

مدل سازی و پیش بینی رشد اقتصادی با تحلیل سری های زمانی غیر خطی

فریده سبحانی فرد^۱، محمدرضا شهرکی^{۲*}، محمد انیسه^۳

۱. دانشجوی کارشناسی ارشد مهندسی صنایع دانشگاه سیستان و بلوچستان

۲. استادیار گروه مهندسی صنایع دانشگاه سیستان و بلوچستان

۳. استادیار گروه مدیریت صنعتی دانشگاه بین المللی امام خمینی

(تاریخ دریافت: ۱۰/۰۲/۹۶، تاریخ دریافت روایت اصلاح شده: ۰۷/۰۳/۹۸، تاریخ تصویب ۲۹/۰۳/۹۸)

چکیده

پیش بینی صحیح رشد اقتصادی در سیاست گذاری ها و برنامه ریزی های بلندمدت توسعه پایدار نقشی مهم را ایفا می کند. یکی از مسائل مهم در پیش بینی سری های زمانی استفاده از روش هایی برای شناسایی الگوهای زمانی با هدف کنترل پیچیدگی ها و بهینه سازی خطای حاصل از پیش بینی است. در این پژوهش، تحلیل سری های زمانی تولید ناخالص داخلی به صورت غیر خطی به منظور پیش بینی مسیر حرکت رشد اقتصادی به کمک روش شبکه عصبی مصنوعی بیزی، برای انعطاف بیشتر مدل غیر خطی در برخورد با پیچیدگی های مسئله و انطباق بیشتر با شرایط واقعی صورت گرفت. در ادامه با استفاده از ترکیب الگوریتم فراابتکاری ژنتیک در آموزش شبکه، به بهبود کارایی مدل در مقایسه با نتایج روش های قدیمی تر پرداخته شد. در تخمین مدل از داده های دوره ۱۳۷۱ تا ۱۳۹۲ استفاده شد. سپس بررسی کارایی آن برای داده های فصلی ۱۳۹۳ تا دو فصل اول ۱۳۹۵ با استفاده از معیارهای میانگین مربعات خطا و خطای استاندارد میانگین صورت گرفت. براساس نتایج، اصلاح پیچیدگی ها در آموزش شبکه نقش بسزایی در بهینه سازی خطای مدل خواهد داشت.

واژه های کلیدی: پیش بینی، سری زمانی، شبکه عصبی بیزی، شبیه سازی زنجیره مارکوف مونت کارلو، مدل سازی.

مقدمه

در رشد اقتصادی یک کشور دخیل خواهند بود [۴]. پیش بینی براساس مدل های چندمتغیری اقتصادسنجی با محدودیت های بسیاری همراه است؛ بنابراین روش جایگزین، استفاده از مدل های پیش بینی ریاضی است؛ از این رو در بخش دوم این مطالعه، ادبیات و پیشینه پژوهش، و بر همین اساس در بخش سوم مدلی برای فرایند پیش بینی های اقتصادی ارائه شد. سپس در بخش بعد تجزیه و تحلیل کمی اطلاعات با توجه به روش مورد نظر صورت گرفت و نتایج پژوهش در مقایسه با نتایج دیگر پژوهش ها بررسی شد.

پیشینه پژوهش

یکی از کاربردهای مدل سازی سیستم های واقعی در پیش بینی، وضعیت و رفتار آن سیستم در آینده است. پیش بینی بیانیه ای در مورد رویدادی نامشخص و احتمالی است که بیشتر براساس تجربه یا دانش صورت می گیرد. در بسیاری از کاربردها مانند آنالیز سری های زمانی، ممکن است مدل ها به کمک مشاهدات مورد بررسی تخمین زده

داده های اقتصادی پیش شرط لازم برای مدیریت مؤثر اقتصاد کلان هستند. پیچیدگی های اقتصاد مدرن یک کشور باید از ظرفیت شناسایی سریع و روند مطلوب در اقتصاد و اقدامات اصلاحی مناسب برخوردار باشد [۱]. این مهم بدون در اختیار داشتن داده های اقتصادی کامل و دقیق امکان پذیر نخواهد بود. رشد اقتصادی پدیده ای کمی، تغییرات در میزان تولیدات و یا به عبارت دیگر تغییرات در تولید خالص سالانه ملی و در نتیجه درآمد ملی است [۲]. در واقع رشد اقتصادی، بر افزایش تولید یا درآمد سرانه ملی دلالت دارد که با افزایش تولید کالا یا خدمات در یک کشور رخ می دهد [۳]. برخی اقتصاددانان از جمله میل، برو، نورث، لویس و واگنر، عوامل اجتماعی را در بررسی رشد اقتصادی کشورها سهیم می دانند. به نظر آن ها کشوری که برای فعالیت ها، تخصص ها و مهارت افراد سرمایه گذاری می کند، در تجارت بین المللی و رقابت جهانی جایگاه ویژه ای خواهد داشت و از نهادهای اقتصادی پایداری برخوردار خواهد بود؛ بنابراین عوامل تعیین کننده بسیاری

تولید ناخالص داخلی و نرخ تورم، نرخ بهره اوراق قرضه شش ماهه و نرخ بیکاری را برای اقتصاد کلان آمریکا پیش بینی کرد. همچنین از داده های فصلی پیش بینی تولید ناخالص داخلی برای دوره ۱۹۹۱ تا ۲۰۰۹ و از داده های ماهانه برای متغیرهای ۱۹۵۰ تا ۲۰۰۹ استفاده کرد. در واقع او به مقایسه روش ANFIS با روش ARIMA خطی و ستاره غیرخطی پرداخت و نتیجه گرفت نتایج روش او از نظر RMSE و میانگین خطای مطلق، و نتایج روش ARIMA بهتر از STAR است [۱۴].

Tealab مدل جدیدی از شبکه عصبی هیبرید را به کمک منطق فازی ارائه داد و نتیجه گرفت شبکه عصبی معمولی برای درک رفتار غیرخطی سری های زمانی مناسب نیست. معمولاً تعداد زیادی از فرایندهای آماری در یک فاصله زمانی تعریف می شوند. این فرایندها با ارزش های گذشته خود بر مقادیر مربوط به زمان های اخیر تأثیرگذارند و ساختاری بر پایه روش های سری زمانی خواهند داشت [۱۵]. رویکردی را که Box و همکاران [۱۶] ارائه دادند، مرجعی کلاسیک در تکنیک های سری زمانی برای مدل کردن ساختار عملکردی است. در این میان، روش های دیگری از جمله روش حداقل مربعات معمولی (OLS) برای برآورد پارامترها وجود دارد [۱۷]. در برخی موارد، به منظور تطبیق بهتر مدل با شرایط واقعی به جای مدل های خطی سری زمانی می توان از مدل های غیرخطی مانند Bilinear، Threshold و Autoregressive (AR) استفاده کرد [۱۸]؛ البته با وجود عملکرد خوب این مدل ها، در آن ها محدودیت های ذاتی نیز وجود دارد [۱۹]. اول از همه اینکه تشخیص مدل مناسب به متخصص نیاز دارد و ممکن است یک ساختار تابع غیرمناسب ایجاد شود. دوم اینکه برخی رفتارهای غیرخطی ممکن است به دلیل ساختار عملکردی مدل سازی نشوند. به منظور غلبه بر این محدودیت ها، از شبکه عصبی مصنوعی برای مدل های سری زمانی غیرخطی استفاده می شود. وقتی شبکه عصبی مصنوعی آموزش می بیند، میانگین مربعات ریسک تجربی را به عنوان تابع خطا برای کنترل پیچیدگی مدل به کار می برد و از برآورد و پیش بینی مدل اطمینان می دهد. به همین ترتیب به کار بردن الگوریتم بیزی (BNNs) در استفاده از رگرسیون سری های زمانی غیرخطی بسیار منعطف عمل می کند و

شوند. همچنین می توان مدل ها را براساس توابع انتقالی فضای پارامتری نیز تعریف کرد که با توجه به آن ها برآورد پیش بینی محاسبه خواهد شد [۵]. پیش بینی رشد اقتصادی با استفاده از مدل سازی ها و روش های مختلفی انجام می شود که به آن ها اشاره خواهد شد.

