

پیش‌بینی شاخص کل قیمت بورس تهران با استفاده از ترکیب خبرگان

محمد رضا یگانگی و دکتر رحیم چینی‌پرداز*

تاریخ وصول: ۱۳۹۰/۱/۳۰ تاریخ پذیرش: ۱۳۹۰/۹/۲۸

چکیده

پیش‌بینی در بازارهای مالی همواره مورد توجه پژوهشگران و سرمایه‌گذاران بوده است. در این میان پیش‌بینی شاخص بازار از اهمیت ویژه‌ای برخوردار است، به طوری که همزمان با توسعه مدل‌های سری زمانی، روش‌های پیش‌بینی شاخص در بازارهای مالی نیز بسیار توسعه یافته‌اند. در این مقاله با استفاده از ترکیب خبرگان، مدلی برای پیش‌بینی شاخص بورس تهران ارائه گردیده است. نتایج این پژوهش نشان داد که مدل ارائه شده توانایی بالایی در مدل‌بندی و پیش‌بینی شاخص بورس تهران دارد.

طبقه بندی JEL: C22, C48, C53

واژه‌های کلیدی: شاخص کل قیمت بورس تهران، پیش‌بینی، مدل خود بازگشت آمیخته، ترکیب خبرگان، شبکه‌های عصبی محلی- سراسری خطی

* به ترتیب، دانشجوی دکتری و استاد گروه آمار دانشگاه شهید چمران اهواز

(rchinipardaz@yahoo.com)

۱- مقدمه

پیش‌بینی شاخص بورس همواره مورد توجه بسیاری از فعالان بازارهای اقتصادی و مالی، همچنین تحلیلگران و برنامه‌ریزان اقتصادی بوده است. در سال‌های اخیر مدل‌های مختلفی برای مدل‌سازی و پیش‌بینی شاخص در بازارهای مختلف جهان به کار گرفته شده است. در این میان شبکه‌های عصبی جایگاه ویژه‌ای را به خود اختصاص داده‌اند، به طوری که بسیاری از مدل‌های ارائه شده، با استفاده از شبکه‌های عصبی طراحی گردیده‌اند (کرولنر^۱ و دیگران، ۲۰۱۰).

اگر چه شبکه‌های عصبی در مدل‌سازی سری‌های زمانی بسیار توانمند هستند، ولی در شرایطی که فضای ورودی شبکه دارای ساختاری خوشه‌ای باشد، به طوری که ساختار خوشه‌ها در طول زمان تغییر کند، شبکه‌های عصبی عموماً توانایی مدل‌سازی کل تغییرات داده‌ها در طول زمان را ندارند. جاکوب^۲ و دیگران (۱۹۹۱) برای رفع این مشکل، تعمیمی از شبکه‌های عصبی را تحت عنوان ترکیب خبرگان ارائه کرده‌اند. مدل‌های ترکیب خبرگان با وجود توانایی بالا در مدل‌سازی، به دلیل پیچیدگی در کاربرد، کمتر مورد استفاده قرار گرفته‌اند.

در بورس اوراق بهادار تهران نیز مدل‌هایی برای پیش‌بینی شاخص کل قیمت ارائه شده است. سینایی و دیگران (۱۳۸۴) و فلاح شمس و دلنواز اصغری (۱۳۸۸) با استفاده از شبکه‌های عصبی به پیش‌بینی شاخص کل بورس تهران پرداخته و نشان داده‌اند که شبکه‌های عصبی در مقایسه با مدل‌های کلاسیک خطی دقت بالاتری دارند. عرفانی (۱۳۸۷) با استفاده از مدل‌های *ARFIMA* به پیش‌بینی شاخص بورس تهران پرداخته و نشان داده‌اند که شاخص بورس تهران یک سری زمانی با حافظه طولانی است. کشاورز حداد و صمدی (۱۳۸۸) با استفاده از این واقعیت که نوسانات شاخص در بازه‌های زمانی مختلف تغییر می‌کند، از یک مدل *ARFIMA-FIGARCH* برای مدل‌سازی شاخص بور تهران استفاده کرده‌اند.

منجمی و دیگران (۱۳۸۸) با استفاده از روش‌های شبکه‌ی عصبی فازی و الگوریتم‌های ژنتیک قیمت سهام در بورس اوراق بهادار تهران را پیش‌بینی و تجزیه و تحلیل کرده‌اند.

¹ Krollner

² Jacobs

در این مقاله به پیش‌بینی شاخص کل بورس اوراق بهادار تهران با استفاده از ترکیب خبرگان پرداخته شده است. برای این منظور یک شبکه‌ی عصبی محلی-سراسری خطی گوسی (یگانگی، ۱۳۸۸)، که یک مدل ترکیب خبرگان با خبره‌های خطی است، استفاده شده است.

در ادامه‌ی مقاله، در بخش دوم به معرفی اجمالی مدل‌های ترکیب خبرگان پرداخته، دو عضو مهم این خانواده معرفی می‌شوند. بخش سوم به معرفی شبکه‌ی عصبی محلی-سراسری خطی گوسی و ویژگی‌های آن می‌پردازد. در بخش چهارم، با استفاده از مدل ارائه شده در بخش سوم، مدلی برای پیش‌بینی شاخص کل بورس اوراق بهادار تهران، ارائه می‌شود. بخش پنجم نیز به بحث در مورد توانایی مدل ارائه شده در پیش‌بینی شاخص کل و ارائه نتایج عملی برای پژوهش‌های آتی اختصاص داده شده است.

۲- ترکیب خبرگان

برای مدل‌سازی در شرایطی که مشاهدات به زیر مجموعه‌های مختلفی در بخش‌های مختلف فضای مشاهدات تقسیم شوند، خانواده‌ای از مدل‌ها با عنوان مدل‌های ترکیب خبرگان توسط همپ‌شیر و وایبل^۳ (۱۹۸۹) و جاکوب و همکاران (۱۹۹۱) معرفی شده است. مدل‌های ترکیب خبرگان را می‌توان تعمیمی از مدل‌های رگرسیون خطی محلی، مدل‌های وابسته به حالت (پریستلی،^۴ ۱۹۸۰ و ۱۹۸۸) و توزیع‌های آمیخته (پیرسون،^۵ ۱۸۹۴) دانست. ایده‌ی اصلی در مدل‌های ترکیب خبرگان، استفاده از مدل‌های مختلف در بخش‌های مختلف فضای مشاهدات است. به طوری که هر یک از مدل‌های مورد استفاده برای بخشی از فضای مشاهدات برازش خوبی ارائه کنند یا به عبارتی دیگر خبره باشند.

یک مدل ترکیب خبرگان، سیستمی شامل مجموعه‌ای از خبرگان و یک شبکه قطع متناوب است. هر خبره یک شبکه‌ی عصبی پیش‌خور یا در حالت کلی تر یک مدل برای برازش به داده‌ها است. همچنین ورودی تمامی خبره‌ها یکسان است. خروجی‌های هر خبره نیز مقدار پیش‌بینی آن خبره از متغیر مورد پیش‌بینی است.

