارگرفته و نمادهایی با بیشترین همبستگی جهت پیش‌بینی انتخاب‌شده است؛ بنابراین، داده‌های شش نماد و پارس، فملی، وتوصا، خبهمن، شبندر و سشرق بررسی‌شده است.

آنچه در این تحقیق موردبررسی قرار می‌گیرد مبتنی بر تحرکات حجمی، نسبت‌های خریدوفروش و قدرت پول می‌باشد که در بورس اوراق بهادار تهران اصطلاحاً به آن تابلو خوانی گفته می‌شود. تغییرات

حجم معاملات معمولاً هم‌زمان، پیش و بعد از تغییرات قیمتی تغییرات معناداری می‌کند که بسته به‌اندازه شرکت‌ها، میزان شناوری و رفتار معامله گران در نمادهای مختلف متفاوت است که با بررسی رفتار این متغیر درگذشته و بررسی میزان همبستگی این متغیر با تغییرات قیمت سهم می‌توان رفتار سهم در آینده را پیش‌بینی کرد.

سرانه تعداد سهم خریدوفروش هر کد معاملاتی در یک جلسه با تغییرات قیمتی سهم ارتباط دارد و میزان این ارتباط بسته به‌اندازه شرکت‌ها، میزان شناوری و رفتار معامله گران در نمادهای مختلف متفاوت است. سرانه مقدار ریالی که هر کد معاملاتی جهت خریدوفروش نیز در یک جلسه با تغییرات قیمتی سهم ارتباط دارد و میزان این ارتباط بسته به‌اندازه شرکت‌ها، میزان شناوری و رفتار معامله گران در نمادهای مختلف متفاوت است که با بررسی رفتار این متغیر درگذشته و بررسی میزان همبستگی این متغیر با تغییرات قیمت سهم می‌توان رفتار سهم در آینده را پیش‌بینی کرد.

میزان همبستگی متغیرها تک‌تک در هر نماد با نرم‌افزار SPSSمحاسبه و نتایج در جدول ۱ تجمیع گردیده است. با توجه به اینکه مقدار همبستگی موردنظر است نه مستقیم یا غیرمستقیم بودن، قدر مطلق همبستگی در جدول قرارگرفته است. در جدول ۲ میزان همبستگی پیرسون و اسپیرمن در ۷ متغیر مستقل با متغیر درصد تغییرات قیمت پایانی ۲ روزه هر نماد موردبررسی قرارگرفته است که همبستگی سه نماد وتوصا، و پارس و شبندر معنادارتر می‌باشند.

جدول ۱: همبستگی پیرسون و اسپیرمن متغیرهای مستقل با درصد تغییرات ۲ روزه قیمت پایانی

متغیر مستقل	همبستگی پیرسون (درصد تغییرات ۲ روزه قیمت)					
	فملی	خبهمن	سشرق	و پارس	شبندر	وتوصا
حجم کل معاملات	۰,۱	۰,۰	۰,۱	۰,۱	۰,۱	۰,۱
سرانه حجم معاملات هر کد خریدار حقیقی	۰,۳	۰,۵	۰,۴	۰,۳	۰,۳	۰,۳
سرانه مبلغ معاملات هر کد خریدار حقیقی	۰,۳	۰,۲	۰,۲	۰,۲	۰,۳	۰,۴
نسبت قدرت خریدار حقیقی	۰,۳	۰,۳	۰,۲	۰,۳	۰,۳	۰,۳
سرانه حجم معاملات هر کد فروشنده حقیقی	۰,۱	۰,۰	۰,۰	۰,۰	۰,۲	۰,۲
سرانه مبلغ معاملات هر کد فروشنده حقیقی	۰,۱	۰,۲	۰,۱	۰,۳	۰,۱	۰,۳
نسبت قدرت فروشنده حقیقی	۰,۱	۰,۳	۰,۱	۰,۳	۰,۲	۰,۱