Stock و Watson نقش متغیرهای مالی و برخی سری های زمانی اقتصاد کلان را در پیش بینی نرخ تولید ناخالص داخلی و تورم با استفاده از داده های ۱۹۷۱ تا ۱۹۸۴ و ۱۹۸۵ تا ۱۹۹۹ در هفت کشور OECD بررسی کردند [۶]. Baffigi و همکاران نیز به بررسی نرخ رشد تولید ناخالص داخلی کشورهای فرانسه، ایتالیا و آلمان با استفاده از مدل بریج (BM) و مدل های ARIMA و VAR پرداختند. داده های مورد استفاده مربوط به سال های ۱۹۸۰-۱۹۹۰ بود. براساس نتایج این پژوهش، مدل BM در شرایطی خاص بهتر عمل می کند [۷].

Schumacher تولید ناخالص داخلی کشور آلمان را به کمک داده های فصلی برای ۱۲۴ فصل از سال ۱۹۷۸ تا ۲۰۰۴ با معیار میانگین مربعات خطا (MSE) بررسی کرد و نتیجه گرفت پیش بینی خارج نمونه یک گام و چهار گام سه مدل عاملی از مدل الگوی اتورگرسیو بهتر است. وی کار خود را به کمک داده های ماهانه و فصلی با استفاده از الگوریتم EM و برآوردگر مؤلفه اصلی بهبود داد [۸]. در پژوهش Schumacher و Breitung پیش بینی تولید ناخالص داخلی آلمان به دقت بهتری رسید [۹]. محققین دیگری نیز برای پیش بینی تولید ناخالص داخلی فصلی اروپا از شاخص های ماهانه در DFM استفاده کردند [۱۰، ۱۱]. Qi و Zhang سری زمانی ماهانه اقتصادی آمریکا را به کمک شبکه های عصبی مصنوعی و مدل ARIMA پیش بینی کردند. آن ها بر حذف روند و الگوهای فصلی تمرکز کردند، اما نتیجه گرفتند شبکه های عصبی با حذف این گونه الگوها قادر به کاهش خطای پیش بینی نیستند [۱۲].

Pouzols و همکاران از سیستم های استنتاج فازی و تخمین ناپارامتری واریانس باقی مانده ها برای پیش بینی کوتاه مدت و بلندمدت سری های زمانی استفاده کردند و نتیجه گرفتند مدل آن ها از مدل پایه اتورگرسیون ها عملکرد بهتری دارد [۱۳]. Giovanis با استفاده از روش سیستم استنتاج فازی عصبی سازگار (ANFIS)، پیش بینی

براین اساس داده‌ها بر طبق متغیر تولید ناخالص داخلی در سری زمانی معینی جمع‌آوری شدند. پس از آن برای چک کردن نرمال بودن سری‌ها، از آزمون کلموگروف-اسمیرنوف (K-S) در نرم‌افزار SPSS نسخه ۲۲ استفاده شد. در این پژوهش، تصادفی کردن داده‌ها صورت گرفت تا استفاده از شبکه‌های عصبی مصنوعی به دلیل تأثیر عوامل مختلف در جهت خاصی نباشد و نتیجه حاصل از تحلیل‌ها برای سایر داده‌ها نیز قابل استفاده باشد. پس از برچسب‌گذاری ستون‌ها و ردیف‌ها نیز ستون‌های مربوط به مقادیر ورودی و خروجی مشخص شد. همچنین این مسئله که چند ردیف به آموزش، چند ردیف به اعتبارسنجی داده‌ها و چند ردیف به آزمون شبکه اختصاص یابد، بررسی شد. مقادیر پیشنهادی نرم‌افزار برای تخصیص ردیف به آموزش، آزمون شبکه و اعتبارسنجی داده‌ها به ترتیب ۶۰، ۲۵ و ۱۵ درصد داده‌هاست. در نهایت، نمونه شاخص و تنظیمات ساختار شبکه تولید شد. در این مطالعه، تانژانت هیپربولیک به عنوان تابع انتقال به دلیل قابلیت تشخیص بهتر با کمترین خطای ممکن و محاسبه سریع‌تر از تابع سیگموئید انتخاب شد.

در ادامه، داده‌های سری زمانی به منظور آموزش شبکه عصبی مصنوعی پیش‌خور به سه بخش آموزش، اعتبارسنجی و آزمون تقسیم شدند. در روش k -fold اعتبارسنجی داده‌ها به K زیرمجموعه افراز می‌شوند. از این K ، هر بار یکی برای اعتبارسنجی و $K-1$ تای دیگر برای آموزش استفاده می‌شوند. این روال K بار تکرار می‌شود و همه داده‌ها دقیقاً یک‌بار برای آموزش و یک‌بار برای اعتبارسنجی به کار می‌روند. در نهایت میانگین نتیجه این K بار اعتبارسنجی به عنوان تخمین نهایی برگزیده می‌شود. در روش K -Fold طبقه‌ای، نسبت داده‌های هر کلاس در هر زیرمجموعه و در مجموعه اصلی یکسان است. در این آزمون، بهترین خطای جذر میانگین مربعات برابر $0/027$ و از تکنیک 10-fold به دست آمده است. در نهایت به منظور اطمینان از تعمیم روش، از تکنیک‌های اعتبارسنجی One-Leave-Out و Cross-Validation استفاده شد. چرخه Cross-Validation تا تعداد مشاهدات باقی‌مانده در سری زمانی آموزش و رسیدن به داده‌های آزمون تکرار شد و در اعتبارسنجی داده‌ها، خطاها در طول فرایند آموزش تحت

استفاده از آن در شبکه‌های عصبی مصنوعی بر پایه تقریب‌های گاوسی صورت می‌گیرد.

شبیه‌سازی زنجیره مارکوف مونت کارلو (MCMC) نیز یکی از روش‌های بیزی است که البته استفاده از آن ممکن است هزینه زیادی داشته باشد یا به همگرایی نرسد؛ بنابراین به منظور رفع کاستی‌های موجود الگوریتم بیزین از روش‌های گوناگونی استفاده می‌شود. برخی از محققین کاربرد شبکه‌های بیزی را برای آموزش شبکه عصبی مصنوعی بازبینی کردند [۲۰،۲۱،۲۲]. همچنین مطالعاتی با تمرکز بر مسائل شبکه عصبی مصنوعی و رویکرد بیزی انجام شده است که در آنها به بررسی انتخاب تعداد لایه‌های پنهان برای مدل‌سازی بهتر شبکه عصبی مصنوعی پرداخته شده است [۲۱،۲۳]. Freitas نیز کاربردی از مونت کارلو را در شبکه عصبی مصنوعی بیزی ارائه داد [۲۴].

به منظور اندازه‌گیری پیچیدگی مدل، تخمین پارامترها و هایپر پارامترها در شبکه‌های عصبی مصنوعی بیزی، روش‌های ترکیبی به تکنیک‌های کلاسیک سری‌های زمانی ترجیح داده می‌شوند. همچنین آموزش شبکه عصبی مصنوعی بیزی و تکنیک زنجیره مارکوف مونت کارلو با الگوریتم بهینه‌سازی گرادیان تلفیق شده است. استفاده از تکنیک زنجیره مارکوف مونت کارلو در کاهش زمان آموزش بطور وسیعی مورد استفاده قرار گرفته است. همچنین مطالعات متعددی در زمینه شبکه‌های عصبی، مسائل غیرخطی و جزئیات الگوریتم‌های کاربردی در آن انجام شده است [۱۹،۲۱،۲۲]. در این میان، استفاده از الگوریتم‌های فراابتکاری به منظور کاهش زمان آموزش برای تخمین پارامترها در شبکه‌های عصبی مصنوعی بسیار مورد توجه است.

روش پژوهش

شبکه‌های عصبی ابزاری مفید برای حل انواع مسائل از جمله بهینه‌سازی‌های محدودیت‌دار هستند. با ترکیب این شبکه‌ها و بیزی می‌توان از آنها در تجزیه و تحلیل سری‌های زمانی استفاده کرد در این پژوهش، مدل‌سازی و پیش‌بینی متغیر رشد اقتصادی براساس روش شبکه‌های عصبی مصنوعی با الگوریتم آموزش بیزین صورت گرفت.

