Arshive of SID

مجله مهندسی مالی و مدیریت اوراق بهادار
شماره سی‌وچهارم / بهار ۱۳۹۷

پیش‌بینی دامنه تغییرات طلا با استفاده از مدل ترکیبی ARIMA و شبکه عصبی

شاپور محمدی[۱]
رضا راعی[۲]
محمدرضا رحیمی[۳]

تاریخ دریافت: ۹۶/۰۶/۲۶ تاریخ پذیرش: ۹۶/۰۹/۱۱

چکیده

مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) که تحت عنوان روش باکس و جنکینز شناخته می‌شود، یکی از پرکاربردترین مدل‌ها در پیش‌بینی سری‌های زمانی است. اما پیش فرض اصلی این مدل خطی بودن سری‌های زمانی می‌باشد. از سوی دیگر شبکه‌ی عصبی یک تخمین زننده‌ی عمومی است که الگوهای غیر خطی را بسیار خوب مدل‌سازی می‌نماید. دانستن الگوی داده‌ها مبنی بر خطی و غیر خطی بودن در واقعیت کمی دشوار است، بنابراین این ایده در ذهن ایجاد می‌گردد که تلفیق مدل‌های خطی و غیرخطی می‌تواند منجر به افزایش دقت پیش‌بینی گردد. از این رو، در این پژوهش بخش خطی را بوسیله‌ی مدل ARIMA پیش‌بینی کرده، آن‌گاه پسماندهای غیر خطی را بوسیله‌ی شبکه‌ی عصبی پیش‌خور مدل سازی نموده و پیش‌بینی حاصل از آن را به مدل ARIMA ، به‌منظور پیش‌بینی حد بالای قیمت، حد پایین قیمت و قیمت پایانی اونس طلا (برای یک مرحله پیش‌رو) اضافه می‌نماییم. نتایج بررسی دقت مدل ترکیبی نسبت بر هر یک از مدل‌های ARIMA و شبکه‌ی عصبی بر اساس دو معیار MSE و MAE با استفاده از آزمون-های مقایسه زوجی و دایبولد- ماریانو دال بر عملکرد بهتر مدل ترکیبی است.

واژه‌های کلیدی: خود رگرسیو میانگین متحرک انباشته (ARIMA) ، شبکه عصبی پیش‌خور، مدل ترکیبی.

۱- دانشیار، اقتصاد سنجی، دانشگاه تهران، تهران، ایران
۲- دانشیار، مدیریت مالی، دانشگاه تهران، تهران، ایران
۳- دانشجوی دکتری، مدیریت مالی، دانشگاه تهران، تهران، ایران (نویسنده مسئول) m.rahimi66@ut.ac.ir

www.SID.ir

Arshive of SID

۱- مقدمه

پیش‌بینی قیمت همواره یکی از چالش برانگیزترین مسائل پیش‌روی سفته‌بازان، معامله‌گران، کارگزاران و ... می‌باشد. از طرفی در تحلیل فاصله‌ای[۱] فرض می‌شود که مشاهدات و تخمین‌ها در دنیای واقعی معمولاً کامل و مطمئن نمی‌باشند، بنابراین برای افزایش دقت، باید داده‌ها را به صورت فاصله‌هایی که کمیت‌های واقعی را در بر می‌گیرند بیان نمود، از این رو، در این پژوهش علاوه بر پیش‌بینی قیمت پایانی به پیش‌بینی دامنه‌ی نوسان قیمت (بالاترین و پایین‌ترین قیمت) پرداخته شده است.

خودرگرسیو میانگین متحرک انباشته (ARIMA) که خلاصه شده‌ی (Autoregressive Integrated Moving Average) می‌باشد و تحت عنوان روش باکس و جنکینز[۲] شناخته می‌شود، یکی از پرکاربردترین مدل‌ها در پیش‌بینی سری‌های زمانی در طول سه دهه‌ی گذشته بوده است، اما پیش‌فرض اصلی آن این است که رابطه‌ی خطی میان ارزش‌های سری بر قرار است بنابراین رابطه‌های غیرخطی بوسیله‌ی مدل خودرگرسیو میانگین متحرک انباشته نمی‌توانند خوب توضیح داده شوند.