منبع: نتایج تحقیق

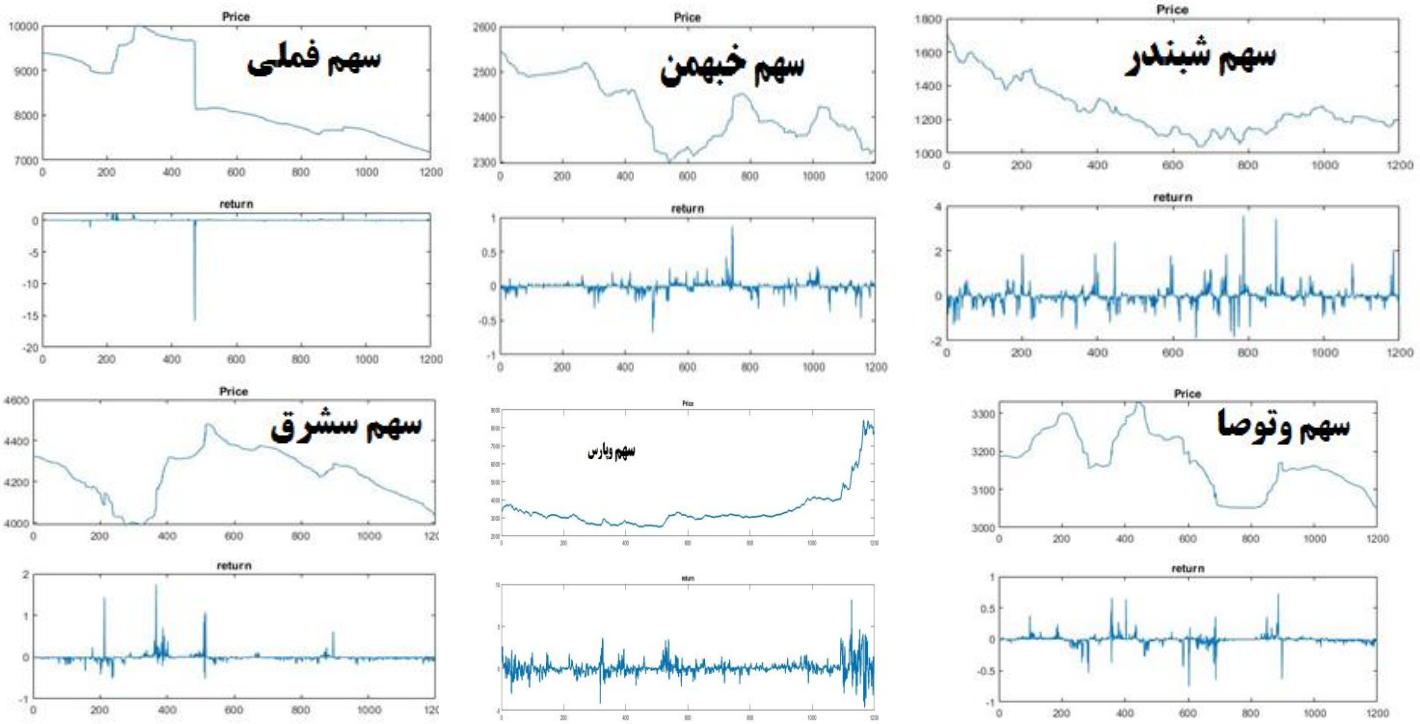
برای ملاحظه روند قیمتی و بازدهی سهام مذکور توسط نرم‌افزار متلب، نمودارهای آن‌ها ترسیم‌شده و نتایج آن در شکل ۴ آورده شده است. همچنین، برای بررسی نرمال بودن داده‌های شش نماد

موردبررسی، دو متغیر قیمت و درصد تغییرات دوروزه قیمت را در محیط spss موردبررسی قرار می‌دهیم. مطابق نتایج شکل ۵ مشاهده می‌شود که داده‌ها نرمال نمی‌باشند، لذا برای بررسی بیشتر جهت اطمینان از مناسب بودن مدل رگرسیونی می‌بایست آزمون‌های دیگری نظیر آزمون ایستایی و دوربین واتسون انجام شود. آزمون دوربین واتسون روشی برای تحلیل و تشخیص همبستگی در باقی‌مانده‌های مدل با یک واحد تأخیر می‌باشد. میزان همبستگی در این آزمون از ۰ تا ۴ می‌باشد و عدد ۲ نشانگر بیشترین همبستگی در باقی‌مانده‌ها با ۱ واحد تأخیر می‌باشد به‌طوری‌که بیشترین همبستگی عدد ۲ و این عدد هرچه از ۲ به سمت صفر میل کند میزان همبستگی مثبت کمتر شده و هرچه از ۲ به سمت ۴ میل کند میزان همبستگی منفی کمتر می‌شود. برای تأیید این استقلال می‌بایست در همبستگی مستقیم این مقدار کمتر از ۱ و در همبستگی معکوس بیشتر از ۳ باشد. یکی از فرض‌های اساسی مناسب بودن مدل رگرسیونی مستقل بودن داده‌های خطا می‌باشد. نتایج محاسبات در جدول شماره ۲ ملاحظه می‌شود.