گاوسی داده‌ها به کمک رابطه ۴ مدل‌سازی می‌شوند. در رابطه ۵ نیز فرایندهای نویز \mathcal{E} مشاهده می‌شود.

$$f_s(x_i, \theta) = b_i^H + \sum_{k=1}^m w_{ik}^H A_k (w_k^I x_i + b_k^I) \quad s=1, 2, \dots, r \quad (۴)$$

$$y_{ii} = f_i(x_i, \theta) + \mathcal{E} \quad (۵)$$

که در آن $x_i = [x_{i1}, x_{i2}, \dots, x_{i3}]$ و $y_i = [y_{i1}, y_{i2}, \dots, y_{ir}]$ به ترتیب لایه‌های ورودی و هدف مشاهده i ام هستند. $w_k^I = [w_{1k}^I, w_{2k}^I, \dots, w_{pk}^I]$ یک لایه سطری است که شامل وزن همه ورودی‌ها و k امین نورون در لایه پنهان است. $w_i^H = [w_{1i}^H, w_{2i}^H, \dots, w_{ms}^H]$ لایه سطحی شامل وزن همه نورون‌ها و s امین خروجی است. b_k^I و b_k^H بایاس‌ها به ترتیب برای k امین نورون و s امین خروجی هستند و تمام $\theta = [w_k^I, w_k^H, b_k^I, b_k^H]$ لایه‌ای پارامتری است که تمام وزن‌ها و بایاس‌ها را شامل می‌شود. $i=1, 2, \dots, N$; $j=1, 2, \dots, p$; $k=1, 2, \dots, m$; $s=1, 2, \dots, r$. اگر \mathcal{E} یک نویز گاوسی با میانگین صفر و انحراف استاندارد σ_{noise} داشته باشد. احتمال مقادیر داده برای پارامتر لایه θ به صورت رابطه ۶ است.

$$p(y_i | x_i, \theta, \beta) = \frac{1}{(2\pi)^{1/2} \beta^{-1/2}} \exp\left(-\frac{\beta}{2} \sum_{i=1}^N \{y_{ii} - f_i(x_i, \theta)\}^2\right) \quad (۶)$$

که در آن $\beta = \frac{1}{\sigma_{noise}^2}$ دقت و کنترل نویز واریانس نامیده می‌شود. در اینجا فرض این است که نقاط داده‌ها به‌طور مستقل از هم نشان داده می‌شوند. این توزیع می‌تواند شامل N مشاهده به‌صورت جداگانه براساس رابطه ۷ باشد که تابع احتمال آن به‌صورت رابطه ۸ نوشته می‌شود.

$$p(D | \theta, \beta) = \prod_{i=1}^N p(y_i | \theta, x_i) = \frac{1}{(2\pi)^{N/2} \beta^{-N/2}} \exp\left(-\frac{\beta}{2} \sum_{i=1}^N \{y_{ii} - f_i(x_i, \theta)\}^2\right) \quad (۷)$$

$$p(D | \theta, \beta) = \frac{1}{Z_D(\beta)} \exp(-\beta E_D) \quad (۸)$$

تابع $Z_D(\beta)$ نرمالیزه شده نام دارد و به‌صورت رابطه ۹ یکپارچه‌سازی می‌شود.

$$Z_D(\beta) = \int \exp(-\beta E_D) dD = (2\pi)^{N/2} \beta^{-N/2} \quad (۹)$$

در چارچوب بیزی، توزیع احتمال بیش از مقادیر پارامترها نشان‌دهنده دانش قبلی در مورد شبکه است؛ بنابراین توزیع پارامترها به‌صورت نمایی براساس رابطه ۱۰ تعریف می‌شود.

نظر قرار گرفت. در این چرخه، خطاها در هر اعتبارسنجی داده‌ها برای احتمال توقف زود هنگام نمایش داده شدند. وقتی خطای اعتبارسنجی افزایش می‌یابد، برای پنج شکست اعتبارسنجی ماکزیمم در هر گام آموزش در چرخه Cross-Validation آموزش به‌صورت اتوماتیک به‌وسیله الگوریتم پایان می‌یابد و سپس به گام بعدی می‌رود. علاوه بر شکست اعتبارسنجی، معیارهای دیگر توقف، رسیدن به حداکثر تکرار یعنی ۵۰۰۰ و تحمل خطا بین $(1 \cdot 10^{-10})$ تکرار موفق و بزرگی شیب $(1 \cdot 10^{-5})$ است.

تقریب گاوسی برای آموزش بیزی

در چارچوب آموزش بیزی، از دانش اولیه پارامترها به‌صورت احتمالی استفاده می‌شود؛ به‌طوری‌که سنجش برازش داده‌های یک مدل پاسخ‌گو به هر گروه از پارامترها، به کمک احتمالات به‌خوبی سنجیده می‌شود. نتیجه یادگیری بیزی، توزیع احتمال است که با در نظر گرفتن پارامترهای مدل، نظر ما را در مورد اینکه ارزش پارامترهای مختلف چگونه است بیان می‌کند. برای شروع فرایند یادگیری باید توزیع $p(\theta)$ برای پارامترها تعریف شود. با مشاهده داده $D = \{x_{i:N}, y_{i:N}\}$ توزیع قدیمی‌تر به کمک توزیع جدیدتر و با قوانین بیزی براساس رابطه ۱ به‌روز می‌شود.

$$p(\theta | D) = \frac{P(D | \theta) p(\theta)}{p(D)} = \frac{P(D | \theta) P(\theta)}{\int P(D | \theta) P(\theta) d\theta} \quad (۱)$$

به کمک توزیع پیشین، احتمال پیش‌بینی داده‌ها به‌صورت رابطه ۲ و مقادیر پیش‌بینی شده مقدار مدنظر از رابطه ۳ به‌دست می‌آید. در این مرحله، باید پیش‌بینی تمام مقادیر ممکن از شبکه‌های مبتنی بر پارامترها به کمک احتمال حضور آن‌ها در آموزش داده‌ها وزن داده شود.

$$P(y_{i:N+1} | x_{i:N+1}, y_{i:N}) = \int P(y_{i:N+1} | \theta, x_{i:N+1}) p(\theta | x_{i:N}, y_{i:N}) d\theta \quad (۲)$$

$$E(y_{i:N+1} | x_{i:N+1}, y_{i:N}) = \int \hat{f}(y_{i:N+1} | \theta, x_{i:N+1}) p(\theta | x_{i:N}, y_{i:N}) d\theta \quad (۳)$$

براساس تقریب گاوسی، عدم قطعیت در فضای پارامترها به توزیع احتمال با در نظر گرفتن درجه باور به مقادیر مختلف لایه‌های پارامتر تعیین می‌شود. با این تعریف، به کمک بیشینه کردن مقدار توزیع پیشین پارامترها، محتمل‌ترین پارامترها می‌توانند تعیین شوند. در تقریب

دوره اول بیشتر و بیشتر غالب می‌شود تا اینکه تأثیر دوره دوم به حداقل می‌رسد؛ در نتیجه بیشینه راه‌حل ماکسیمم برای انتخاب محتمل‌ترین HMP برای N بزرگ به تقریب خوبی می‌رسد. برعکس، برای مجموعه داده‌های کوچک، دوره‌های پیشین سختی بیشتری را در تشخیص موقعیت محتمل‌ترین راه‌حل دارند. در عمل برای سنجش توزیع احتمال شبکه پیش‌بینی به یکپارچه‌سازی فضای پارامتری نیاز است؛ بنابراین این انتگرال باید رام شود، اما نرمال‌سازی فاکتور $Z_S(\alpha, \beta)$ به صورت تحلیلی سنجیده نشده است [20]. با استفاده از تقریب گاوسی برای احتمال پیشین به کمک قوانین احتمال، توزیع خروجی برای تخصیص ورودی به لایه جدید به کمک رابطه ۱۷ صورت می‌گیرد.

$$p(y_{t,N+1} \setminus x_{N+1}^{new}, D) = \int p(y_{t,N+1} \setminus x_{N+1}, \theta) p(\theta \setminus D) d\theta \quad (17)$$

که $p(\theta \setminus D)$ توزیع خلفی از پارامترها و $\{x_{t,N}, y_{t,N}\}$ مجموعه داده‌هاست.