یکی دیگر از روش‌های مورد استفاده در پیش‌بینی سری‌های زمانی شبکه‌ی عصبی است که توان تخمین روابط غیرخطی مختلفی را دارا می‌باشد (اصطلاحاً به شبکه‌ی عصبی تخمین زننده‌ی همگانی[۳] می‌گویند). اما استفاده از شبکه، طبق ادبیات پژوهش در روابط خطی نتایج پیچیده‌ای در بر داشته است. دانستن الگوی داده‌ها، مبنی بر خطی و غیرخطی بودن در دنیای واقعی کمی دشوار است و به ندرت سری‌های زمانی به طور خالص خطی و غیرخطی می‌باشند و اغلب از هر دو الگو تبعیت می‌کنند. بنابراین مسأله اینجاست که چگونه می‌توانیم قیمت پایانی و دامنه‌ی نوسان قیمت را با خطای کمتری پیش‌بینی کنیم؟

۲- مبانی نظری و مروری بر پیشینه پژوهش

در زمینه مدل‌سازی سرهای زمانی، روش‌های متفاوتی مورد استفاده قرار می‌گیرند. مدل‌های سنتی مانند میانگین متحرک، هموارسازی نمایی و خودرگرسیو میانگین متحرک انباشته پیش‌بینی آینده را به روابط خطی از گذشته محدود می‌نمایند، از این مدل‌ها به دلیل سادگی در فهم و کاربرد در دهه‌های اخیر بسیار استفاده شده است.

به دلیل مشاهده الگوهای غیرخطی در دنیای واقعی، یک سری از مدل‌های غیرخطی مانند ناهمسانی واریانس شرطی و ناهمسانی واریانس شرطی تعمیم یافته مطرح گردیدند. همگی این مدل‌ها، الگوهای غیرخطی بخصوصی را توضیح می‌دهند، اما شبکه عصبی مصنوعی (ANN[۴]) توان و

www.SID.ir

Arshive of SID

قدرت پیش‌بینی روابط غیرخطی متفاوتی را داراست و کاملاً انعطاف پذیر عمل می‌نماید، ولی آنچه از ادبیات پژوهش بر می‌آید وجود نتایج متفاوت در استفاده از شبکه‌ی عصبی برای روابط خطی است. برای مثال مارخام و راکز (۱۹۹۸) اذعان داشتند، عملکرد شبکه‌ی عصبی برای مسأله‌های رگرسیون خطی وابسته به اندازه نمونه و سطح شوک[۵] می‌باشد[16]. بنابراین چون دانستن خصوصیت داده‌ها مبنی بر خطی و یا غیرخطی بودن دشوار است و از طرفی دیگر به ندرت روابط کاملاً خطی و یا کاملاً غیرخطی مشاهده می‌شوند (اغلب ترکیبی از دو الگو را دارا می‌باشند)، بنابراین ترکیبی از مدل‌های خودرگرسیو میانگین متحرک انباشته و شبکه‌ی عصبی می‌تواند به افزایش دقت پیش‌بینی منجر گردد. مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) مناسب، مدلی است که پسماندهای آن همبستگی خطی نداشته باشند، اما تحلیل‌های پسماند قادر به شناسایی الگوهای غیرخطی نمی‌باشند؛ بنابراین وجود روابط غیرخطی در پسماندها می‌تواند محدودیتی برای مدل خودرگرسیو میانگین متحرک انباشته به حساب آید.

در مدل ترکیبی سعی بر این است که روابط غیرخطی در پسماندها با استفاده از شبکه‌ی عصبی شناسایی شده و پیش‌بینی پسماندهای حاصل از شبکه به مدل خودرگرسیو میانگین متحرک انباشته اضافه گردد. با این ترکیب، بخش خطی بوسیله‌ی مدل خودرگرسیو میانگین متحرک انباشته و بخش غیرخطی بوسیله‌ی شبکه پیش‌بینی خواهد شد.

از آن‌جا که افزایش روز افزون قیمت جهانی طلا بیانگر رشد تقاضا برای این فلز گرانبها و تمایل جهانی برای در اختیار داشتن آن به عنوان یک کالای سرمایه‌ای است، در این پژوهش با استفاده مدل ترکیبی، علاوه بر پیش‌بینی قیمت پایانی، به پیش‌بینی بالاترین و پایین‌ترین قیمت هفتگی اونس طلا پرداخته شده است. در واقع علاوه بر پیش‌بینی خود قیمت، دامنه‌ی نوسان قیمت نیز پیش‌بینی گردیده و فرضیه‌های زیر را مورد آزمون قرار گرفته است:

۱) پیش‌بینی مدل ترکیبی در حد بالای قیمت طلا نسبت به پیش‌بینی مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری می‌باشد.
۲) پیش‌بینی مدل ترکیبی در حد بالای قیمت طلا نسبت به پیش‌بینی شبکه دارای خطای کمتری می‌باشد.
۳) پیش‌بینی مدل ترکیبی در حد پایین قیمت طلا نسبت به پیش‌بینی مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری می‌باشد.
۴) پیش‌بینی مدل ترکیبی در حد پایین قیمت طلا نسبت به پیش‌بینی شبکه دارای خطای کمتری می‌باشد.

www.SID.ir

Arshive of SID

۵) پیش‌بینی مدل ترکیبی در قیمت پایانی هفتگی نسبت به پیش‌بینی مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری است.
۶) پیش‌بینی مدل ترکیبی در قیمت پایانی هفتگی نسبت به پیش‌بینی مدل شبکه‌ی عصبی پیش خور دارای خطای کمتری است.

پیشینه پژوهش

در این بخش بعد از مروری کوتاه بر مبانی نظری مدل‌های سری‌زمانی تک‌متغیره و شبکه‌ی عصبی به بررسی پیشینه تجربی پرداخته می‌شود و در نهایت مدل ترکیبی به‌کار رفته در پژوهش ارائه می‌گردد.

مدل‌های سری زمانی تک متغیره مدل‌هایی هستند که برای پیش‌بینی ارزش یک متغیر تنها از اطلاعات گذشته‌ی همان متغیر و اطلاعات حال و گذشته‌ی اجزای اخلال بهره می‌گیرند.

مدل‌های سری‌زمانی در مقابل مدل‌های ساختاری[۶] مطرح می‌گردند. اما این مدل‌ها نسبت به مدل‌های ساختاری دارای دو مزیت عمده به قرار زیر می‌باشند:

۱) مدل‌های ساختاری یک پیش‌فرض تئوریک را مبنای برقراری رابطه‌ی علت و معلول بین متغیرهای مستقل و وابسته قرار می‌دهند، در حالی که مدل‌های سری زمانی نیازی به این پیش فرض ندارند.
۲) مدل‌های ساختاری در پیش‌بینی خارج از نمونه‌ای چندان خوب عمل نمی‌کنند.

فرآیندهای خودرگرسیو[۷] (AR)

فرآیندهای خودرگرسیو فرآیندهایی هستند که در آن ارزش جاری متغیر X_t تنها به ارزش-های گذشته خود متغیر و جزء اخلال، وابسته می‌باشد. یک فرآیند AR از مرتبه P را به صورت های زیر می‌توان نمایش داد:

$$X_t = \mu + Q_1X_{t-1} + Q_2X_{t-2} + \cdots + Q_pX_{t-p} + U_t \qquad AR(p) \qquad \text{رابطه(۱) :}$$

$$X_t = \mu + \sum_{i=1}^{p} Q_iL^iX_t + U_t$$

$$\begin{cases} Q(L)X_t = \mu + U_t \\ Q(L) = 1 - Q_1L^1 - Q_2L^2 - \cdots - Q_pL^p \end{cases}$$

www.SID.ir

Arshive of SID

L: عملگر وقفه[8]

U_t: یک فرآیند نوفه سفید

یک فرآیند خودرگرسیو زمانی مانا است که $|Q|<1$ باشد.[6]

فرآیندهای میانگین متحرک[9] (MA)

فرآیندهای میانگین متحرک، ترکیب خطی ساده‌ای از فرآیندهای نوفه سفید می‌باشند که در آن ارزش جاری X_t به ارزش‌های جاری و گذشته اجزاء اخلال، وابسته است.[6]

فرآیندهای میانگین متحرک از مرتبه ی q را به صورت های زیر می‌توان نمایش داد:

رابطه (۲): $X_t = \mu + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \ldots + \theta_q U_{t-q} + U_t \qquad \mu A(q)$

$$X_t = \mu + \sum_{j=1}^{q} \theta_j L^j U_t + U_t$$

$$\begin{cases} X_t = \mu + \theta(L)U_t \\ \theta(L) = 1 + \theta_1 L^1 + \theta_2 L^2 + \cdots + \theta_p L^p \end{cases}$$

فرآیندهای خودرگرسیو میانگین متحرک[10] (ARMA)

مدل‌های ARMA(p,q) از ترکیب مدل‌های خودرگرسیو از مرتبه‌ی p و میانگین متحرک از مرتبه‌ی q حاصل می‌شوند. در این مدل‌ها ارزش جاری سری زمانی X_t به صورت خطی به ارزش-های گذشته‌ی خودش و ارزش‌های جاری و گذشته‌ی اجزاء اخلال، وابسته است.