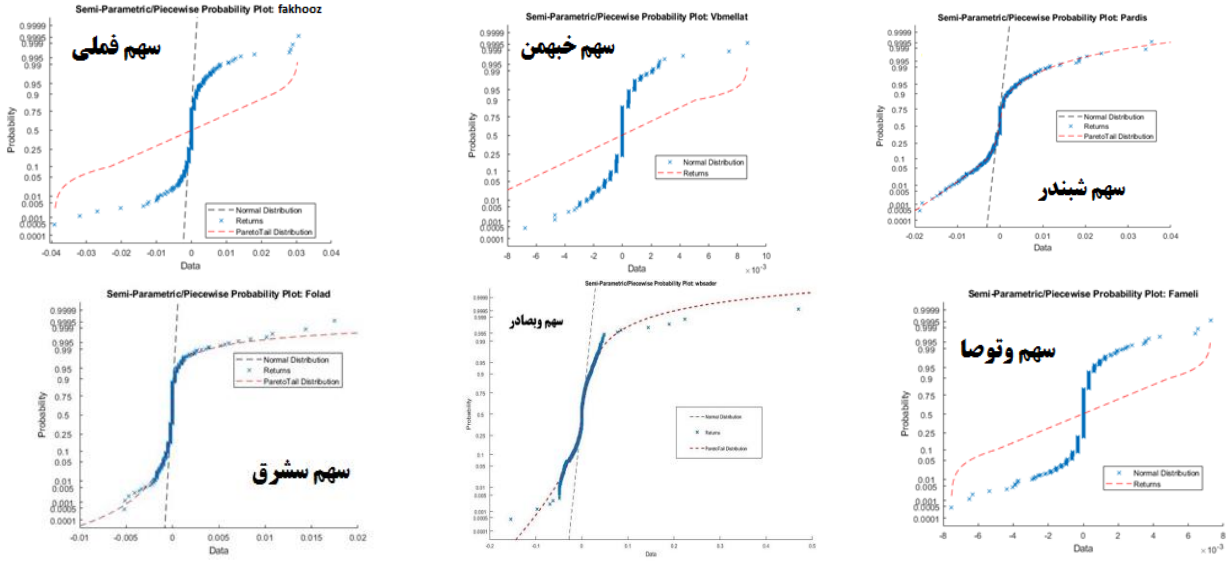
جدول ۲: نتایج آزمون دوربین-واتسون در نمادهای منتخب

Stock	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
فملی	۰,۳۵۲	۰,۱۲۴	۰,۰۵۷	۶,۲۱۴۳۶	۰,۱۲۴	۱,۸۵۶	۷	۹۲	۰,۰۸۶	۰,۸۱۷
خبهمن	۰,۴۳۱	۰,۱۸۶	۰,۱۴۵	۶,۲۱۸۵۳	۰,۱۸۶	۴,۵۶۲	۷	۱۴۰	۰,۰۰۰	۰,۸۷۰
شبندر	۰,۳۷۴	۰,۱۴۰	۰,۰۹۷	۵,۸۸۷۹۲	۰,۱۴۰	۳,۲۳۳	۷	۱۳۹	۰,۰۰۳	۰,۷۷۱
سشرق	۰,۴۵۵	۰,۲۰۷	۰,۱۶۸	۶,۰۴۵۹۸	۰,۲۰۷	۵,۲۹۰	۷	۱۴۲	۰,۰۰۰	۰,۹۲۱
و پارس	۰,۴۹۵	۰,۲۴۵	۰,۱۸۴	۶,۱۶۴۸۹	۰,۲۴۵	۴,۰۳۱	۷	۸۷	۰,۰۰۱	۰,۸۲۱
وتوصا	۰,۴۱۰	۰,۱۶۵	۰,۱۲۷	۵,۳۴۴۳۲	۰,۱۶۸	۴,۰۲۱	۷	۱۳۹	۰,۰۰۰	۱,۰۳۹