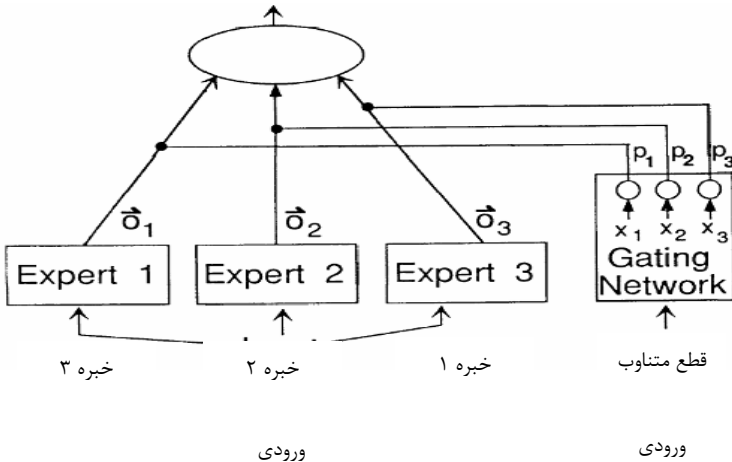
³ Hampshire and Waibel

⁴ Priestly

⁵ Pearson

شبکه‌ی قطع متناوب یک شبکه‌ی عصبی پیش‌خور است که می‌تواند ورودی‌های مشترک با خبره‌ها داشته باشد. یا ورودی‌های متفاوتی در مقایسه با خبره‌ها دریافت کند. تعداد خروجی‌های شبکه‌ی قطع متناوب نیز برابر با تعداد خبره‌های مورد استفاده در سیستم است. اگر خروجی‌های نرمال شده شبکه‌ی قطع متناوب در نظر گرفته شود، خروجی زام شبکه قطع متناوب را می‌توان به‌عنوان احتمال انتخاب خبره زام برای یک ورودی معین در نظر گرفت.^۶ نقش شبکه‌ی قطع متناوب در مدل‌های ترکیب خبرگان، افزایش فضای مشاهدات به زیربخش‌هایی است که هر یک از خبرگان در آنها کارایی مطلوبی داشته باشند. نمودار (۱) نمای کلی یک مدل ترکیب خبرگان را نشان می‌دهد (جاکوبز و دیگران، ۱۹۹۱).

نمودار ۱: مدل ترکیب خبرگان شامل سه خبره



مأخذ: جاکوبز و دیگران (۱۹۹۱)

برای برآورد پارامترهای مدل‌های ترکیب خبرگان الگوریتم‌های مختلفی ارائه شده است. به‌عنوان نمونه، می‌توان به برآورد بیشینه درست‌نمایی و الگوریتم EM که توسط جاکوبز و دیگران (۱۹۹۱) و جردن و جاکوبز^۷ (۱۹۹۲) اشاره کرد.

^۶ خروجی‌های شبکه قطع متناوب لزوماً اندازه‌های احتمال نیستند، ولی با در نظر گرفتن شرایطی می‌توان خروجی‌ها را به اندازه‌های احتمال تبدیل کرد.

^۷ Jordan and Jacobs

به عنوان یک حالت خاص از مدل‌های ترکیب خبرگان، می‌توان خروجی‌های شبکه‌ی قطع متناوب را ثابت در نظر گرفت. به این ترتیب، شبکه‌های عصبی پیش‌خور، مدل‌های آمیخته متناهی، توزیع‌های آمیخته و مدل‌های MAR حالات خاصی از مدل‌های ترکیب خبرگان به‌شمار خواهند آمد. در دو دهه‌ی اخیر، با استفاده از ایده ترکیب خبرگان، مدل‌های پیچیده‌تری برای مدل‌سازی و پیش‌بین سری‌های زمانی ارائه شده است. در ادامه این بخش دو خانواده از این مدل‌ها مورد بررسی قرار می‌گیرند.

۲-۱- مدل خود بازگشت آمیخته

فرض کنید y_t به شرط y_{t-1}, \dots, y_{t-p} دارای تابع چگالی احتمال آمیخته^۸ با m مؤلفه‌ی چگالی نرمال باشد، در این صورت سری زمانی y_t را می‌توان به صورت یک مدل خود بازگشت آمیخته^۹ از مرتبه‌ی $(m, p_1, p_2, \dots, p_m)$ ، یا به اختصار $MAR(m, p_1, p_2, \dots, p_m)$ ، به صورت زیر فرمول بندی کرد:

$$y_t = \sum_{i=1}^m \pi_i (\phi_{i1} y_{t-1} + \phi_{i2} y_{t-2} + \dots + \phi_{ip_i} y_{t-p_i}) + \varepsilon_t \quad (1)$$

مدل MAR یک ترکیب محدب از m مدل خود بازگشت است. ونگ و لی^{۱۰} (۲۰۰۰) نشان داده‌اند که واریانس شرطی در مدل‌های $MAR(m, p_1, p_2, \dots, p_m)$ ثابت نیست و به این ترتیب، این مدل‌ها برای مدل‌سازی داده‌هایی با واریانس متغیر مناسب‌اند. علاوه بر این، به دلیل شکل توزیع‌های آمیخته، مدل‌های MAR به سادگی می‌توانند مشاهداتی را که دارای توزیع‌های دم‌سنگین هستند، مدل‌بندی کنند (ونگ و لی، ۲۰۰۰).

از دیدگاه ترکیب خبرگان، MAR یک مدل ترکیب خبرگان است که در آن خروجی‌های شبکه‌ی قطع متناوب مقادیر ثابتی هستند. به این ترتیب، مدل MAR برای مدل‌سازی، فضای مشاهدات را به m زیر بخش تقسیم می‌کند و برای هر یک از این بخش‌ها یک تابع توزیع احتمال در نظر می‌گیرد. بنابراین توزیعی که برای مدل‌سازی مورد استفاده قرار می‌گیرد، نسبت به مدل‌های پارامتری انعطاف پذیرتر است. با این وجود، وزن‌های ترکیب مؤلفه‌های تابع توزیع در تمام فضای مشاهدات

⁸ Mixture distribution function

⁹ Mixture Autoregressive

¹⁰ Wong and Li

ثابت فرض شده‌اند. از این رو، اگر ساختار توزیع در طول زمان تغییر کند یا وزن-های مؤلفه‌ها به زیربخش‌های مختلف فضای مشاهدات وابسته باشد، مدل‌های MAR چندان کارا نیستند. در این شرایط می‌توان از مدل‌های کلی‌تر ترکیب خبرگان استفاده کرد.

۲-۲- شبکه‌های عصبی محلی- سراسری خطی

مدل شبکه‌ی عصبی محلی- سراسری خطی^{۱۱}، L^xGNN ، یک نگاهت غیر خطی به صورت زیر است:

$$y_t = \sum_{i=1}^m (d_i x_t + b_i) B(x_t; \varphi_{B_i})$$

که در آن $x_t \in \mathbb{R}^q$ بردار مقادیر گذشته سری y_t و در صورت وجود، متغیرهای خارجی اثرگذار بر سری و ε_t یک فرآیند اغتشاش گوسی محض است. توابع $L(x_t) = (a'x_t + b)$ و $B(\cdot)$ از فضای \mathbb{R}^q به \mathbb{R} تعریف شده است و به ترتیب با نام تابع تقریب و تابع فعالیت شناخته می‌شوند. تابع تقریب یک تابع سراسری است، یعنی ناحیه‌ی فعالیت آن حد اقل از یک طرف بی‌کران است. در مقابل تابع فعالیت دارای ویژگی محلی است، یعنی به‌ازای مقادیر کراندار از بردارهای x_t خروجی غیر صفر دارد و مقادیر خروجی آن در بقیه‌ی فضای مشاهدات صفر خواهد بود. به‌عنوان یک نمونه از تابع فعالیت می‌توان از تفاضل دو تابع لجستیک استفاده کرد (فاریناس^{۱۲} و دیگران، ۲۰۰۴). در این صورت تابع فعالیت به شکل زیر تعریف خواهد شد.

$$B(x_t; \varphi_{B_i}) = \left[\frac{1}{1 + \exp(\gamma_i (d_i' x_t - \beta_i^{(1)}))} - \frac{1}{1 + \exp(\gamma_i (d_i' x_t - \beta_i^{(2)}))} \right]$$