مدل‌ها خودرگرسیو میانگین متحرک را به صورت‌های زیر می‌توان نمایش داد:

رابطه (۳):

$$X_t = \mu + Q_1 x_{t-1} + Q_2 x_{t-2} + .. + Q_p x_{t-p} + \theta_1 U_{t-1} + \ldots + \theta_q U_{t-q} + U_t$$

$$X_t = \mu + \sum_{i=1}^{p} Q_i L^i X_t + \sum_{j=1}^{q} \theta_j L^j U_t + U_t$$

$$\begin{cases} Q(L)X_t = \mu + \theta(L)U_t \\ Q(L) = 1 - Q_1 L^1 - Q_2 L^2 - \cdots - Q_p L^p \\ \theta(L) = 1 + \theta_1 L^1 + \theta_2 L^2 + \cdots + \theta_p L^p \end{cases}$$

www.SID.ir

Arshive of SID

مدل‌های رگرسیو میانگین متحرک، هم خصوصیات مدل‌های خودرگسیو (AR) و هم خصوصیات مدل‌های میانگین متحرک (MA) را در بر می‌گیرند. در مدل‌های خودرگسیو، تابع خودهمبستگی به صورت نمایی کاهش پیدا می‌کند در حالی که تابع خودهمبستگی جزئی تعداد وقفه‌های مورد نیاز برای مدل را معین می‌سازد، اما در مدل‌های میانگین متحرک، این دو تابع برعکس مدل‌های خودرگسیو، عمل می‌نمایند.[6]

مانایی

اگر تابع توزیع یک سری زمانی مانند $X(t)$ از مرتبه ی n ام نسبت به تغییر مبدأ زمانی ناوردا[11] باشد، آن گاه سری زمانی $X(t)$ را یک سری مانای قوی[12] می نامیم.[15]

رابطه (۴): $$F(x_1, x_2, ...x_n; t_1, t_2, ...t_n) = F(x_1, x_2, ...x_n; t_1 + S, t_2 + S, ...t_n + S)$$

اما از آن جایی که در عمل به سختی می‌توانیم نشان دهیم که یک سری زمانی مانای قوی است (به جزء در توزیع نرمال)، بنابراین به جای در نظر گرفتن تابع توزیع مشترک n مرتبه‌ای، تابع توزیع یک یا دو مرتبه‌ای را لحاظ می‌کنیم. بنابراین اگر تابع توزیع یک سری مانند $X(t)$ از مرتبه‌ی اول و یا دوم نسبت به تغییر مبدأ زمانی ناوردا باشد، داریم:

رابطه(۵): $$F(x;t) = F(x;t+S) \quad \forall t, S$$

$$F(x_1, x_2; t_1, t_2) = F(x_1, x_2; t_1 + S, t_2 + S) \quad \forall t_1, t_2, S$$

از معادلات بالا می‌توان نتیجه گرفت که میانگین سری زمانی $X(t)$ در طول زمان ثابت و کوواریانس بین $X(t)$ و $X(t+S)$ تنها به فاصله‌ی زمانی یعنی S وابسته است.

انواع نامانایی

اغلب دو مدل به منظور مشخص کردن نامانایی به کار گرفته می‌شود:

۱. مدل گشت تصادفی با رانش[13]

رابطه (۶): $$X_t = \mu + x_{t-1} + U_t$$

۲. فرآیند روند- مانا[14] (اطراف خط روند مانا می باشند)

www.SID.ir

Arshive of SID

رابطه (۷):

$$X_t = \alpha + \beta T + U_t$$

که مدل اول را نامانای تصادفی[۱۵] و مدل دوم را نامانای معین[۱۶] می نامند.
برای مانا کردن مدل‌های گشت تصادفی، معمولاً از تفاضل[۱۷] و برای مانا کردن مدل‌های روند- مانا معمولاً از روند زدایی استفاده می‌گردد.[6]

مدل‌های خودرگرسیو میانگین متحرک انباشته (ARIMA)[۱۸]