منبع: نتایج تحقیق



شکل ۴: نمودار قیمت و بازدهی سهام مختلف



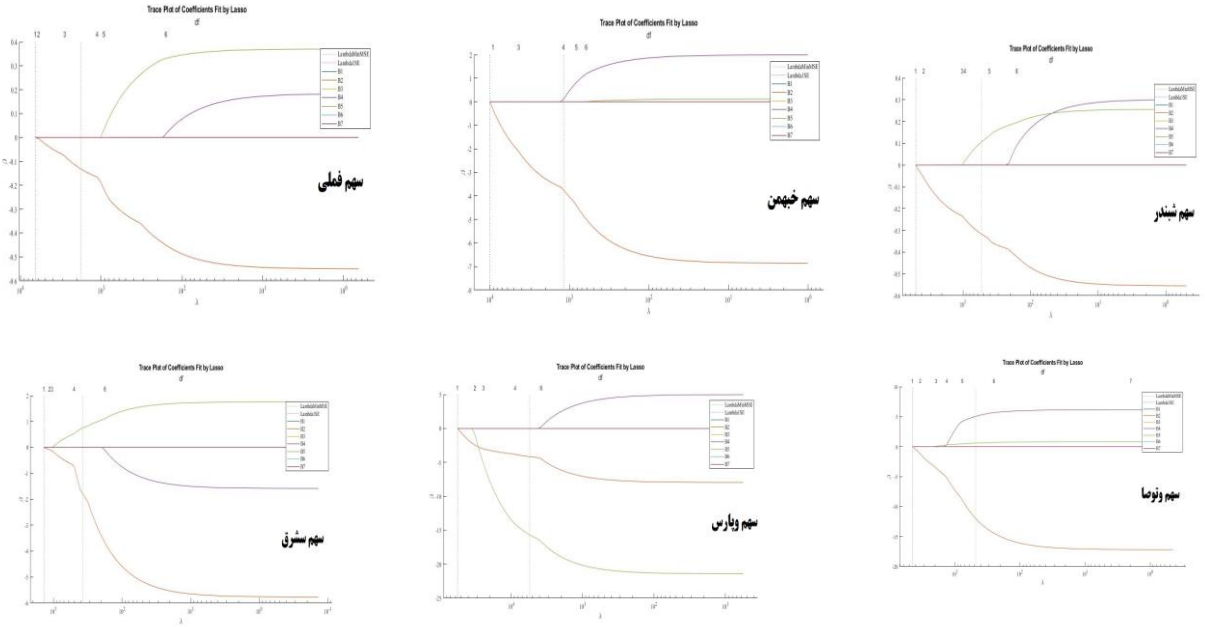
شکل ۵: نمودار احتمال قطعه به قطعه نیمه پارامتریک سهام نمونه