از نظر معماری شبکه‌های عصبی، یک شبکه‌ی عصبی محلی- سراسری، یک شبکه پیش‌خور با m جفت نرون است. از هر جفت نرون یکی دارای تابع فعالیت محلی و دیگری دارای تابع فعالیت سراسری است. در مدل L^xGNN تابع فعالیت در نرون سراسری، یک تابع همانی است. مدل L^xGNN یک مدل ترکیب خبرگان با خبره‌های خطی است. نرون‌های محلی نیز شبکه‌ی قطع متناوب را در این مدل

¹¹ Linear Local-Global Neural Network

¹² Farinas

تشکیل می‌دهند (فاریناس و دیگران، ۲۰۰۴). علاوه بر این، بسیاری از مدل‌های سری زمانی غیر خطی را می‌توان به صورت یک $L^X GNN$ نمایش داد (همان منبع). به دلیل پیچیدگی بالای مدل $L^X GNN$ ، پیاده‌سازی و استفاده عملی از این مدل‌ها با محدودیت‌هایی همراه است. به عنوان نمونه می‌توان به مسائل مربوط به برآورد بیشینه درست‌نمایی پارامترهای $L^X GNN$ و ویژگی‌های احتمالی آن اشاره کرد (فاریناس و دیگران، ۲۰۰۴ و یگانگی، ۱۳۸۸). در بخش بعد با استفاده از ایده‌ی مدل‌های MAR و $L^X GNN$ ، خانواده جدیدی از مدل‌های ترکیب خبرگان معرفی می‌شود.

۳- مدل خود بازگشت آمیخته با تابع وزن گوسی

یک مدل خود بازگشت آمیخته با تابع وزن گوسی به صورت زیر تعریف می‌شود (یگانگی، ۱۳۸۸):

$$y_t = E(y_t | \mathbf{x}_t) + \varepsilon_t \quad (2)$$

$$= \sum_{i=1}^m g^{(i)}(\mathbf{x}_t) \left(\mu_{y_t}^{(i)} + \Sigma_{y_t}^{(i)} (\Sigma_{y_t}^{(i)})^{-1} (\mathbf{x}_t - \mu_{\mathbf{x}_t}^{(i)}) \right) + \varepsilon_t$$

$$Var(y_t | \mathbf{x}_t) = \sum_{i=1}^m g^{(i)}(\mathbf{x}_t) \left(\sigma_{y_t}^{2(i)} + \Sigma_{y_t}^{(i)} (\Sigma_{y_t}^{(i)})^{-1} \Sigma_{y_t}^{(i)} \right) \quad (3)$$

$$+ \sum_{i=1}^m g^{(i)}(\mathbf{x}_t) \left[\left(\mu_{y_t}^{(i)} + \Sigma_{y_t}^{(i)} (\Sigma_{y_t}^{(i)})^{-1} (\mathbf{x}_t - \mu_{\mathbf{x}_t}^{(i)}) \right)^2 - \left(\sum_{i=1}^m g^{(i)}(\mathbf{x}_t) \left(\mu_{y_t}^{(i)} + \Sigma_{y_t}^{(i)} (\Sigma_{y_t}^{(i)})^{-1} (\mathbf{x}_t - \mu_{\mathbf{x}_t}^{(i)}) \right) \right)^2 \right]$$

که در آن $\mathbf{x}_t = (y_{t-1}, y_{t-2}, \dots, y_{t-p}) \in \mathfrak{R}^p$ میانگین y_t و $\mu_{\mathbf{x}_t}^{(i)}$ بردار میانگین \mathbf{x}_t در مؤلفه i ام هستند. ماتریس کوواریانس و توابع وزن نیز به صورت زیر تعریف می‌شوند:

$$Cov^{(i)}(y_t, \mathbf{x}_t) = \begin{pmatrix} \sigma_{y_t}^{2(i)} & \Sigma_{y_t}^{(i)} \\ \Sigma_{y_t}^{(i)} & \Sigma_{y_t}^{(i)} \end{pmatrix}_{(p+1) \times (p+1)}$$

$$g^{(i)}(\mathbf{x}_t) = \frac{\pi_i |\Sigma_{y_t}^{(i)}|^{-\frac{1}{\nu}} \exp \left\{ -\frac{1}{\nu} (\mathbf{x}_t - \mu_{\mathbf{x}_t}^{(i)})' (\Sigma_{y_t}^{(i)})^{-1} (\mathbf{x}_t - \mu_{\mathbf{x}_t}^{(i)}) \right\}}{\sum_{j=1}^m \pi_j |\Sigma_{y_t}^{(j)}|^{-\frac{1}{\nu}} \exp \left\{ -\frac{1}{\nu} (\mathbf{x}_t - \mu_{\mathbf{x}_t}^{(j)})' (\Sigma_{y_t}^{(j)})^{-1} (\mathbf{x}_t - \mu_{\mathbf{x}_t}^{(j)}) \right\}}$$

مدل (۲) یک مدل شبکه‌ی عصبی محلی- سراسری است. توابع وزن در این مدل، نقش توابع فعالیت در شبکه‌ی عصبی محلی- سراسری ایفا می‌کنند. از دیدگاه ترکیب خبرگان نیز، توابع وزن یک شبکه‌ی قطع متناوب را تشکیل می‌دهند که خروجی‌های آن احتمالات شرطی هستند. با توجه به آنکه توابع وزن در مدل (۲) از روی توزیع گوسی ساخته شده‌اند، مدل (۲) را می‌توان یک مدل آمیخته‌ی متناهی با توابع وزن گوسی دانست. همچنین به دلیل آنکه این توابع وزن، توابع فعالیت محلی هستند و مدل‌های خطی، نقش توابع تقریب سراسری (خبره‌ها در مدل ترکیب خبرگان) را ایفا می‌کنند، مدل (۲) را می‌توان مدل شبکه‌ی عصبی محلی- سراسری خطی گوسی یا به اختصار $GL^X GNN$ نامید.

در مورد مدل $GL^X GNN$ باید به چند نکته دقت شود. نخست آنکه ضرایب و ورودی‌های توابع وزن با ضرایب و ورودی‌های خبره‌ها مشترکند. بنابراین برای برآورد ضرایب از الگوریتم‌هایی مثل درست‌نمایی متمرکز که ضرایب خبره‌های خطی و ضرایب شبکه قطع متناوب را به‌طور جداگانه برآورد می‌کند، نمی‌توان استفاده کرد. با این وجود، الگوریتم‌های مبتنی بر مشتقات جزئی مرتبه‌ی دوم و صافی کالمن که در آموزش شبکه‌های عصبی بسیار مرسوم هستند، می‌توانند مورد استفاده قرار گیرند. همچنین به دلیل استفاده از توزیع‌های آمیخته در طراحی این مدل‌ها، به‌سادگی می‌توان از الگوریتم‌هایی مانند الگوریتم EM برای یافتن برآورد بیشینه درست‌نمایی پارامترهای مدل استفاده کرد.