$p(y_{t,N+1} \setminus x_{N+1}^{new}, \theta)$ به منظور توزیع نویز بر داده‌های هدف برای مقادیر ثابت شده از لایه پارامتری، به سادگی مدل‌سازی می‌شود. انتگرال رابطه ۱۷ می‌تواند با تقریب گاوسی به سادگی برآورد شود [۲۲، ۲۵، ۲۶] که در نهایت رابطه ۱۸ را خواهیم داشت:

$$p(y_{t,N+1} \setminus x_{N+1}^{new}, D) = \frac{1}{(2\pi\sigma_{t,N+1}^2)^{1/2}} \exp\left(-\frac{(y_{t,N+1} - f_t(x_{N+1}, \hat{\theta}_{MP}))^2}{2\sigma_{t,N+1}^2}\right) \quad (18)$$

که در آن نرمالیزه‌سازی فاکتور به صورت واضح بازسازی شد؛ در نتیجه این توزیع دارای میانگین، به صورت رابطه ۱۹ و واریانس آن به صورت رابطه ۲۰ است.

$$\bar{y}_{t,N+1} = \hat{f}_t(x_{N+1}, \hat{\theta}_{MP}) \quad (19)$$

$$\sigma_{t,N+1}^2 = \frac{1}{\beta} + g'A^{-1}g \quad (20)$$

در اینجا انحراف استاندارد σ از توزیع پیش‌بینی برای $y_{t,N+1}$ می‌تواند نوار خطا در مقدار میانگین $\bar{y}_{t,N+1}$ باشد. برای این خطا دو دلیل مطرح می‌شود؛ نخست نویز ذاتی بر داده‌های هدف مربوط به دوره اول رابطه ۲۰ و دوم عرض توزیع پیشین پارامترهای شبکه براساس دومین دوره در رابطه ۲۰. تاکنون فرض می‌شد که پارامترها می‌توانند به‌طور مستقیم از تابع خط با به‌کارگیری هایپر پارامترهای

$$p(\theta) = \frac{1}{Z_\theta(\alpha)} \exp(-\alpha E_\theta) \quad (10)$$

که $Z_\theta(\alpha) = \int \exp(-\alpha E_\theta) d\theta$ فاکتور نرمالیزه‌شده از گرفته شده که در آن $\int p(\theta) d\theta = 1$ است؛ بنابراین توزیع اولیه پارامترها از رابطه ۱۱ به دست می‌آید که در آن ضریب نرمالیزه‌شده $Z_\theta(\alpha)$ به صورت رابطه ۱۲ است.

$$p(\theta) = \frac{1}{Z_\theta(\alpha)} \exp\left(-\frac{\alpha}{2} \|\theta\|^2\right) \quad (11)$$

$$Z_\theta(\alpha) = \int \exp(-\alpha E_\theta) d\theta = (2\pi/\alpha)^{w/2} \quad (12)$$

که w تعداد پارامترهای کل است. اگر $\|\theta\|$ بزرگ باشد، E_θ ممکن است به اندازه کافی بزرگ بوده و $p(\theta)$ کوچک باشد. به دلیل اینکه پارامتر α توزیع پارامترهای دیگر را کنترل می‌کند، هایپر پارامتر نامیده می‌شود. مزیت توزیع گاوسی ضرایب نرمالیزه‌سازی $Z_D(\beta)$ و $Z_\theta(\alpha)$ به صورت تحلیلی است. انواعی از روش‌های جایگزین برای برآورد پارامتر پیشین وجود دارد. در رابطه ۱۳ لگاریتم منفی احتمال شرطی با توجه به مقادیر $s(\theta)$ در رابطه ۱۴ و $Z_S(\alpha, \beta)$ در رابطه ۱۵ کمینه‌سازی خواهد شد.

$$p(\theta \setminus D) = \frac{1}{Z_S(\alpha, \beta)} \exp(-\beta E_D - \alpha E_\theta) = \frac{1}{Z_S(\alpha, \beta)} \exp(-s_\theta) \quad (13)$$

که

$$s(\theta) = \beta E_D + \alpha E_\theta \quad (14)$$

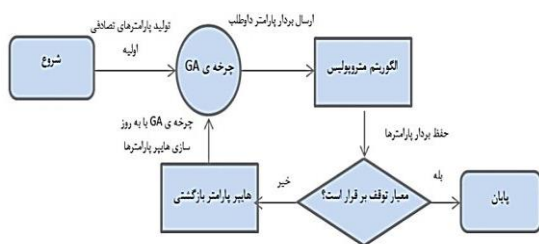
$$Z_S(\alpha, \beta) = \int \exp(-\beta E_D - \alpha E_\theta) d\theta \quad (15)$$

در اینجا مسئله اصلی یافتن لایه پارامتر θ_{MP} براساس ماکزیمم‌سازی تابع پیشین است که می‌تواند با کمینه‌کردن الگوریتم منفی رابطه ۱۴ قابل‌دستیابی باشد؛ زیرا نرمالیزه‌سازی فاکتور در معادله ۱۲ مستقل از پارامترهاست. استخراج بیشینه توزیع پیشین معادل مینیمم‌سازی $s(\theta)$ در رابطه ۱۶ است؛ بنابراین $s(\theta)$ را می‌توان به صورت رابطه ۱۶ نوشت.

$$s(\theta) = \frac{\beta}{2} \sum_{i=1}^N (y_{ii} - \hat{f}_t(x_i, \theta))^2 + \frac{\alpha}{2} \|\theta\|^2 \quad (16)$$

محتمل‌ترین مقدار لایه پارامتر به کمک [] نشان داده می‌شود که می‌تواند با کمینه‌کردن طرف راست رابطه ۱۶ انجام شود. اگر α و β ثابت باشند، N افزایش می‌یابد و

مطالعات پیشین از تابع عضویت فازی به منظور در نظر گرفتن عدم قطعیت‌ها به کمک استنتاج فازی استفاده شده است. در روند امتیازدهی بیش از یک تابع تناسب در روند تولید کاربرد دارد. علاوه بر این پارامترهای تولیدشده مداوم در شبیه‌سازی MC، براساس احتمال پیشین با یکدیگر ترکیب شده‌اند. به منظور اجرای GAMC، ابتدا مقادیر اولیه برای داده‌های پارامتر و هایپر پارامتر در رابطه ۱۶ مقدار می‌گیرند و پارامترها از توزیع یکنواخت تصادفی تولید می‌شوند. سپس a و b براساس رابطه ۲۵ به دست می‌آیند. کاربرد این الگوریتم هایپر پارامترها را ثابت می‌کند و اگر برازش داده‌ها بیش از حد مهم باشد، باید نسبت $\frac{a}{b}$ نزدیک به صفر یا ۰/۱ یا ۰/۰۱ باشد. پس از اقدامات اولیه، پارامترها با چرخه GA به روز می‌شوند. در شکل ۱، چرخه الگوریتم پیشنهادی آمده است.



شکل ۱. چرخه رویکرد مونت کارلو ژنتیک

تجزیه و تحلیل داده‌ها

رشد اقتصادی در ایران با در نظر گرفتن نرخ تولید ناخالص ملی یا تولید ناخالص داخلی در سال مورد نظر، به نسبت مقدار آن در یک سال پایه محاسبه می‌شود. به دلیل هم‌خوانی دقیق نرخ رشد اقتصادی و نرخ تولید ناخالص داخلی، در این پژوهش از آمارهای اعلام‌شده در مرکز آمار استفاده شد.

مدل‌سازی و سیستم پیش‌بینی

در شکل ۲، سری‌های زمانی فصلی نرخ رشد تولید ناخالص داخلی آمده است. داده‌های آزمون نیز به منظور پیش‌بینی نتایج آینده برای ۱۴ سری فصلی مطابق شکل ۳ است. سری‌ها شامل ۹۸ مورد مشاهده هستند. به منظور چک کردن نرمال بودن این سری‌ها از آزمون کلموگروف-اسمیرنوف (K-S) استفاده شد. براساس نتایج این تست، سری‌های زمانی در

ثابت‌شده α و β تعریف‌شده در رابطه ۱۸ برآورد شوند. در این مقاله، مقادیر تأثیرگذار از α و β به صورت اتوماتیک با پیشنهاد الگوریتم تکاملی تکراری برآورد شد که بر پایه آنالیز تحلیلی در تقریب عمل می‌کند. همچنین برای فرمول‌بندی برآورد α و β از انتگرال رابطه ۲۱ استفاده شد.

$$P(\theta \setminus D) = \iint p(\theta, \alpha, \beta \setminus D) d\alpha d\beta \quad (21)$$

$$= \frac{1}{p(D)} \iint p(D \setminus \theta, \beta) p(\alpha) p(\beta) d\alpha d\beta$$

[] و $P(\theta \setminus \alpha, \beta) = p(\theta \setminus \alpha)$ در آن وجود دارد؛ زیرا احتمال دوره مستقل از α و مطابق آن، پیشینه پارامترها مستقل از β است. علاوه بر این، به دلیل استقلال هایپر پارامترها $P(\alpha, \beta) = p(\alpha) p(\beta)$ می‌توان به انتخاب معینی برای پیشینه هایپر پارامتر پرداخت تا انتگرال رابطه ۲۱ سنجیده شود. با تغییر حالت از مقدار قدیمی به مقدار جدید می‌توان مقادیر را به صورت زیر تبدیل کرد.