یافتن مقادیر بهینه ضرایب β و اعتبار سنجی بینابینی توسط رگرسیون لاسو: در این بخش، با به‌کارگیری نرم‌افزار متلب مقادیر ضرایب مجهول β برای سهام نمونه محاسبه و به‌صورت عددی و نمودار نمایش داده‌شده است. شکل شماره ۶، تغییرات مقدار بهینه ضرایب مجهول β را به ازای مقادیر مختلف λ برای درصد تغییرات دو روزه قیمت نشان می‌دهد. نتیجه‌ای که بسیار حائز اهمیت است، ترتیب به صفر رسیدن ضرایب مجهول با افزایش مقدار پنالتی است. شایان‌ذکر است، مقایسه اهمیت این عوامل با در نظرگیری مقدار عددی این ضرایب به ازای هر λ ثابت، امکان‌پذیر نیست؛ زیرا به دلیل امکان تفاوت واحدها یا اندازه‌گیری متغیرهای مستقل و نیز، تفاوت مفهومی آن‌ها نمی‌توان به این مقایسه اطمینان کرد. در این تصویر ترتیب به صفر رسیدن ضرایب مجهول با افزایش مقدار جریمه است که اولویت‌بندی عوامل را ازنظر اهمیت نشان می‌دهد. برای مثال، اگر نمودار $B7$ مربوط به نسبت قدرت فروش حقیقی و $B4$ مربوط به نسبت قدرت خرید حقیقی در نظر گرفته شود، با افزایش λ، ابتدا $B7$ و سپس $B4$ صفر خواهد شد. این رویداد بیان‌کننده آن است که اثر عامل ششم در بهبود جواب‌ها از عامل سوم کمتر است. یکی از نتایج مهمی که از شکل شماره ۶ به دست می‌آید، می‌توان تأثیر تمام پارمترهای مستقل در پیش‌بینی درصد تغییرات دو روزه قیمت را بررسی کرد. با توجه به مقدار λ_{MSE} مشاهده می‌شود که بعضی از نمودارها قبل از این مقدار و به ازای مقادیر کمتر λ صفر شده‌اند. درنتیجه قیمت سهام به تمام متغیرهای مستقل وابسته نیست. جدول شماره ۳ میزان ضرایب مجهول β را برای متغیرهای مستقل در λ_{MSE} برای درصد تغییرات دو روزه قیمت را نشان می‌دهد که با کد نویسی متلب استخراج‌شده است.

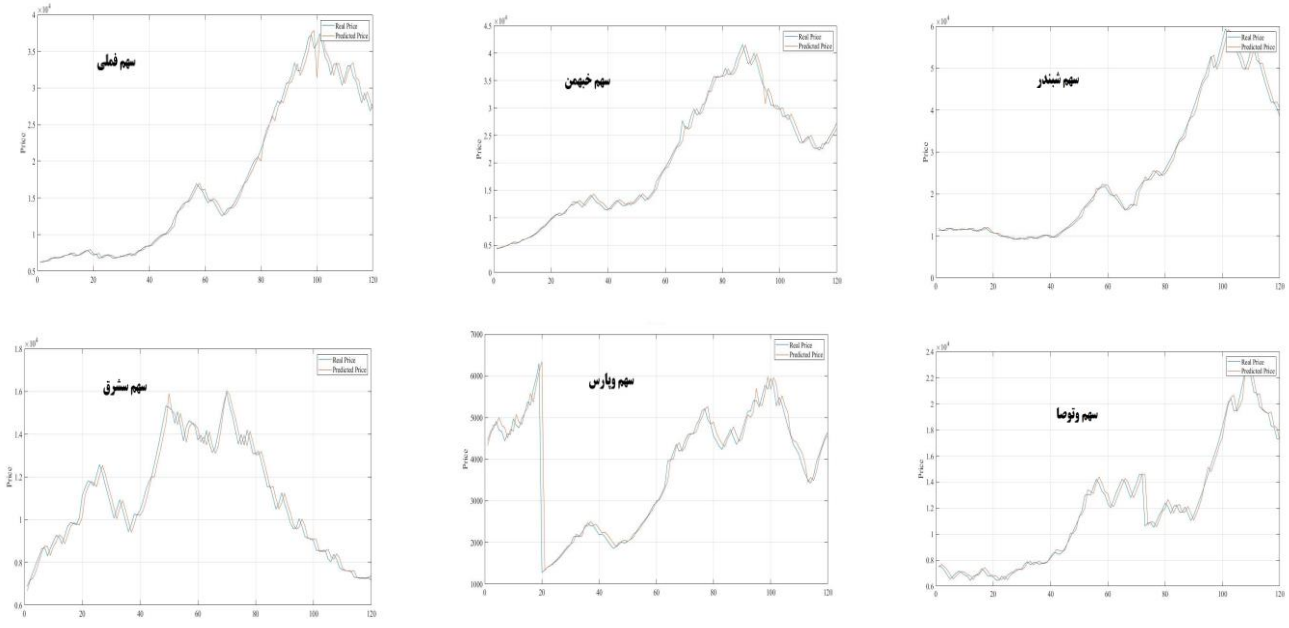
جدول ۳: ضرایب مجهول β را برای متغیرهای مستقل در λ_{MSE}