مسأله‌ی دیگری که در استفاده از مدل $GL^X GNN$ باید مد نظر قرار گیرد، قسمت سیستماتیک مدل تنها شامل امید ریاضی شرطی نیست. چنان‌که در ابتدای همین بخش اشاره شد، بر اساس رابطه‌ی (۳)، واریانس جمله خطا را می‌توان مدل‌بندی کرد. به این ترتیب خطاهای مدل عبارت خواهند بود از:

$$w_t = \frac{y_t - E(y_t | \mathbf{x}_t)}{\sqrt{Var(y_t | \mathbf{x}_t)}} \quad (4)$$

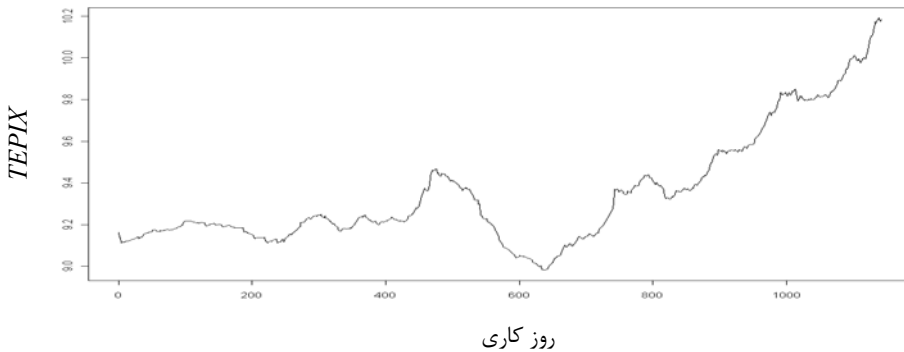
که در آن w_t ها دارای میانگین صفر و واریانس یک هستند.

۴- مدل سازی و پیش‌بینی شاخص کل قیمت بورس تهران

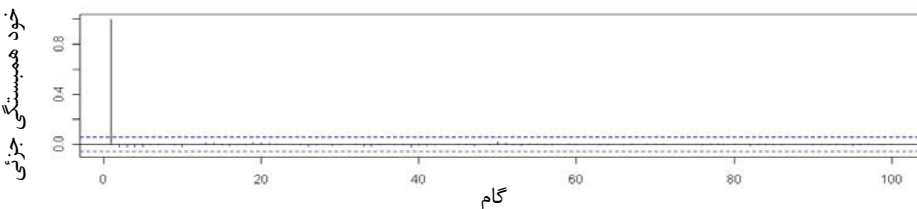
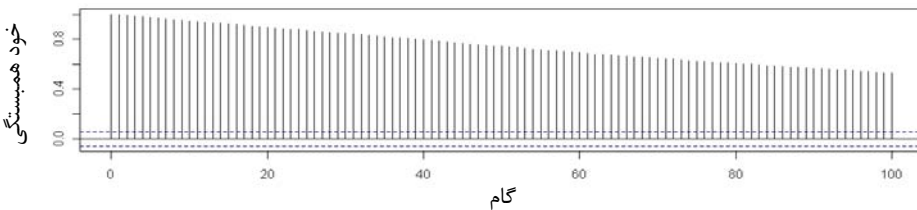
داده‌های مورد بررسی در این بخش، لگاریتم داده‌های شاخص کل قیمت بورس اوراق بهادار تهران ($TEPIX$) هستند که به‌صورت روزانه توسط سایت رسمی شرکت بورس اوراق بهادار تهران در اختیار عموم قرار می‌گیرند. به این ترتیب، y_t عبارت

از لگاریتم عدد شاخص کل قیمت بورس تهران در روز کاری t ام است. دلیل استفاده از تبدیل لگاریتمی روی داده‌ها، وجود نوسانات شدید در ساختار داده‌های شاخص بورس است. تبدیل لگاریتمی بیشترین تأثیر را بر داده‌های بزرگتر خواهد داشت. به این ترتیب، با کاهش نوسانات موجود در داده‌ها، برای مدل سازی نیاز به تعداد مولفه‌های کمتری خواهد بود. بازه‌ی مورد بررسی نیز از تاریخ ۱۳۸۵/۵/۱ تا ۱۳۹۰/۱/۳۱ است ($t=1,2,\dots,1142$). نمودار سری زمانی مورد بررسی و توابع خود همبستگی و خود همبستگی جزئی آن در نمودار (۲) و (۳) نشان داده شده است.

نمودار ۲: نمودار سری زمانی $TEPIX$



نمودار ۳: نمودار تابع خود همبستگی و خود همبستگی جزئی سری زمانی $TEPIX$



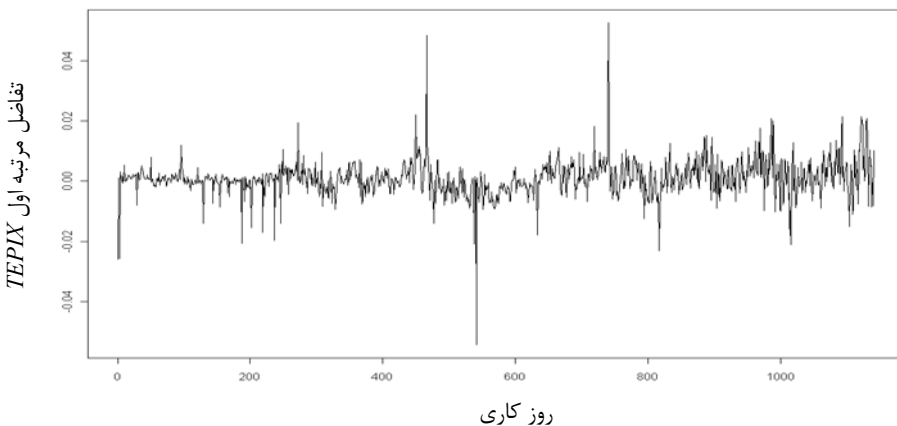
با توجه به نمودارهای (۲) و (۳) سری زمانی *TEPIX* نا ایستا است. با یک مرحله‌ی تفاضل‌گیری، سری زمانی به حالت ایستایی نزدیکتر خواهد شد. همچنین واریانس سری زمانی به‌طور چشمگیری کاهش می‌یابد. کاهش واریانس سری زمانی می‌تواند به کاهش پیچیدگی مدل و افزایش سرعت الگوریتم آموزش کمک کند. با افزایش مرتبه تفاضل‌گیری واریانس سری افزایش می‌یابد. به این ترتیب مدل به سری جدیدی به‌صورت زیر برازش خواهد شد:

$$z_t = (1 - B)y_t = y_t - y_{t-1}$$

که در آن سری z_t در حقیقت نشانگر تغییر شاخص در روز t ام نسبت به روز قبل از آن است.

نمودارهای (۴) و (۵)، نمودار سری *TEPIX* و نمودارهای خود همبستگی و خودهمبستگی جزئی را پس از یک بار تفاضل‌گیری نشان می‌دهند.

نمودار ۴: نمودار سری زمانی حاصل از تفاضل‌گیری مرتبه‌ی اول *TEPIX*

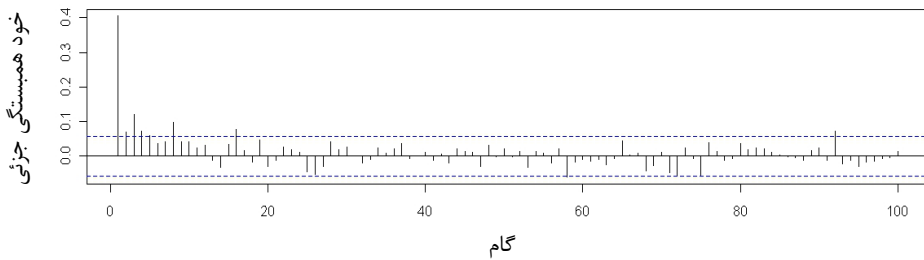
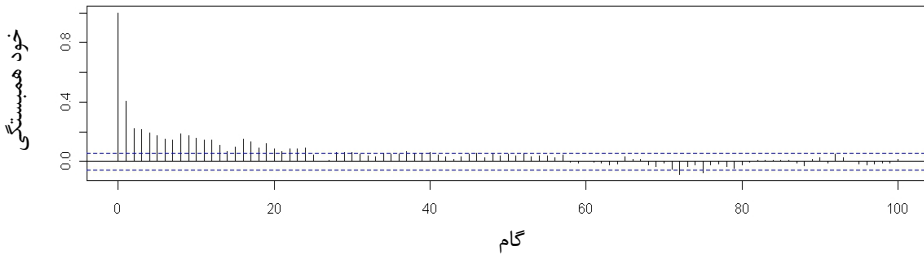