$p(\alpha) = 1/\alpha$ و $p(\beta) = 1/\beta$ $0 < \alpha < \infty$ $0 < \beta < \infty$
در این چارچوب، از آنالیز انتگرال برای پارامترها استفاده می‌شود. در این میان با در نظر گرفتن روابط ۸ و ۱۰ می‌توان انتگرال ۲۱ را ارزیابی کرد.

$P(\theta) = \int_0^\infty p(\theta \setminus \alpha) p(\alpha) d\alpha = \Gamma(w/2) / (2\pi E_\theta)^{w/2}$ (22)
که در آن Γ تابع گامای استاندارد است و می‌تواند به صورت رابطه ۲۳ تعمیم بیابد.

$P(D \setminus \theta) = \int_0^\infty p(D \setminus \theta, \beta) p(\beta) d\beta = \Gamma(N/2) / (2\pi E_\theta)^{N/2}$ (23)
با توجه به رابطه ۲۲ و ۲۳ رابطه ۲۴ و ۲۵ به دست می‌آید:

$$P(\theta \setminus D) = \frac{1}{P(D)} \frac{\Gamma(N/2)}{(2\pi E_D)^{N/2}} \frac{\Gamma(w/2)}{(2\pi E_\theta)^{w/2}} \quad (24)$$

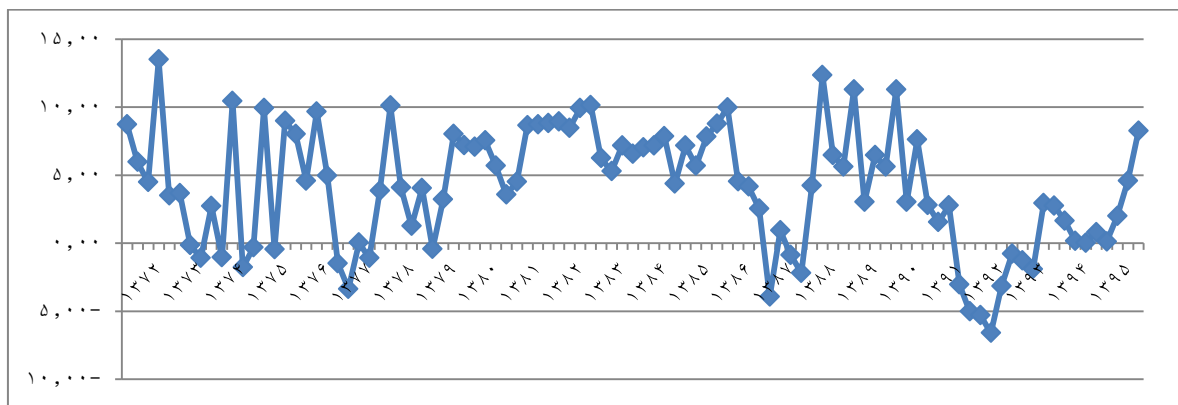
$$\alpha_{eff} = \frac{W}{2E_\theta}$$

$$\beta_{eff} = \frac{N}{2E_D} \quad (25)$$

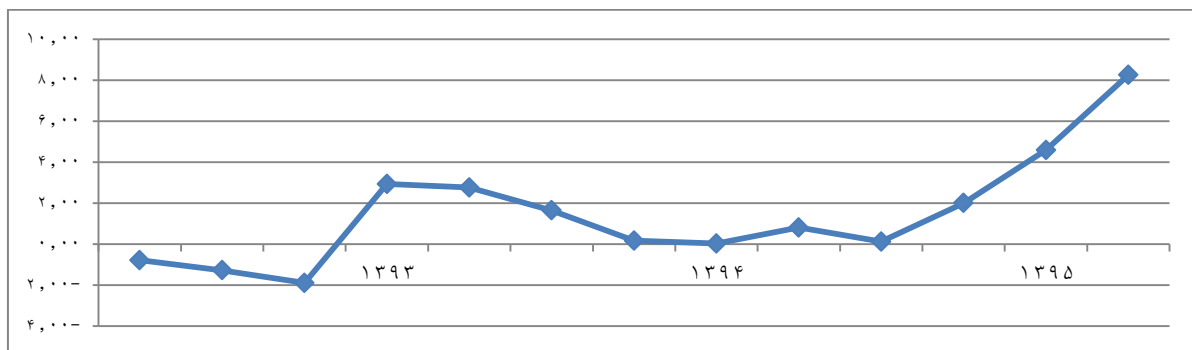
در نهایت الگوریتم آموزشی GAMC که در تحقیق Aşikgil و Kocadağlı معرفی شده مورد استفاده قرار گرفته است [۲۷]. هدف توسعه الگوریتم آموزش بیزی، کنترل پیچیدگی استفاده از هایپر پارامترهای α_{eff} و β_{eff} و تخمین پارامترهای تقریب گاوسی شبکه عصبی مصنوعی بیزی به طور صحیح است. در این فرایند، برخلاف

به‌دلیل مزیت‌های فراوان آن برای برآورد ماتریس هسین استفاده شد. به‌منظور تقریب گاووسی از هایپر پارامترهای بازگشتی شبکه‌های عصبی مصنوعی بیزی، الگوریتم MC-GA به‌کار رفت. بدین‌منظور عملکرد شبکه‌های عصبی مصنوعی و شبکه‌های عصبی مصنوعی بیزی در برابر فضای هایپر پارامتری چندبعدی با به‌کارگیری تعداد گوناگونی از نوروها در لایه‌های پنهان ارزیابی شدند.

سطح معناداری $\alpha = 0.05$ و $P\text{-value} = 0.02$ از توزیع نرمال برخوردار بودند. همچنین به‌کمک نمودار جعبه‌ای مشخص شد مشاهدات قابل توجه دور از مرکزی وجود ندارد. در این مطالعه، به‌منظور آموزش شبکه‌های عصبی مصنوعی پیش‌خور چندلایه از روش الگوریتم بهینه‌سازی گرادیان نزولی برای همگرایی سریع این روش به مقدار بهینه و همچنین تقریب گاووسی با هایپر پارامترهای ثابت در شبکه عصبی مصنوعی بیزی استفاده شد. سپس از روش شبه‌نیوتن



شکل ۲. داده‌های نرخ رشد تولید ناخالص داخلی ایران از سال ۱۳۷۱-۱۳۹۵



شکل ۳. داده‌های آزمون به‌منظور پیش‌بینی نتایج آینده برای ۱۲ سری فصلی

مقایسه روش‌های تخمین بیان می‌شود.

آنالیز سری‌های زمانی

پیش از شروع آنالیز، ابتدا لگاریتم طبیعی سری‌های زمانی برای تشخیص نویزها مشخص شد؛ بنابراین بررسی آلترناتیوهای مدل ARMA صورت گرفت. براساس شکل ۴ نیز نتایج هم‌بستگی لگاریتمی انجام‌شده روی داده‌های سری زمانی تولید ناخالص داخلی ایران ذکر شد و تخمین و سنجش بهترین مدل انتخابی با استفاده از OLS صورت

همچنین برای تخمین تعداد نوروهایی که باید در لایه پنهان حضور داشته باشند از BIC^1 ، $AICc$ و AIC استفاده شد. علاوه‌براین، به‌منظور تعیین عملکرد مدل تخمین زده‌شده، به‌ترتیب از معیارهای میانگین مربعات خطا و خطای استاندارد میانگین در داده‌های آموزش و آزمون استفاده شد. پس از نوشتن برنامه روش پیشنهادی در Matlab، برای آنالیز سری‌های زمانی کلاسیک از EViews 7 استفاده شد. در ادامه نتایج پیش‌بینی و

می توان مشاهده کرد که مدل $ARMA(0,3)(0,0)$ با مقدار خطای استاندارد میانگین کمتر و سایر مدل ها شامل $MA(3)$ نشان دهنده بهترین برآزش برای داده های آموزش هستند. ممکن است سایر مدل های مرتب شده نیز قابلیت خوبی در پیش بینی داشته باشند، اما مدل $MA(3)$ با توجه به پیچیدگی مدل مناسب تر است. در جدول ۲، نتایج مقایسه ای اجرای مدل های $ARMA$ آمده است.