متغیر مستقل	ضرایب مجهول β
حجم معاملات حقیقی	0.000000
سرانه حجم خرید هر کد حقیقی	0.000076
سرانه مبلغ خرید هر کد حقیقی	0.002376
نسبت قدرت خریدار حقیقی	0.102964
سرانه حجم فروش هر کد حقیقی	-0.000022
سرانه مبلغ فروش هر کد حقیقی	0.000000
نسبت قدرت فروشنده حقیقی	0.000000

منبع: نتایج تحقیق



شکل ۶: نمودار احتمال قطعه به قطعه نیمه پارامتریک سهام نمونه



شکل ۷: نمودار مقایسه قیمت پیش‌بینی‌شده و قیمت واقعی سهام

نتیجه‌گیری

در این مقاله پیش‌بینی قیمت سهام در بورس اوراق بهادار تهران توسط رویکرد رگرسیون لاسو با به‌کارگیری برنامه‌نویسی در نرم‌افزار MATLAB پرداخته شد. نتایج حاصله حاکی از آن دارد که این رگرسیون به‌خوبی می‌تواند قیمت را بدون مشکلات رگرسیون‌های دیگر همانند خطی، چندگانه و ستیغی تقریب بزند. از جمله نتایج حاصله دیگر این پژوهش بررسی تأثیر رفتار معامله گران به‌عنوان بازیگران اصلی در تغییرات قیمت سهم در روزهای آتی می‌باشد. برای این منظور هفت متغیر پرکاربرد در تابلوی معاملات به‌عنوان متغیرهای مستقل در نظر گرفته شدند و تأثیرات این متغیرها بر درصد تغییرات دو روزه قیمت موردبررسی قرار گرفت. به همین منظور هفت پارامتر کلیدی در تابلو خوانی به‌عنوان متغیرهای مستقل با عناوین:

- حجم معاملات
- سرانه حجم خریدار حقیقی
- سرانه مبلغ خرید هر کد حقیقی
- نسبت قدرت خریدار
- سرانه حجم فروشنده حقیقی
- سرانه مبلغ فروش هر کد حقیقی
- نسبت قدرت فروشنده

و یک متغیر وابسته با عنوان درصد تغییرات قیمت دو روزه موردبررسی قرار گرفت. پس از حصول اطمینان از مناسب بودن مدل رگرسیونی با انتخاب دو نماد که از لحاظ برآوردها و آزمون‌های آماری مناسب‌تر بودند با بهره‌گیری از کد نویسی در محیط پایتون پیش‌بینی تغییرات قیمت دو روزه با رویکرد رگرسیون خطی لاسو برای ۲ شرکت بانک پارسیان و پالایشگاه نفت بندرعباس با نمادهای و پارس و شبندر انجام شد. همچنین، سعی شد با جداسازی متغیرهای مناسب‌تر از بقیه، مدل ساده‌تری مبتنی بر رگرسیون لاسو ارائه گردد. بدین منظور مقدار پارامتر تنظیم‌کننده λ به روش اعتبار سنجی بینابینی و بر اساس کمترین میانگین مربعات خطا تعیین و مقادیر بهینه ضرایب مجهول β متغیرهایی که تأثیر کمتری در پیش‌بینی دارند جریمه می‌شوند تا تأثیر کمتری در مدل رگرسیونی داشته باشند.

برای پاسخ به سؤالات تحقیق در خصوص دقت پیش‌بینی، مقادیر جذر میانگین مربعات خطا و درصد میانگین خطای مطلق برای هر یک از نمادها برای دو پیش‌بینی انجام شد که در پیش‌بینی

مبتنی بر رگرسیون خطی لاسو با میانگین قدر مطلق خطا کمتر از ۵ درصد در هر دو نماد موردبررسی مناسب بودن مدل رگرسیونی را تأیید می‌کند. همچنین دقت مدل برای پیش‌بینی روزهای مثبت برای نماد شبندر ۶۲ درصد و برای روزهای منفی ۸۸ درصد می‌باشد که این مقدار برای پیش‌بینی روزهای منفی بسیار قابل‌توجه می‌باشد. بعلاوه، دقت مدل برای پیش‌بینی روزهای مثبت برای نماد و پارس ۶۷,۵ درصد و برای روزهای منفی ۵۶ در صد می‌باشد که دقت مدل برای پیش‌بینی نماد و پارس کمتر از شبندر می‌باشد.