با افزایش مرتبه تفاضل‌گیری واریانس سری افزایش می‌یابد. بنابراین مدل‌های کلاسیک *ARIMA* برای حل این مسأله‌ی مفید نخواهند بود (پالما،^{۱۳} ۲۰۰۷) و استفاده از مدل‌های غیر خطی برای برازش به داده‌ها مناسب‌تر به نظر می‌رسد. نکته‌ی دیگری که باید به آن توجه کرد آن است که شکل توزیع احتمال این داده‌ها یک شکل نامتقارن است (نمودارهای (۶) و (۷)) که شباهتی به

¹³ Palma

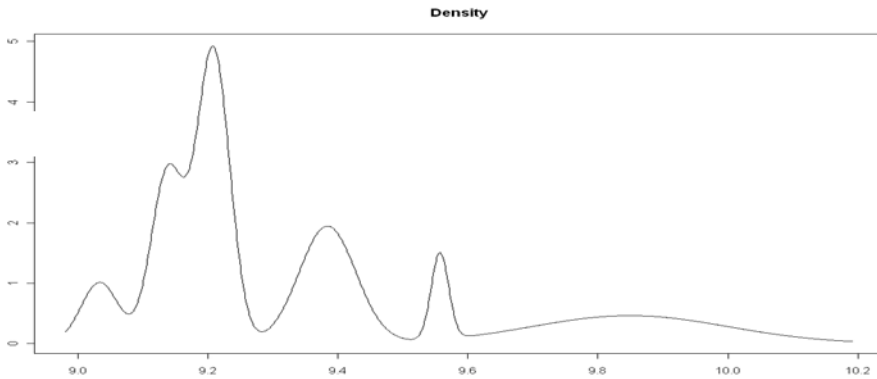
توزیع‌های مورد استفاده در مدل‌های غیر خطی پارامتری ندارد. لازم به یادآوری است که در مدل‌های سری زمانی پارامتری از توزیع‌های متقارن مانند توزیع نرمال و یا توزیع t استیودنت استفاده می‌شود.

نمودار ۵: نمودار خود همبستگی و خودهمبستگی جزئی سری زمانی Z_t

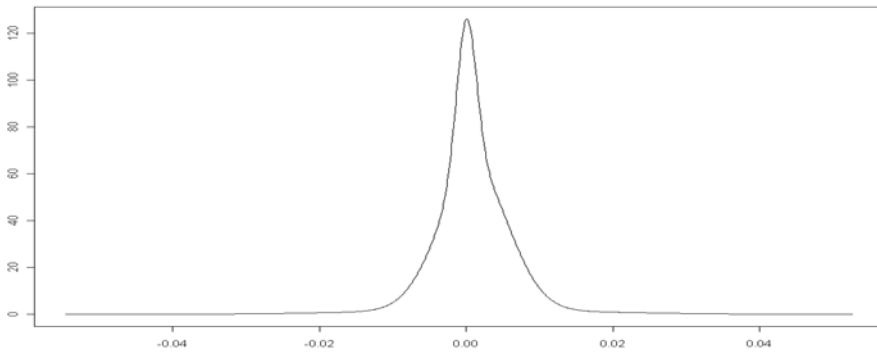


بنابراین با توجه به نمودارهای (۶) و (۷) توزیع برازش شده به داده‌ها، شکل متقارنی نداشته است و بنابراین مدل‌های غیر خطی پارامتری برای برازش به این داده‌ها مناسب نخواهند بود.

نمودار ۶: نمودار برآورد تابع چگالی $TEPIX$ (توزیع آمیخته نرمال)



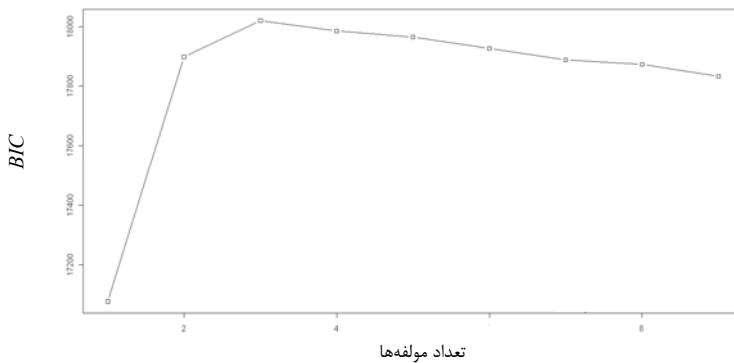
نمودار ۷: نمودار برآورد تابع چگالی Z_t (توزیع آمیخته نرمال)



علاوه بر این، برای مدل سازی Z_t باید دقت شود که تغییرات شدید نوسانات سری در طول زمان ممکن است موجب ناپایداری در واریانس شود. بنابراین مدل‌هایی مثل شبکه‌های عصبی پرسپترون برازش مناسبی برای این داده‌ها ارائه نخواهند کرد (یگانگی، ۱۳۸۸). با توجه به اینکه مدل GL^*GNN دارای واریانس شرطی متغیر است، انتظار می‌رود که با استفاده از این مدل‌ها برازش مناسب‌تری نسبت به دیگر مدل‌های شبکه‌ی عصبی به دست آید. برای انتخاب تعداد گام‌های وابستگی (ورودی‌های مدل) و تعداد مؤلفه‌ها در مدل GL^*GNN می‌توان از معیار BIC استفاده نمود. در این پژوهش برای برازش مدل و برآورد پارامترها، از الگوریتم EM استفاده شده است. برای پیاده سازی الگوریتم EM نیز از بسته $mclust$ نرم‌افزار R نسخه ۲/۱۲ استفاده شده است. با توجه به

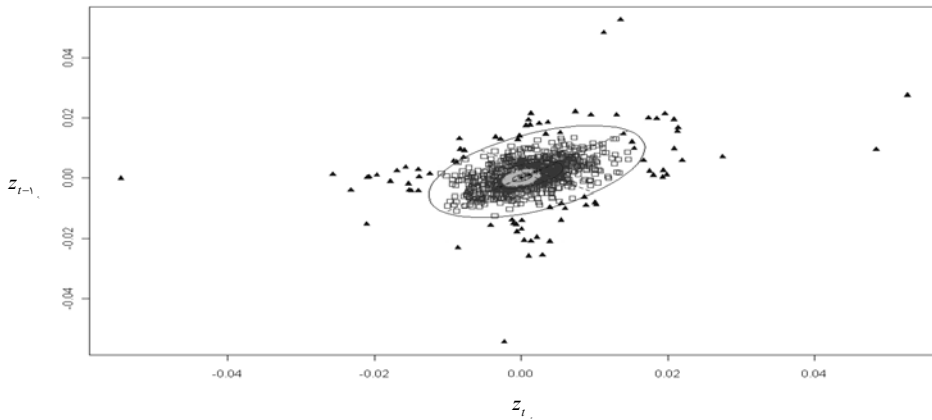
مقادیر BIC ، بهترین مدل برازش شده دارای یک گام وابستگی و سه مؤلفه است. نمودار (۸) نمودار مقادیر BIC به ازای یک ورودی را نشان می‌دهد. نمودار (۹) نیز مؤلفه‌های برازش شده به توزیع توأم Z_t و Z_{t-1} را نشان می‌دهد. با توجه به نمودار (۹)، تفاوت مؤلفه‌ها بیشتر در اندازه‌ی واریانس آنها است. به بیان دیگر، برای مقادیر مختلف واریانس در طول زمان، مدل‌های مختلفی به عنوان خبره در نظر گرفته شده است.