پیش بینی با شبکه های عصبی پیش خور

نرخ های مختلف یادگیری برای تعیین بهترین گام اندازه گیری شد i و بهترین آن ها در جدول ۳ آمده است. براساس این برآزش پیش بینی برحسب شبکه های عصبی انجام شد که نتایج آن در ادامه بیان می شود. ۷۵ درصد داده ها برای آموزش، ۱۵ درصد در اعتبارسنجی و ۱۰ درصد در آزمون استفاده شده اند. به منظور دستیابی به عملکرد پیش بینی خوب، شبکه های عصبی مصنوعی پیش خور با بهترین مدل ورودی پیشنهادی استفاده از ترکیب های مختلف نورون ها، به عنوان داده های ورودی تشکیل شده اند. به منظور تخمین اینکه چند نورون باید در لایه پنهان وجود داشته باشند، از روش های AIC ، $AICC$ و BIC استفاده شد. عملکرد مدل تخمین و پیش بینی به ترتیب با اندازه گیری میانگین مربعات خطا و خطای استاندارد میانگین در آموزش و تست داده ها صورت گرفت. همچنین به منظور اندازه گیری خطا، ارتباط میان هدف و خروجی شبکه و سه معیار پیچیدگی برای جست و جوی بهترین مدل در نظر گرفته شد. در جدول ۳، بهترین مدل ها با فونت ضخیم آمده است. این مدل ها کوچک ترین میانگین خطای استاندارد میانگین و مقدار پیچیدگی منطقی را در دسته ورودی های متفاوت دارند.

رویکرد مونت کارلو ژنتیک

در این پژوهش، آلفا و بتای بازگشتی در رویکرد گاوسی به صورت اتوماتیک با $MCGA$ به جای ثابت نگه داشتن آن ها تخمین زده می شوند. در این الگوریتم، چرخه اصلی برای چرخه GA به منظور به روز کردن پارامترها و هایپر پارامترها در هر چرخه صورت می گیرد. به عبارت دیگر، چرخه GA در داخل تعدادی نسل تعریف شده در هر تکرار از چرخه اصلی انجام می شود و در پایان به روزرسانی هر تکرار صورت می گیرد. به منظور تخمین پارامترها و هایپر پارامترها،

گرفت. به منظور بررسی عملکرد شبکه های عصبی مصنوعی سنتی و شبکه های عصبی مصنوعی بیزی، سری زمانی با ساختار $MA(3)$ به عنوان ورودی تعریف شدند. همچنین پس از معرفی y_t به عنوان هدف، در نهایت مقادیر y_t با به کارگیری مدل های شبکه های عصبی مصنوعی و شبکه های عصبی مصنوعی بیزی برآورد شدند. با توجه به اینکه برآزش مدل میانگین متحرک به دلیل عدم قابلیت مشاهده جملات، تأخیر پیچیده تری از مدل خود هم بسته دارند، در نتیجه می توان بیشینه تعداد تأخیر مناسب برای تخمین را با بررسی تابع خود هم بستگی نمونه تعیین کرد.

روش هموارسازی نمایی

این روش برای سری های زمانی داده های فصلی تولید ناخالص داخلی به منظور پیش بینی بهتر کاربرد دارد. در جدول ۱، نتایج این روش آمده است. براین اساس، رویکردهای $Single$ Exponential و $Holt-Winters$ از نظر مقدار میانگین مربعات خطا تأثیر یکسانی بر سری های زمانی دارند.

مدل اتورگرسیو

پیش از به کار بردن OLS در آنالیز مدل آلترناتیو $ARMA$ ، سری ها از نظر ایستایی چک شدند و برای این منظور از آزمون $Augmented$ Dickey-Fuller استفاده شد. براساس این آزمون، سری های زمانی و مدل $ARMA$ باید به اثبات برسند. از این خروجی می توان نتیجه گرفت که مدل داده های بررسی شده غیر مانا و غیر خطی است.

Sample: 1371Q1 1395Q2
Included observations: 98

	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
1	0.495	0.495	24.736	0.000		
2	0.427	0.242	43.379	0.000		
3	0.376	0.135	57.983	0.000		
4	0.064	-0.306	58.411	0.000		
5	0.104	0.054	59.555	0.000		
6	0.106	0.121	60.751	0.000		
7	0.141	0.208	62.904	0.000		
8	0.041	-0.241	63.086	0.000		
9	0.093	0.020	64.039	0.000		
10	0.072	0.027	64.619	0.000		
11	-0.120	-0.125	66.236	0.000		
12	0.040	0.097	66.417	0.000		
13	0.020	0.078	66.461	0.000		
14	-0.060	-0.079	66.877	0.000		
15	0.126	0.104	68.756	0.000		
16	0.067	-0.004	69.288	0.000		
17	-0.024	-0.113	69.358	0.000		
18	0.054	0.061	69.716	0.000		
19	-0.028	-0.061	69.816	0.000		
20	-0.071	-0.004	70.446	0.000		

شکل ۴. نتایج هم بستگی لگاریتمی داده ها

MCGA به تعریف تعداد تکرارها برای چرخه اصلی و تعداد تولید نسل برای چرخه GA نیاز دارد.

جدول ۱. نتایج اجرای رویکرد هموارسازی نمایی

Exponential smoothing	RMSE	MSE	Test SSR	Test SSE	Mean	Alpha	Beta
Single exponential	۳/۹۰۹۱۹۷	۰/۰۱۴۵۹	۴۹۷/۶۱۹	۰/۱۶۹۸	۵/۲۲۶۹۵۰	۰/۴۵۲۰	---
Double exponential	۴/۱۲۸۸۴۶	۰/۰۱۷۹۲	۱۶۷۰/۶۴۲	۰/۲۱۷۶	۴/۷۸۷۸۳۳	۰/۲۰۸۰	---
Holt-Winters (no seasonal)	۳/۹۴۹۱۶۱	۰/۰۱۵۲۶	۱۵۲۴/۵۲۸	۰/۱۷۳۴	۵/۸۱۶۶۳۰	۰/۵۱۰۰	۰/۰۰۰۰

جدول ۲. نتایج مقایسه‌ای اجرای مدل‌های ARMA

ARIMA Model	[]	Correlation coefficient	MSE	AIC	AICc	BIC	Test SSE	Durbin Watson stat
(۰/۳) (۰/۰)	۴	۰/۶۶۴۷۸۹	۰/۰۴۳۷۴۶	۵/۳۹۳۸۷۴	۵/۴۰۱۵۰۳	۵/۵۲۵۷۶۰	۰/۱۰۹۰۸۷	۲/۰۱۶۷۱۳
(۱/۳) (۰/۰)	۶	۰/۶۷۴۹۵۵۴	۰/۰۳۴۵۶۴۹	۵/۴۰۶۹۵۹	۵/۴۳۵۶۵۲	۵/۵۶۷۸۹۱	۰/۱۰۹۹۳۷۸	۱/۹۸۹۴۷۸
(۰/۴) (۰/۰)	۵	۰/۶۹۱۱۵۸۷	۰/۰۳۱۸۵۹۶	۵/۴۱۱۶۳۵	۵/۴۵۶۹۱۱	۵/۵۶۹۲۹۶	۰/۱۸۸۵۷۹	۱/۸۶۵۶۱۲
(۲/۳) (۰/۰)	۷	۰/۶۹۵۳۲۵۱	۰/۰۲۴۶۵۵۵	۵/۴۲۱۸۷۵	۵/۴۶۷۸۲۰	۵/۶۱۲۲۲۵	۰/۱۲۷۷۶۴	۱/۹۷۸۳۷۵
(۱/۴) (۰/۰)	۷	۰/۶۹۸۶۷۳۴	۰/۰۱۷۵۶۹۳	۵/۴۲۹۳۸۷	۵/۴۷۱۱۵۷	۵/۶۱۳۹۸۷	۱/۱۳۹۰۵۴	۱/۸۲۴۹۷۵