شرط بهره‌گیری از مدل‌های خودرگرسیو میانگین متحرک (ARMA)، مانا بودن سری‌های زمانی می‌باشد. برای مانا کردن سری‌های زمانی مالی و اقتصادی اغلب از روش تفاضل استفاده می-شود. جزء یکپارچگی در مدل خودرگرسیو میانگین متحرک انباشته بیان‌گر میزان تفاضل مورد نیاز برای مانا کردن سری زمانی است.[9]

مدل‌های خودرگرسیو میانگین متحرک انباشته را معمولاً به صورت ARIMA(p,d,q) نمایش می‌دهند که در آن p مرتبه‌ی مورد نیاز برای خودرگرسیو، d میزان تفاضل مورد نیاز برای مانا کردن سری زمانی و q میزان مرتبه‌ی مورد نیاز برای میانگین متحرک است.

مراحل ساخت مدل های ARIMA

۱) مانا کردن سری زمانی: با استفاده از روش تفاضل، به منظور استفاده از مدل‌های خودرگرسیو میانگین متحرک انباشته ، سری زمانی را مانا می‌کنیم.
۲) شناسایی مدل[۱۹]: در این مرحله مرتبه‌های مورد نیاز برای ساخت مدل تعیین می‌گردد. برای این منظور می‌توان از توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) بهره گرفت.
۳) تخمین مدل[۲۰]: تخمین ضرایب با استفاده از تکنیک‌هایی مانند حداقل مربعات، حداکثر درست نمایی و از این قبیل.
۴) بررسی مدل[۲۱]: در این مرحله با استفاده از توابع خودهمبستگی و خودهمبستگی جزئی و آماره‌ی باکس و الژانگ، مستقل بودن سری پسماندها، مورد آزمون قرار می‌گیرد. مدل خودرگرسیو میانگین متحرک انباشته بهینه مدلی است که پسماندهای آن فاقد همبستگی خطی باشند.
۵) پیش‌بینی: با استفاده از مدل خودرگرسیو میانگین متحرک انباشته بهینه، ارزش متغیر در آینده را پیش‌بینی می‌کنیم.

www.SID.ir

Arshive of SID

شبکه های عصبی مصنوعی

شبکه‌های عصبی مصنوعی از عناصر عملیاتی ساده‌ای ساخته می‌شوند که به صورت موازی در کنار یکدیگر عمل می‌کنند. این عناصر که از سیستم‌های عصبی زیستی الهام گرفته شده‌اند، در تلاش‌اند که به صورت ناپارامتریک، مغز انسان را شبیه سازی نمایند. [2]

در شبکه‌ی عصبی مصنوعی به این عناصر پردازشی، نرون‌های مصنوعی می‌گویند که این نرون‌ها مشابه نرون‌های بیولوژیکی مغز انسان می‌باشند.

مغز انسان دارای 10^{11} نرون است که هر کدام از آن‌ها دارای هزاران ارتباط با یکدیگر می باشند. هر نرون از ۳ بخش اصلی تشکیل می شود که عبارتند از: دندریت[۲۲]، هسته[۲۳] و آکسون[۲۴] نرون‌ها از طریق دندریت‌ها سیگنال‌ها را از سایر نرون‌ها دریافت کرده و از طریق آکسون‌ها به دیگر نرون‌ها ارسال می‌کنند. به ارتباط میان آکسون و دندریت، سیناپس[۲۵] گفته می‌شود. در واقع، ورودی یک شبکه‌ی عصبی مصنوعی، مشابه ایمپلاسهای الکتروشیمیایی است که توسط دندریت-های نرون‌های بیولوژیکی دریافت می‌گردد و خروجی نرون‌های مصنوعی مشابه سیگنال‌های خروجی از آکسون یک نرون بیولوژیکی است.

شبکه‌های پس‌انتشار

شبکه پس‌انتشار، که به اختصار آن را با BP نمایش می‌دهیم، یک شبکه چند لایه با تابع انتقال غیرخطی و قاعده یادگیری Widrow-Hoff می‌باشد. از بردار ورودی و هدف در راستای آموزش این نوع شبکه، برای تقریب زدن یک تابع، یافتن رابطه‌ی بین ورودی و خروجی و دسته‌بندی ورودی‌ها استفاده می‌گردد.