نمودار ۸: نمودار مقادیر BIC به ازای یک ورودی



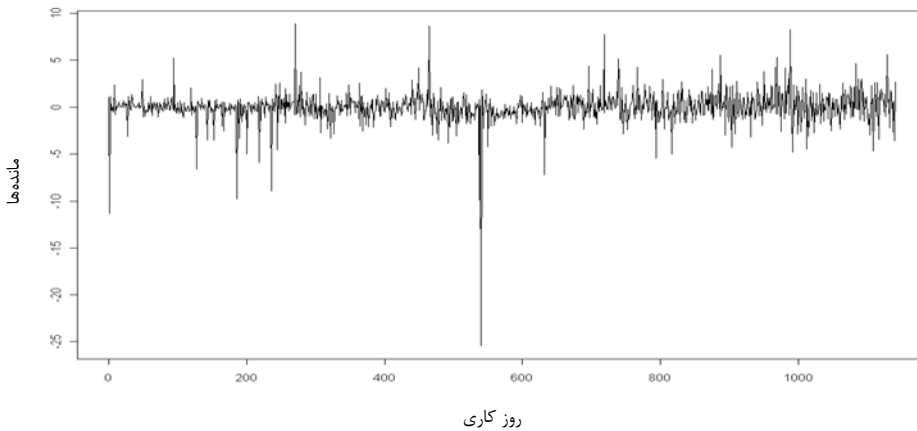
برآورد انحراف معیار پسماندهای مدل (رابطه‌ی (۴))، پس از برازش به سری زمانی Z_t ، برابر با $2/5$ است که با مقدار ۱ فاصله دارد (بر اساس آنچه در مورد رابطه‌ی (۴) گفته شد، واریانس مانده‌های مدل باید برابر با ۱ باشد). بررسی نمودار روند مانده‌های مدل و همچنین نمودارهای همبستگی و همبستگی جزئی پسماندها نشان می‌دهد. (نمودارهای (۱۰) و (۱۱)) مانده‌های مدل برازش شده، دارای خود همبستگی هستند.

نمودار ۹: مؤلفه های برازش شده به توزیع توأم Z_t و Z_{t-1}

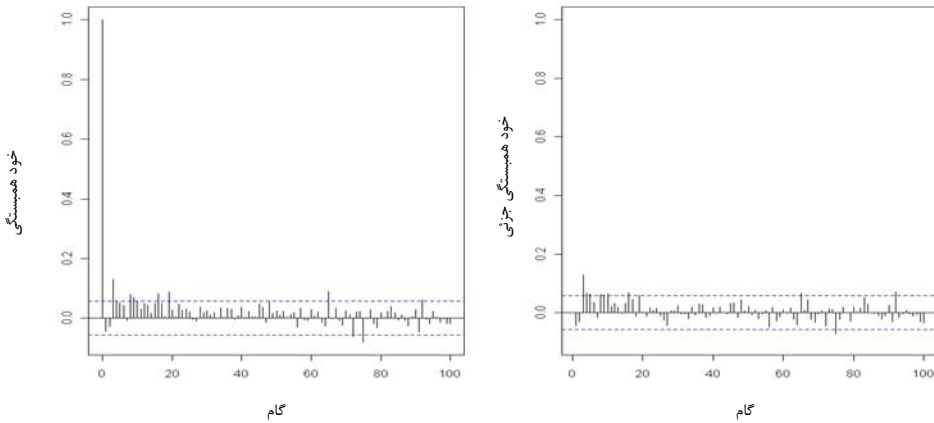


برای حذف خودهمبستگی در پسماندهای مدل، می توان از یک مدل $ARMA$ استفاده کرد. با توجه به تعداد گام های معنی دار در نمودارهای فوق، به نظر می رسد پسماندها از یک مدل $(ARMA\ 2,2)$ پیروی می کنند.

نمودار ۱۰: نمودار روند مانده های مدل

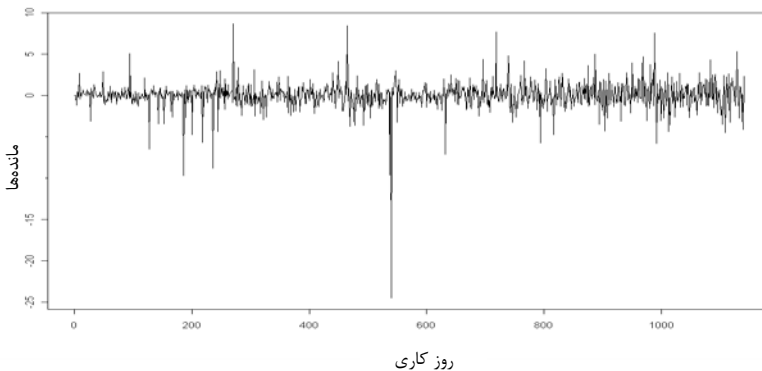


نمودار ۱۱: نمودارهای خود همبستگی و خود همبستگی جزئی مانده‌های مدل

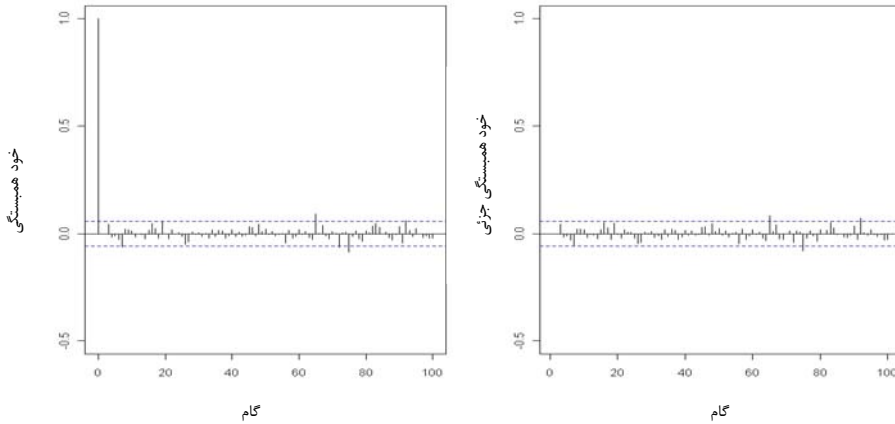


با استفاده از معیار BIC برای مدل $ARMA$ برازش شده نیز می‌توان به نتیجه مناسبی رسید. با برازش یک مدل $(ARMA(2,2))$ ، اثر خود همبستگی از مانده‌ها حذف خواهد شد. نمودار (۱۲) و (۱۳) روند پسماندهای نهایی مدل و نمودارهای خود همبستگی و خود همبستگی جزئی پسماندها را نشان می‌دهد.

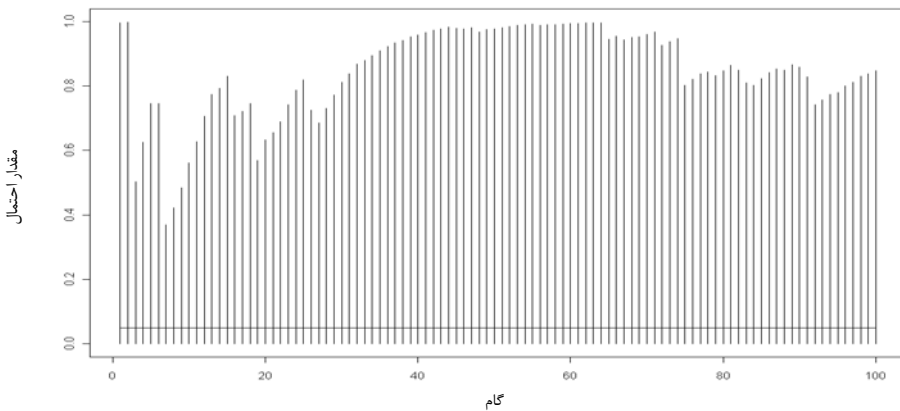
نمودار ۱۲: نمودار روند مانده‌های مدل نهایی



نمودار ۱۳: نمودارهای خود همبستگی و خودهمبستگی جزئی مانده های مدل نهایی



همچنین با استفاده از آماره *Ljung-Box* می توان مشاهده کرد که مانده ها پس از برازش مدل *ARMA* دارای خود همبستگی نیستند. نمودار (۱۴) مقادیر احتمال معنی داری را برای ۱۰۰ گام اول این آماره نشان می دهد که همگی بزرگتر از $0/3$ هستند.