جدول ۳. اجرای شبکه عصبی پیش‌خور

Inputs Neuron	Par	Correlation coefficient	MSE	AIC	AICc	BIC	Test SSE	Learning rate	Iter
(۰/۰)	۲	۰/۸۴۶۹	۰/۰۱۱۶۵۹	۴/۳۹۵۶۳	۴/۳۰۶۹۵۸۶	۳/۸۸۷۰۶۱	۰/۱۲۸۶۳۱	۰/۵	۱۰۰۰
(۰/۳)	۳	۰/۸۶۹۱	۰/۰۱۱۵۶۹	۴/۵۶۹۸۳	۴/۷۳۶۹۸۵۵	۴/۰۶۱۴۸۹	۰/۱۱۹۳۸۷	۰/۱	۱۰۰۰
	۴	۰/۸۶۱۴	۰/۰۱۰۲۵۹	۴/۹۳۶۵۸	۵/۴۵۶۹۸۳۱	۴/۴۶۸۸۹۷	۰/۹۹۸۵۷۹	۰/۵	۱۰۰۰
	۵	۰/۸۷۱۱	۰/۰۰۹۸۶۶	۴/۱۷۵۶۹	۵/۵۳۹۸۶۱۱	۴/۹۶۹۲۸۹	۰/۸۸۴۱۲۳	۰/۰۵	۱۰

بر اساس بخش‌بندی داده‌ها از Leave-One-Out و تکنیک Cross-Validation استفاده شد. سپس شبکه عصبی مصنوعی بیزی بر دسته داده‌های مختلف به کمک الگوریتم MCGA آموزش می‌بیند. مقدار پارامترها در چرخه Cross-Validation به عنوان نقطه اولیه برای گام آموزش بعدی کاربرد دارد؛ بنابراین کارایی MCGA بهبود می‌یابد. مطابق با چرخه GA تعداد نسل‌ها و حد تلورانس برای تابع برازش به کار می‌رود. بر این اساس مدل خطای استاندارد میانگین بیشتری دارد، اما این مقدار شامل داده‌های آزمون و پیچیدگی مدل نمی‌شود؛ بنابراین نتایج آموزش مطابق با جدول ۴ در تعداد نوروں‌های مشخصی پایان یافته است. با توجه به جدول ۴، میانگین خطای استاندارد بر داده‌های آزمون با تعداد نوروں‌هایی مشخص کاهش می‌یابد. در یک رویکرد بیزی، برای مدل تخمین زده شده با استفاده از ورودی‌های مشابه به دلیل جریمه ذاتی

در شبیه‌سازی، تعریف تعداد تکرارها و تولید نسل امتحان می‌شود و سپس به ارزش منطقی به کاررفته در به کارگیری هزینه و فضای محاسباتی توجه می‌شود. به منظور دستیابی به پیش‌بینی قوی‌تر، تکرار کل به دو صورت مشتعل‌شدن و آموزش تقسیم می‌شود. در شبیه‌سازی، پارامترهای نخبه‌گرایی و تقاطع و جهش در مقادیر ۰/۲، ۰/۸ و ۰/۱ ثابت شدند. شبیه‌سازی با توجه به شبیه‌سازی انجام شده در پژوهش Chambers [۲۸]، Marwala [۲۹] و Michalewicz [۳۰] صورت گرفت. به کمک این مقادیر می‌توان از تنوع تولید نسل در چرخه GA اطمینان یافت. در آنالیز، تغییر نسل و اندازه جمعیت به تعداد نوروں‌ها و ورودی‌های تعریف شده در شبکه وابسته است. تغییر در تولید نسل و اندازه جمعیت تأثیر بسیاری در اجرای آموزش و پیچیدگی مدل تخمینی ندارد. این کمیت‌ها در جدول ۴ ثابت هستند.

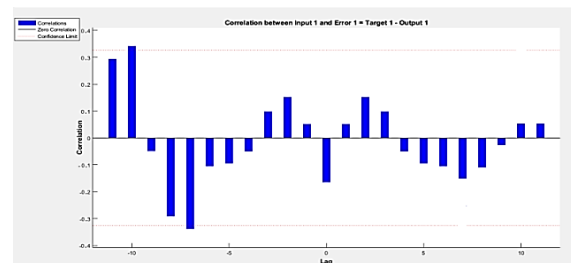
آموزش داده‌ها و بالاترین هم‌بستگی را میان داده‌های هدف و خروجی دارد.

نتیجه‌گیری

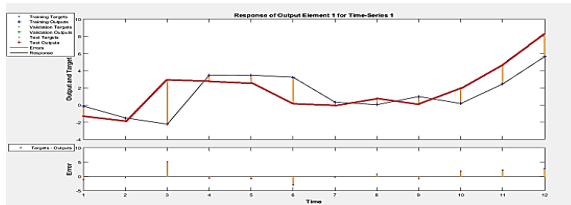
در این پژوهش، از روش MCGA در بهینه‌سازی نتایج حاصل از شبکه عصبی مصنوعی در تحلیل سری‌های زمانی با میانگین متحرک برای پیش‌بینی رشد اقتصادی استفاده شد. همچنین الگوریتم مونت کارلو برای آموزش شبکه عصبی مصنوعی بیزی در پیش‌بینی سری‌های زمانی پیشنهاد شد. در واقع در پژوهش حاضر، رویکرد پیشنهاد شده شبیه‌سازی مونت کارلو و الگوریتم ژنتیک و تابع عضویت فازی را برای پیش‌بینی کاربردی با مقدار خطای کمتر ارائه شد و شبکه‌های عصبی مصنوعی بیزی به کمک سیستم ترکیبی برای پیش‌بینی سری‌های زمانی آموزش دیده شدند. به بیان دیگر، آموزش بیزین راهی طبیعی برای کنترل پیچیدگی مدل است. توابع عضویت فازی نیز برای اندازه‌گیری عدم قطعیت در فرایند پیش‌بینی الگوریتم GA و شبیه‌سازی داده‌های مفقودی در نظر گرفته شدند. همچنین مقایسه روش پیشنهادی از نظر عملکرد با روش‌های سری زمانی کلاسیک و شبکه عصبی پیش‌خور صورت گرفت و مشاهده شد در روش پیشنهادی مزایای فراوانی در مقایسه با روش‌های دیگر وجود دارد؛ برای مثال، به کمک رویکرد پیشنهادی می‌توان به برآورد مدل پرداخت؛ به گونه‌ای که در برابر بیش‌برازش داده‌های آموزش به صورت خودکار جریمه در نظر گرفته شود؛ از این‌رو برآورد مدلی بهتر در جهت بهینه‌سازی خطا ارائه می‌شود. به‌منظور تعمیم مدل نیز استفاده از مدل‌های دیگر در مورد سری‌های زمانی دیگر در آینده پیشنهاد می‌شود.

آموزش شبکه عصبی مصنوعی پیش‌خور و الگوریتم کاهش گرادیان بر پایه کمینه‌کردن میانگین مربعات خطا، به برخی روش‌های پیچیده مانند توقف زود هنگام برای کنترل پیچیدگی، یا تشخیص نرخ یادگیری مناسب برای جست‌وجوی گرادیان نیازمند است. اگر فرایند آموزش به دلیل کاهش گرادیان بسیار زودتر متوقف شود، دستیابی به اتصالات کمتر اجتناب‌ناپذیر خواهد بود، حتی اگر افزایش تکرارها با افزایش نوروها سبب کاهش میانگین مربعات خطا در داده‌های میانگین شود. در نتیجه ممکن است میانگین خطای استاندارد کاهش نیابد. با وجود افزایش تعداد نوروها و ساختار شبکه‌های عصبی مصنوعی، اگر

در مدل‌های پیچیده، با افزایش تعداد نوروها، میانگین مربعات خطا افزایش می‌یابد. با این حال آموزش بیزی از بیش‌برازش به داده‌های آموزش به‌طور اتوماتیک جلوگیری می‌کند. MCGA تعمیم خوبی از داده‌های آزمون ایجاد می‌کند. بهترین مدل به صورت فونت ضخیم در جدول ۴ آمده است. به‌منظور مشخص شدن بهترین مدل، میانگین خطای استاندارد با معیارهای پیچیدگی در نظر گرفته می‌شود. مقادیر خودهم‌بستگی تابع انرژی داده‌ها و تناسب پیش‌بینی در مورد داده‌های هدف با توجه به داده‌های آزمون به ترتیب در شکل‌های ۵ و ۶ آمده است.