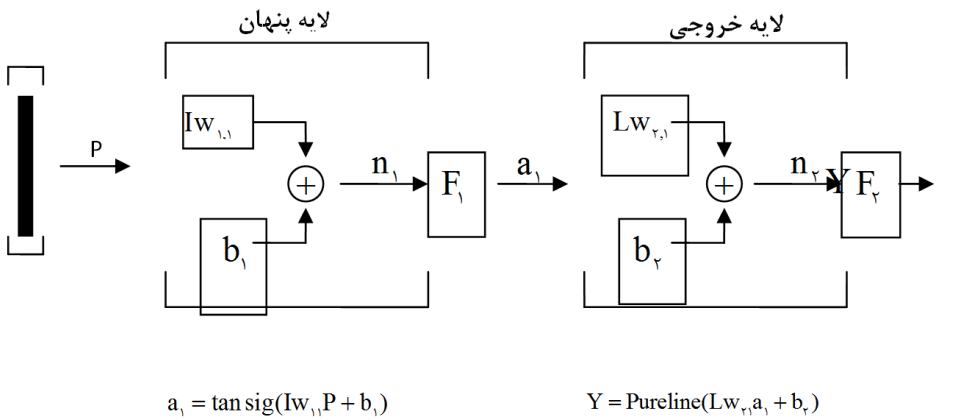
یکی از مرسوم‌ترین معماری‌های شبکه‌های BP، شبکه‌های چند لایه پیش‌خور می‌باشند[۲۶] که به منظور تخمین مدل، طبقه‌بندی و مدل‌سازی به طور گسترده مورد استفاده قرار می‌گیرد. [13]

شبکه‌های پیش‌خور اغلب دارای یک یا چند لایه‌ی پنهان از نرون‌های sigmoid بوده و در لایه‌ی خروجی خود از تابع فعال سازی خطی[۲۷] استفاده می‌کنند.

شبکه‌ی پیش‌خور (ff) را می‌توان به صورت $I\times H\times O$ بیان نمود به طوری که I میزان نرون‌ها در لایه‌ی ورودی و H میزان نرون‌ها در لایه‌ی پنهان و O میزان نرون‌ها در لایه‌ی خروجی را نشان می‌دهد. شکل زیر یک شبکه‌ی پیش‌خور را به نمایش می‌کشد.

www.SID.ir

Arshive of SID



شکل ۱- شبکه‌ی عصبی پیش‌خور با تابع فعال سازی تانژانت هیپربولیک

مطالعات زیادی در زمینه‌ی مقایسه‌ی روش‌های خطی از قبیل مدل خودرگرسیو میانگین متحرک انباشته و غیرخطی از قبیل شبکه‌های عصبی انجام پذیرفته است. آن‌چه از این مطالعات بر می‌آید، وجود نتایج متفاوت در استفاده از شبکه‌های عصبی است برای مثال:

شاردا و پاتیل (۱۹۹۰) در مقاله‌ای از ۷۵ سری زمانی برگرفته از مقاله‌ی به منظور مقایسه عملکرد شبکه‌ی عصبی و خودرگرسیو میانگین متحرک انباشته (ARIMA) بهره جستند که در ۳۹ سری زمانی عملکرد شبکه، و در ۳۶ سری زمانی، عملکرد ARIMA دقیق‌تر بود.[18]

فاستر و همکاران (۱۹۹۲) به مقایسه‌ی عملکرد شبکه‌ی عصبی با مدل رگرسیون خطی و میانگین شش مدل هموارسازی نمایی در ۳۸۴ سری زمانی اقتصادی و جمعیت‌شناختی پرداختند. نتایج نشان داد که استفاده از رگرسیون خطی به طور معناداری از شبکه‌ی عصبی به عنوان تخمین زننده، در این سری های تصادفی، بهتر است. [11]

تانگ و فیش ویچ (۱۹۹۳) در مقاله‌ی خود اذعان داشتند، عملکرد شبکه‌ی به ماهیت داده‌ها وابسته می‌باشد و برای سری‌های زمانی بی‌قاعده و پیش‌بینی چند مرحله‌ای، این مدل نسبت به مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری است.[19]

ایندرو و همکاران (۱۹۹۹) اذعان داشتند که، عملکرد شبکه عصبی نسبت به رگرسیون خطی، در صندوق‌های مشترک سرمایه‌گذاری به نوع صندوق وابسته است.[12]

www.SID.ir

Arshive of SID

کلین و راسین (۱۹۹۹) بیان داشتند که، عملکرد شبکه‌ی عصبی وابسته به ماهیت داده‌ها می-باشد و کاهش نرخ خطا و مقدار خطا می‌تواند منجر به بهبود عملکرد شبکه‌ی عصبی در پیش‌بینی خارج از نمونه‌ای گردد.[14]