درنهایت پیشنهادهای بیشتر برای تحقیقات آتی به‌منظور پیش‌بینی قیمت سهام در بورس اوراق بهادار تهران با استفاده از روش‌های شبکه عصبی بر پایه متغیرهای تابلو خوانی، استفاده از رگرسیون لجستیک بر پایه متغیرهای تابلو خوانی و استفاده از روش‌های هوش مصنوعی بر پایه متغیرهای تابلو خوانی داده می‌شود.

منابع

۱) خنجرپناه، حسین؛ داود دوروش، سعید شوال پور، آرمین جبارزاده. ۱۳۹۷. کاربرد روش تکنیکال برای پیش‌بینی قیمت سهام: رویکرد مدل‌های احتمال غیرخطی و شبکه‌های عصبی مصنوعی. راهبرد مدیریت مالی. ۵۹-۷۹.

۲) سیف زاده، تارا و اسدی صمدی، دانیال و کفاش، عطیه،۱۳۹۹. ارائه یک سیستم خبره پیش‌بینی قیمت سهام در بورس مبتنی بر شبکه عصبی فازی، پنجمین کنفرانس ملی مهندسی کامپیوتر و بلاک چین ایران، تهران.

۳) فتح علیان، سمانه و نبوی چاشمی، سیدعلی و چیرانی، ابراهیم،۱۳۹۹. تبیین الگوی بهینه ارزیابی و قیمت‌گذاری عرضه اولیه عمومی سهام با استفاده از تکنیک‌های تصمیم‌گیری چند معیاره فازی، رگرسیون، شبکه عصبی و الگوریتم ژنتیک، مهندسی مالی و مدیریت اوراق بهادار، دوره: ۱۱، شماره:۴۲

۴) مشاری، محمد؛ حسین دیده خانی، کاوه خلیلی دامغانی، ابراهیم عباسی. ۱۳۹۸. طراحی مدل هوشمند ترکیبی جهت پیش‌بینی نقاط طلایی قیمت سهام. دانش سرمایه‌گذاری. ۴۵-۶۵.

۵) موسوی، سید علیرضا؛ افسانه غلامی. ۱۳۹۷. استفاده از الگوریتم ترکیبی عصبی کرم شب‌تاب و روش رگولاسیون بیزین جهت پیش‌بینی قیمت سهام. مهندسی مالی و مدیریت اوراق بهادار. ۲۹۵-۳۲۱.

6) Beyaz, E. Tekiner, F. Zeng, X. J. & Keane, J. (2018, June). Comparing technical and fundamental indicators in stock price forecasting. In 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE ۱۶th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) (pp. 1607-1613). IEEE.

7) Chen, Y. J. Chen, Y. M. Tsao, S. T. & Hsieh, S. F. (2018). A novel technical analysis-based method for stock market forecasting. Soft Computing, 22(4), 1295-1312

8) Ciner, C. (2019). Do industry returns predict the stock market? A reprise using the random forest. The Quarterly Review of Economics and Finance, 72, 152-158.

9) Daelemans, B. Daniels, J. P. & Nourzad, F. (2018). Free Trade Agreements and Volatility of Stock Returns and Exchange Rates: Evidence from NAFTA. Open Economies Review, 29(1), 141-163.

10) Feng Zhou Qun Zhang, Didier Sornette, Liu Jiang.(2019). Cascading logistic regression onto gradient boosted decision trees for forecasting and trading stock indices. Applied Soft Computing Journal 84 (2019) 105747

11) Guang Liu & Xiaojie Wang(2019). & A new metric for individual stock trend prediction in Engineering Applications of Artificial Intelligence 82 (2019) 1-12-Elsevier

12) Lahmiri, S. (2018). Minute-ahead stock price forecasting based on singular spectrum analysis and support vector regression. Applied Mathematics and Computation, 320, 444-451

13) Mittnik, S. Robinzonov, N. & Spindler, M. (2015). Stock market volatility: identifying major drivers and the Nature of Their Impact. Journal of Banking & Finance, 58, 1-14.