نمودار ۱۴: مقادیر آماره *Ljung-Box* برای پسماندهای مدل نهایی

به این ترتیب، شکل نهایی مدل برای سری زمانی Z_t و برآورد پارامترهای آن به صورت زیر خواهند بود:

$$z_t = \sum_{i=1}^r g^{(i)}(z_{t-1}) \left(\mu_{z_t}^{(i)} + \frac{\sigma_{\gamma}^{(i)}}{\sigma_{\gamma}^{(i)}} (z_{t-1} - \mu_{z_{t-1}}^{(i)}) \right) + \sigma_t (\phi \varepsilon_{t-1} + \phi_{\gamma} \varepsilon_{t-\gamma} + \psi_{\gamma} w_{t-1} + \psi_{\gamma} w_{t-\gamma}) + \sigma_t w_t$$

$$g^{(i)}(z_t) = \frac{\pi_i \sigma_{\gamma}^{(i)-\frac{1}{\gamma}} \exp \left\{ -\frac{1}{\gamma \sigma_{\gamma}^{(i)}} (z_t - \mu_{z_t}^{(i)})^{\gamma} \right\}}{\sum_{j=1}^r \pi_j \sigma_{\gamma}^{(j)-\frac{1}{\gamma}} \exp \left\{ -\frac{1}{\gamma \sigma_{\gamma}^{(j)}} (z_t - \mu_{z_t}^{(j)})^{\gamma} \right\}}$$

$$\sigma_t^{\gamma} = \sum_{i=1}^r g^{(i)}(z_{t-1}) \left(\sigma_{\gamma}^{(i)} - \frac{\sigma_{\gamma}^{(i)}}{\sigma_{\gamma}^{(i)}} \right) + \sum_{i=1}^r g^{(i)}(z_{t-1}) \left[\left(\mu_{z_t}^{(i)} + \frac{\sigma_{\gamma}^{(i)}}{\sigma_{\gamma}^{(i)}} (z_{t-1} - \mu_{z_{t-1}}^{(i)}) \right)^{\gamma} - \left(\sum_{i=1}^r g^{(i)}(z_{t-1}) \left(\mu_{z_t}^{(i)} + \frac{\sigma_{\gamma}^{(i)}}{\sigma_{\gamma}^{(i)}} (z_{t-1} - \mu_{z_{t-1}}^{(i)}) \right) \right)^{\gamma} \right]$$

$$\varepsilon_t = \frac{z_t - \sum_{i=1}^r g^{(i)}(z_{t-1}) \left(\mu_{z_t}^{(i)} + \frac{\sigma_{\gamma}^{(i)}}{\sigma_{\gamma}^{(i)}} (z_{t-1} - \mu_{z_{t-1}}^{(i)}) \right)}{\sigma_t}$$

$$w_t = \frac{z_t - \sum_{i=1}^r g^{(i)}(z_{t-1}) \left(\mu_{z_t}^{(i)} + \frac{\sigma_{\gamma}^{(i)}}{\sigma_{\gamma}^{(i)}} (z_{t-1} - \mu_{z_{t-1}}^{(i)}) \right)}{\sigma_t} - (\phi \varepsilon_{t-1} + \phi_{\gamma} \varepsilon_{t-\gamma} + \psi_{\gamma} w_{t-1} + \psi_{\gamma} w_{t-\gamma})$$

جدول ۱: مقادیر برآورد پارامترهای مدل

پارامتر	برآورد	پارامتر	برآورد	پارامتر	برآورد	پارامتر	برآورد
$\mu_{z_t}^{(1)}$	۰/۰۰۲۱۲۹	$\sigma_{11}^{(1)}$	۰/۰۰۰۲۱۶۷	$\sigma_{22}^{(1)}$	۰/۰۰۰۲۳۲۶	π_{γ}	۰/۲۲۶۸۶۵۱
$\mu_{z_t}^{(2)}$	۰/۰۰۱۰۱۰۴	$\sigma_{11}^{(2)}$	۲/۱۸۴ ۵-۵	$\sigma_{22}^{(2)}$	۲/۱۲۷ ۵-۵	ϕ_1	۰/۸۱۵۵
$\mu_{z_t}^{(3)}$	۰/۰۰۰۱۰۶۸	$\sigma_{11}^{(3)}$	۱/۶۸۱ ۵-۶	$\sigma_{22}^{(3)}$	۱/۷۶۰ ۵-۶	ϕ_2	۰/۰۱۳۱
$\mu_{z_{t-1}}^{(1)}$	۰/۰۰۲۱۵۳	$\sigma_{12}^{(1)}$	۰/۰۰۰۱۱۴۶	π_1	۰/۱۰۰۸۷۹۵	ψ_1	-۰/۸۷۸۹
$\mu_{z_{t-1}}^{(2)}$	۰/۰۰۰۹۷۷۳	$\sigma_{12}^{(2)}$	۱/۱۰۲ ۵-۵	π_2	۰/۶۷۲۲۵۵۴	ψ_2	-۰/۰۱۱۴
$\mu_{z_{t-1}}^{(3)}$	۰/۰۰۰۰۵۵۰	$\sigma_{12}^{(3)}$	۸/۷۹۴ ۵-۷				

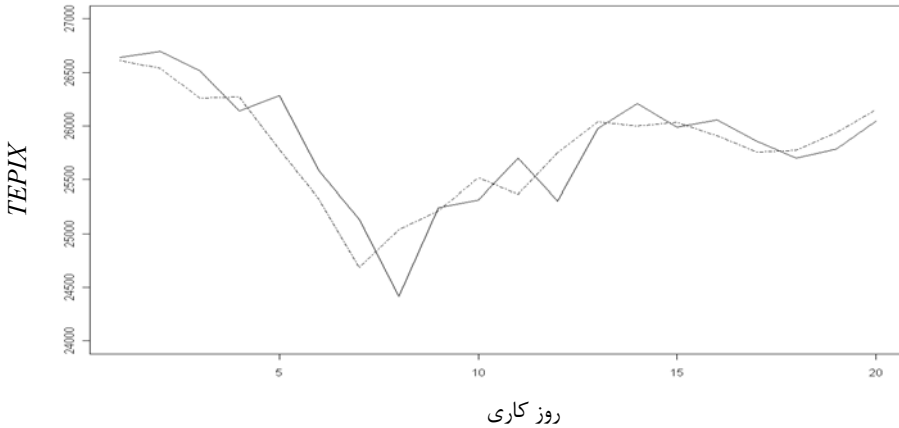
پیش‌بینی یک گام به جلو برای مشاهدات بعد از تاریخ ۹۰/۱/۳۱، نشان می‌دهد که مدل برازش شده کارایی مناسبی در پیش‌بینی سری زمانی z_t دارد. جدول (۲) مقادیر مشاهده شده و مقادیر پیش‌بینی را برای بیست روز کاری (۹۰/۲/۳۱ تا ۹۰/۲/۱) به همراه ریشه‌ی میانگین مربع خطا نشان می‌دهد. همچنین نمودار (۱۵) مقادیر فوق را نمایش می‌دهد.