شکل ۵. مقادیر خودهم‌بستگی تابع انرژی داده‌ها



شکل ۶. تناسب پیش‌بینی در مورد داده‌های هدف با توجه به داده‌های آزمون

مدل‌ها براساس پایین‌ترین میانگین خطای استاندارد برای مقایسه بهترین برآوردها با روش‌های مختلف از نظر پیش‌بینی بر داده‌های آزمون در جدول ۵ آمده است. پیش‌بینی بر داده‌های آزمون MSE روی داده‌های آموزش و هم‌بستگی میان هدف و خروجی به‌عنوان عملکرد آموزش در نظر گرفته شده است. براساس جدول ۵، توانایی پیش‌بینی بر داده‌های آموزش شبکه‌های عصبی مصنوعی و شبکه‌های عصبی مصنوعی بیزی از روش‌های سری زمانی کلاسیک بیشتر است. بر پایه پایین‌ترین میانگین خطای استاندارد بر داده‌های آزمون مدل‌ها به کمک GAMC برآورد شد که بهترین عملکرد پیش‌بینی را دارد. علاوه بر این، مدل میانگین مربعات خطای کمتری برای

انتخاب بیش‌ازحد پیچیده شود، ریسک گیرافتادن در نقطهٔ بهینهٔ محلی افزایش می‌یابد.

جدول ۴. اجرای رویکرد مونت‌کارلو ژنتیک با هایپر پارامترهای بازگشتی

Inputs	Neuron	Par	Correlation coefficient	MSE	AIC	AICc	BIC	Test SSE	Gen	Pop	Learn	Bum
(۰/۳)	۲	۱۱	۰/۸۶۴۷	۰/۰۱۳۷۴	۴/۵۳۵۳	۴/۳۵۹۶	۳/۸۸۷۰۶	۰/۱۱۳۷	۲۰	۵۰	۲۰	۵
(۰/۰)												
	۳	۱۶	۰/۸۵۷۶	۰/۰۱۱۴۷	۴/۶۸۷۵	۴/۴۷۶۵	۴/۰۶۱۴۷	۰/۱۱۹۹۳	۲۰	۵۰	۲۰	۵
	۴	۲۱	۰/۸۵۹۷	۰/۰۱۰۲۵	۴/۴۱۷۷	۴/۲۷۱۸	۴/۴۶۸۴۱	۰/۱۰۸۸۹	۲۰	۵۰	۲۰	۵
	۵	۲۶	۰/۸۵۲۱	۰/۰۰۹۶۷	۴/۴۷۵۴۳	۴/۲۳۱۸۲	۴/۹۶۹۲۴	۰/۰۸۲۴	۲۰	۵۰	۲۰	۵

جدول ۵. مقایسهٔ بهترین مدل تخمینی با به‌کارگیری روش‌های مختلف

SSE	MSE	R	P	روش‌ها
۰/۰۸۲۴	۰/۰۰۹۶۷	۰/۸۵۲۱	۲۱	رویکرد GAMC
۰/۸۸۴۱	۰/۰۰۹۸۶	۰/۸۷۱۱	۲۶	شبکهٔ عصبی پیش‌خور
۰/۱۰۹۰	۰/۰۴۳۷۴	۰/۶۶۴۷	۴	اتورگرسیو میانگین متحرک
۰/۱۶۹۸	۰/۰۱۴۵۹	-	-	هموارسازی نمایی

منابع

- Granger, C. W. J., and Newbold, P. (2014). *Forecasting Economic Time Series*: Academic Press.
- Cutright, P. (1965). Political structure, economic development, and national social security programs. *American journal of sociology*, Vol.70, No.5, PP. 537-550.
- Bjork, G. J. (1999). *The Way it Worked and Why It Won't: Structural Change and the Slowdown of U.S. Economic Growth*: Praeger; Complete Numbers Starting.
- Jones, C. L. (2013). *Introduction to Economic Growth (3 Edition Ed.)*: W. W. Norton and Compan.;
- Einicke, G. A. (2012). *Smoothing, Filtering and Prediction: Estimating the Past, Present and Future*: Intech.
- Stock, J. H., and Watson, M. W. (2003). *Introduction to Econometrics*, Addison Wesley Boston.
- Baffigi, A., Golinelli, R., and Parigi, G. (2004). "Bridge Models to Forecast the Euro Area GDP", *International Journal of Forecasting*, Vol. 20, No. 3, PP. 447-460 .
- Schumacher, C. (2007). "Forecasting German GDP Using Alternative Factor Models Based on Large Datasets", *Journal of Forecasting*, Vol. 26, No. 4, PP. 271-302.
- Schumacher, C., and Breitung, J. (2008). "Real-Time Forecasting of German GDP Based on a Large Factor Model with Monthly and Quarterly Data", *International Journal of Forecasting*, Vol. 24, No. 3, PP. 386-398.
- Marcellino, M., and Schumacher, C. (2010). "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP", *Oxford Bulletin of Economics and Statistics*, Vol. 72, No. 4, PP. 518-550.
- Bañbura, M., and Rünstler, G. (2011). "A Look Into the Factor Model Black Box: Publication Lags and the Role of Hard and Soft Data in Forecasting GDP", *International Journal of Forecasting*, Vol. 27, No. 2, PP. 333-346.
- Zhang, G. P., and Qi, M. (2005). "Neural Network Forecasting for Seasonal and Trend Time Series", *European Journal of Operational Research*, Vol. 160, No. 2, PP. 501-514.
- Pouzols, F. M., Lendasse, A., and Barros, A. B. (2010). "Autoregressive Time Series Prediction by Means of Fuzzy Inference Systems Using Nonparametric Residual Variance Estimation", *Fuzzy Sets and Systems*, Vol. 161, No. 4, PP. 471-497.
- Giovanis, E. (2010), A Study of Panel Logit Model and Adaptive Neuro-Fuzzy Inference System in the Prediction of Financial Distress Periods, *World Academy of Science, Engineering and Technology*, Vol. 64, PP. 646-652.
- Tealab, A. (2018). Time series forecasting using artificial neural networks methodologies: A systematic review, *Future Computing and Informatics Journal*, Vol. 3, No. 2, PP. 334 – 340.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*: John Wiley and Sons.
- Davidson, R., and Mackinnon, J. G. (1993). *Estimation and Inference In Econometrics*.

18. Fuller, W. A. (2009). *Introduction to Statistical Time Series*: John Wiley and Sons.
 19. Geman, S., Bienenstock, E., and Doursat, R. (1992). "Neural Networks and the Bias/Variance Dilemma", *Neural Computation*, Vol. 4, No. 1, PP. 1-58.
 20. Liang, F. (2005). "Bayesian Neural Networks for Nonlinear Time Series Forecasting", *Statistics and Computing*, Vol. 15, No. 1, PP. 13-29.
 21. Mackay, D. J. (1995). "Probable Networks and Plausible Predictions—A Review of Practical Bayesian Methods for Supervised Neural Networks", *Network: Computation in Neural Systems*, Vol. 6, No. 3, PP. 469-505.
 22. Neal, R. M. (1996). "Priors for Infinite Networks", *Bayesian Learning for Neural Networks* (PP. 29-53): Springer.
 23. Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*: Oxford University Press.
 24. Holmes, C., and Mallick, B. (1998). "Bayesian Radial Basis Functions of Variable Dimension", *Neural Computation*, Vol. 10, No. 5, PP. 1217-1233.
 25. Freitas, J. (2000). *Bayesian Methods for Neural Networks* [Ph. D. Thesis]: Trinity College University of Cambridge and Cambridge University Engineering Department, Cambridge, UK.
 26. Mackay, D. J. (1992). "A Practical Bayesian Framework for Backpropagation Networks", *Neural Computation*, Vol. 4, No. 3, PP. 448-472.
 27. Chua, C., and Goh, A. (2003). "Nonlinear Modeling with Confidence Estimation Using Bayesian Neural Networks", *International Journal for Numerical and Analytical Methods in Geomechanics*, No. 27, PP. 651-667.
 28. Kocadağlı, O., and Aşıkıl, B. (2014). "Nonlinear Time Series Forecasting with Bayesian Neural Networks", *Expert Systems with Applications*, Vol. 41, No. 15, PP. 6596-6610.
 29. Chambers JA, Sherliker W and Mandic DP 2000 A normalised gradient algorithm for an adaptive recurrent perceptron. In Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-2000), Vol. 1, PP. 396–399.
 30. Marwala, T. (2007). Bayesian training of neural networks using genetic programming. *Pattern Recognition Letters*, Vol. 28, No. 12, PP. 1452-1458 .
 31. Michalewicz, Z., : *Genetic algorithms + data STRUCTURES = evolution programs*. Springer, New York (1996)
-