14) Padhi, D. K. Padhy, N. & Mishra, J. (2020, March). Intraday Stock Prices Forecasting Using an Autoregressive Model. In 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA) (pp. 1-6). IEEE.

15) R. Dash, P.K. Dash(2016). "A hybrid stock trading framework integrating technical analysis with machine learning techniques", J Finance Data Sci, 2 (1), pp. 42-57.

16) Rodriguez J. V. Torra S. Felix J. A. (2005); "STAR andANN models: Forecasting Performance on SpanishIbex-35 Stock Index", Journal of EmpiricalFinance, no. 12

17) Roy, S. S. Mittal, D. Basu, A. & Abraham, A. (2015). Stock market forecasting using LASSO linear regression model. In Afro-European Conference for Industrial Advancement (pp. 371-381). Springer, Cham.

18) Setiawan, B. (2018). Lasso technique application in stock market modelling: An empirical evidence in indonesia. Sriwijaya International Journal of Dynamic Economics and Business, 2(1), 51-62.

19) Shahrokh Asadi.(2018). Evolutionary fuzzification of RIPPER for regression: case study of stock prediction. Neurocomputing 2019, Pages 121-137

20) Sharma, A. & Mehra, A. (2016). Financial analysis based sectoral portfolio optimization under second order stochastic dominance. Annals of Operations Research, 1-27

21) Teixeira, L. A. & De Oliveira, A. L. I. (2010). A method for automatic stock trading combining technical analysis and nearest neighbor classification. Expert systems with applications, 37(10), 6885-6890

22) Wang, S. Ji, B. Zhao, J. Liu, W. & Xu, T. (2018). Predicting ship fuel consumption based on LASSO regression. Transportation Research Part D: Transport and Environment, 65, 817-824.

23) Yang, X. & Wen, W. (2018). Ridge and lasso regression models for cross-version defect prediction. IEEE Transactions on Reliability, 67(3), 885-896.

24) Yensen Ni,Paoyu Huang, Yuhsin Chen. (2019) Board structure, considerable capital, and stock price overreaction informativeness in terms of technical indicators. The North American Journal of Economics and Finance Volume 48, April 2019, Pages 514-528

25) Zhang, J. Teng, Y. F. & Chen, W. (2019). Support vector regression with modified firefly algorithm for stock price forecasting. Applied Intelligence, 49(5), 1658-1674

واژگان

فارسی	لاتین
تحلیل تکنیکال	Technical analysis
رگرسیون لاسو	LASSO regression
متغیر پاسخ	Response
متغیر توصیفی	Explanatory
بازگشت	Regress
رگرسیون خطی ساده	Simple Linear Regression
رگرسیون خطی چندگانه	Multiple Linear Regression
صفحه	plane
ابر صفحه	Hyper plane
رگرسیون ستیغی (ریج)	Ridge regression
کمینه‌سازی مجموع مربعات خطا	Ordinary least Square
بیش‌برازش	Over fitting
انقباض کمترین قدرمطلق	Least Absolute Shrinkage and Selection Operator
تابع جریمه	Penalty
روش تحلیل سلسله مراتبی فازی	AHP
الگوریتم ریپ	Repeated Incremental Pruning
رگرسیون بردار پشتیبانی	SVR
شبکه عصبی	LM-BP

Predicting stock price via Lasso regression in Tehran stock exchange

Receipt: 14/11/2023 **Acceptance: 23/09/2024**

Amir Sadeghi[1]
Amir Kamali Dolat Abadi[2]

Abstract

Predicting capital market behavior has always been one of the challenges for market participants. Over the years, trends forecasting methods have evolved and knowledge of predicting behavior and stock prices is still evolving. It is important to analyze the past behavior of stock prices based on technical methods. The technical method is often focused on price changes, moving averages and trading volume. In this research, we have tried to study the trading ratios based on table reading and analyze the past behavior and movements of real traders in the form of Lasso linear regression.

Keywords

sign reading, price change percentage, stock price forecast, lasso regression

1-Assistant Professor, Department of Applied Mathematics, Parand branch, Islamic Azad University, Tehran. Iran. (Corresponding Author) Drsadeghi.iau@gmail.com
2-Assistant Professor, Department of Industrial Engineering, Parand branch, Islamic Azad University, Tehran, Iran. amir.kamali2002@gmail.com