جدول ۲: مقادیر پیش‌بینی یک گام به جلو و مقادیر مشاهده شده از تاریخ ۹۰/۲/۱ تا ۹۰/۲/۳۱

تاریخ	پیش‌بینی	مشاهده	تاریخ	پیش‌بینی	مشاهده
۹۰/۲/۳	۲۶۶۳۹/۴۶	۲۶۶۱۰/۱	۹۰/۲/۱۸	۲۵۷۰۰/۸۰	۲۵۳۶۶/۶
۹۰/۲/۴	۲۶۶۹۴/۴۶	۲۶۵۳۹/۷	۹۰/۲/۱۹	۲۵۳۰۴/۴۴	۲۵۷۵۲/۷
۹۰/۲/۵	۲۶۵۱۵/۹۸	۲۶۲۶۱/۵	۹۰/۲/۲۰	۲۵۹۷۸/۳۹	۲۶۰۴۱/۷
۹۰/۲/۶	۲۶۱۴۱/۷۰	۲۶۲۷۳/۶	۹۰/۲/۲۱	۲۶۲۱۱/۳۷	۲۶۰۰۲/۸
۹۰/۲/۷	۲۶۲۸۶/۱۴	۲۵۷۸۱/۸	۹۰/۲/۲۴	۲۵۹۹۰/۷۵	۲۶۰۳۳/۶
۹۰/۲/۱۰	۲۵۵۸۰/۹۲	۲۵۳۲۱/۷	۹۰/۲/۲۵	۲۶۰۵۶/۲۸	۲۵۹۰۸/۵
۹۰/۲/۱۱	۲۵۱۳۵/۱۵	۲۴۶۸۵/۳	۹۰/۲/۲۶	۲۵۸۶۱/۲۱	۲۵۷۵۹/۷
۹۰/۲/۱۲	۲۴۴۱۴/۴۱	۲۵۰۳۷/۷	۹۰/۲/۲۷	۲۵۷۰۰/۵۱	۲۵۷۷۳/۴
۹۰/۲/۱۳	۲۵۲۴۳/۶۳	۲۵۲۱۱/۲	۹۰/۲/۲۸	۲۵۷۸۶/۵۵	۲۵۹۴۰/۸
۹۰/۲/۱۴	۲۵۳۱۷/۸۵	۲۵۵۲۰/۲	۹۰/۲/۳۱	۲۶۰۴۵/۱۰	۲۶۱۵۵/۶
RMSE ¹⁴	۲۷۳/۵۶۹۸۰۱۹				

¹⁴ Root Mean Square Error

نمودار ۱۵: نمودار پیش‌بینی یک گام به جلو (خط هموار)، مقادیر مشاهده شده (خط چین)



۵- بحث و نتیجه‌گیری

در این مقاله بر مبنای مدل‌های ترکیب خبرگان، ساختار مدل $GL^{\lambda}GNN$ معرفی شد و با استفاده از آن، مدلی برای پیش‌بینی شاخص بورس تهران ارائه گردید. نتایج نشان داد که با وجود شکل نامتعارف توزیع داده‌ها، مدل ارائه شده از کارایی مناسبی در مدل‌سازی و پیش‌بینی شاخص بورس تهران برخوردار است. مدل ارائه شده علاوه بر اینکه رفتار شاخص بورس تهران را به خوبی مدل‌سازی می‌کند، در پیش‌بینی یک گام به جلو نیز روند شاخص را به خوبی دنبال می‌کند.

با توجه به اینکه در مسائل اقتصادی و مالی، داده‌های مورد بررسی عموماً دارای توزیعی شبیه به توزیع‌های آماری کلاسیک ندارند، استفاده از مدل‌های ترکیب خبرگان شبیه $GL^{\lambda}GNN$ ، که دارای ساختاری نیمه پارامتری هستند، می‌تواند سودمند باشد. از این‌رو، استفاده از این مدل‌ها در مسائل مختلف اقتصادی از یک سو می‌تواند دقت مدل‌سازی و پیش‌بینی را افزایش دهد و از سوی دیگر مبنای نظری برای استنباط در مورد مدل برازش شده را در اختیار قرار دهد.

فهرست منابع

- سینایی، حسنعلی، سعید ا... مرتضوی و یاسر تیموری اصل. (۱۳۸۴). پیش بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه های عصبی مصنوعی. بررسی های حسابداری و حسابرسی، ۱۲(۴۱): ۵۹-۸۳.
- عرفانی، علیرضا. (۱۳۸۷). بررسی حافظه بلند بودن شاخص کل قیمت بورس اوراق بهادار تهران. پژوهش نامه علوم انسانی و اجتماعی، ۸(۲۸): ۷۷-۹۲.
- فلاح شمس، میر فیض و بیبا دلنواز اصغری. (۱۳۸۸). پیش بینی شاخص بورس اوراق بهادار تهران با استفاده از شبکه های عصبی. فراسوی مدیریت، ۳(۹): ۲۱۲-۱۹۱.
- کشاوری حداد، غلامرضا و باقر صمدی. (۱۳۸۸). برآورد و پیش بینی تلاطم بازدهی در بازار سهام تهران و مقایسه دقت روش ها در تخمی ارزش در معرض خطر: کاربردی از مدل های خانواده FIGARCH. مجله تحقیقات اقتصادی، ۸۶: ۲۳۵-۱۹۳.
- منجمی، امیر حسین، مهدی ابزری و علیرضا رعیتی شوازی. (۱۳۸۸). پیش بینی قیمت سهام در بازار بورس اوراق بهادار. اقتصاد مقداری، ۶(۳): ۲۶-۱.
- یگانگی، محمدرضا. (۱۳۸۸). مدل سازی و پیش بینی سری های زمانی غیر خطی با استفاده از شبکه های عصبی. پایان نامه کارشناسی ارشد آمار اقتصادی- اجتماعی. دانشگاه شهید چمران اهواز.
- Farinas, M.S., C.E. Pedreira & M.C. Medeiros. (2004). Local-Global Neural Networks: A New Approach for Nonlinear Time Series Modeling. *Journal of American Statistical Association*, 99(468):1092-1107.
- Hampshir J. & A. Waibel. (1989). The Meta-Pi Network: Building distributed Knowledge Representations for Robust Pattern Recognition. Technical Report CMU-CS-89-166. Carnegie-Mellon, Pittsburgh, PA, U.S.
- Jacobs, A.R., M.I. Jordan, S.J. Nowlan & G.E. Hinton. (1991). Adaptive Mixture of Local Experts. *Neural Computation*, 3: 79-87.
- Jordan, M.I. & R.A. Jacobs. (1992). Hierarchies of Adaptive Experts. *Advances in Neural Information Processing System 4*. Eds. J.E. Moody, & R.P. Lippmann. San Mateo: Morgan Kaufmann publications, 985- 992.
- Krollner, B., B. Vanstone & G. Finnie. (2010). Financial Time Series Forecasting with Machine Learning Techniques: A Survey. Paper presented at the European Symposium on Artificial Neural Networks: Computational and Machine Learning. Bruges, Belgium. Apr. 2010.
- Palma, W. (2007). Long-Memory Time Series. John Wiley, Hoboken, New Jersey.

- Pearson, K. (1894). Contributions to the Theory of Mathematical Evolution. Philosophical Transaction of the Royal Society of London, 186: 343-411.
- Priestly, M.B. (1980). State Dependent Models: A General Approach to Nonlinear Time Series Analysis. Journal of Time Series Annual, 1: 47-71.
- Priestly, M.B. (1988). Nonlinear Non-Stationary Time Series Analysis. Academic Press, London.
- Wong, C.S. & W.K. Li. (2000). On a Mixture Autoregressive Model. Journal of the Royal Statistical Society, 62: 95-115.