رضا روشن (۱۳۸۵) در پایان‌نامه‌ی خود توان پیش‌بینی تورم مدل‌های خودرگرسیو میانگین متحرک انباشته، ناهمسانی واریانس شرطی و شبکه‌ی عصبی را با هم مقایسه کرده و نتایج بر کارآتر بودن مدل ناهمسانی واریانس شرطی نسبت به مدل خودرگرسیو میانگین متحرک انباشته و مدل شبکه‌های عصبی مصنوعی نسبت به رقبای خود دلالت دارد.[3]

متوسلی و طالب (۱۳۸۵) نتیجه گرفتند که عملکرد مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) به طور معناداری نسبت به شبکه‌های عصبی بهتر است و در واقع مدل خطی خودرگرسیو میانگین متحرک انباشته بهتر از مدل‌های غیرخطی (شبکه‌های عصبی) می‌تواند پیچیدگی سری زمانی قیمت سهام را تجزیه و تحلیل نماید.[5]

بنابراین با توجه به عملکرد متفاوت مدل‌های خطی و غیر خطی در سری‌های زمانی مختلف در بسیاری از مطالعات به بررسی توان پیش‌بینی مدل‌های ترکیبی پرداخته شده است برای نمونه ژانگ (۲۰۰۳) در مقاله‌ای خود به تلفیق ARIMA و شبکه‌ی عصبی پرداخت که نتایج این پژوهش دال بر موفقیت مدل ترکیبی نسبت به مدل‌های ARIMA و شبکه‌ی عصبی می‌باشد. [23]

ودینگ و سیوس (۱۹۹۶) نتیجه گرفتند، اگر پیش‌بینی سری زمانی و فاکتورهای اطمینان که توسط شبکه‌ی عصبی RBF تولید شده است را با مدل باکس و جنکینز ترکیب نماییم، این امر موجب افزایش دقت پیش‌بینی می‌شود.[21]

تی سنگ و همکاران (۲۰۰۲) در مقاله‌ای سعی در ترکیب شبکه‌ی عصبی پس‌انتشار و مدل-های سری زمانی فصلی (SARIMA[۲۸]) نمودند و آن را با (SARIMABP) نمایش دادند. نتایج پژوهش نشان داد که عملکرد مدل SARIMABP نسبت به مدل خودرگرسیو میانگین متحرک انباشته‌ی فصلی و شبکه‌های عصبی با داده‌های تفاضلی و فصل زدا شده به طور معناداری بهتر است.[20]

لیو و پای (۲۰۰۵) برای سری زمانی ده سهم بورس نیویورک، سعی در ترکیب مدل خطی خودرگرسیو میانگین متحرک انباشته و مدل غیر خطی ماشین بردار پشتیبان[۲۹] نمودند و نتایج نشان داد که عملکرد مدل ترکیبی به طور معناداری از سایر مدل‌ها بهتر است.[17]

فاروق (۲۰۱۰) در مقاله خود از سه سری زمانی دمای آب، ماده‌ی بورون (Boron) و اکسیژن حل شده در طول سال‌های ۲۰۰۴-۱۹۹۶ در رودخانه‌ی بویاک مندرز[۳۰] ترکیه برای مقایسه توان پیش‌بینی ترکیب مدل خطی خودرگرسیو میانگین متحرک انباشته‌ی فصلی و شبکه‌ی عصبی

www.SID.ir

Arshive of SID

نسبت به هر یک از مدل‌ها مذکور بهره جست و نتایج دال بر موفقیت مدل ترکیبی نسبت به دو مدل دیگر می‌باشد.[10]

۳- مدل مفهومی

مدل ترکیبی به کار گرفته شده در این پژوهش به این صورت است ابتدا هر سری زمانی مانند Y_t به دو بخش خطی و غیرخطی به صورت زیر تقسیم می‌گردد:

رابطه(۸):

$$Y_t = L_t + N_t$$

L_t: جزء خطی

N_t: جزء غیرخطی

آنگاه با استفاده از مدل خودرگرسیو میانگین متحرک انباشته بخش خطی سری زمانی Y_t پیش-بینی گردیده و پسماندهای حاصل از این پیش‌بینی به قرار زیر محاسبه می‌شود:

رابطه(۹):

$$Y_t - \hat{L}_t = e_t$$

e_t: پسماندهای حاصل از پیش‌بینی مدل ARIMA

$\hat{L}_t$: پیش‌بینی مدل ARIMA

سپس با استفاده از شبکه‌ی عصبی، بخش غیرخطی به صورت زیر محاسبه می‌گردد:

رابطه(۱۰):

$$e_t = f(e_{t-1}, e_{t-2}, ..., e_{t-n}) + \varepsilon_t$$

f: تابع غیرخطی برازش شده توسط شبکه‌ی عصبی

ε_t: جزء اخلال

در نهایت با افزودن پیش‌بینی شبکه‌ی عصبی به پیش‌بینی مدل خطی خودرگرسیو میانگین متحرک انباشته، پیش‌بینی مدل ترکیبی برای سری زمانی Y_t حاصل خواهد شد. در واقع داریم:

رابطه(۱۱):

$$\hat{e}_t = \hat{N}_t$$

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t$$

$\hat{Y}_t$: پیش‌بینی مدل ترکیبی

www.SID.ir

Arshive of SID

$\hat{N}_t$: پیش‌بینی پسماندها توسط شبکه

۴- روش‌شناسی پژوهش

در این پژوهش به‌منظور مقایسه عملکرد مدل ترکیبی ارائه شده در قسمت قبل با هر یک از مدل‌های ARIMA و شبکه‌ی عصبی از سری‌های زمانی حدبالای قیمت هفتگی اونس طلا ، حد پایین قیمت هفتگی اونس طلا و قیمت پایانی هفتگی اونس طلا از ابتدای سال ۲۰۰۶ تا اکتبر ۲۰۱۱ بهره گرفته شده است.

از آن‌جایی که یکی از روش‌های مورد استفاده در محاسبه‌ی ارزش فاصله‌ای[۳۱]، بهره‌گیری از دامنه‌ی تغییرات[۳۲] و نقطه وسط[۳۳] می‌باشد؛ بنابراین، از بالاترین و پایین‌ترین قیمت طلا در طول یک هفته استفاده شده و دو سری زمانی دامنه‌ی تغییرات و نقطه‌ی وسط برای هر هفته به قرار زیر محاسبه گردیده است:

رابطه(۱۲):

$$x_t^c = \frac{x_t^u + x_t^l}{۲}$$

رابطه(۱۳):

$$x_t^r = \frac{x_t^u - x_t^l}{۲}$$

x_t^c: نقطه وسط

x_t^r: دامنه‌ی تغییرات

x_t^u: بالاترین قیمت هفتگی (حد بالا)

X_t^l: پایین‌ترین قیمت هفتگی (حد پایین)

همان‌طور که مشخص است، جمع سری‌های زمانی دامنه‌ی تغییرات و نقطه‌ی وسط بیان‌گر حد بالای قیمت (بالاترین قیمت) در طول یک هفته و تفاضل این دو سری زمانی بیان‌گر حد پایین قیمت (پایین‌ترین قیمت) در طول یک هفته می‌باشد. بنابراین از سری‌های زمانی دامنه‌ی تغییرات و نقطه‌ی وسط به عنوان متغیرهای اصلی برای پیش‌بینی دامنه‌ی نوسان طلا و از سری زمانی قیمت پایانی به عنوان متغیر اصلی برای پیش‌بینی قیمت پایانی در هر یک از مدل‌های خودرگرسیو میانگین متحرک انباشته (ARIMA)، شبکه‌ی عصبی و مدل ترکیبی استفاده شده است.

بعد از مانا کردن سری‌های زمانی نقطه‌ی وسط، دامنه‌ی تغییرات و قیمت پایانی (از طریق محاسبه‌ی بازده‌های لگاریتمی)، ابتدا با ۲۷۳ مشاهده از سری‌های مربوطه مدل خودرگرسیو

www.SID.ir

Arshive of SID

میانگین متحرک انباشته برازش شده و معیارهای میانگین مجذور خطا و میانگین قدر مطلق خطا برای مشاهده ۲۷۴ام از حد بالای قیمت، حد پایین قیمت و قیمت پایانی به روش زیر محاسبه گردیده است:

رابطه(۱۴):

$$MSE = \frac{\sum_{i=1}^{n}(Y_{R_i} - Y_{F_i})^2}{n}$$

رابطه(۱۵):

$$MAE = \frac{\sum_{i=1}^{n}\left|Y_{R_i} - Y_{F_i}\right|}{n}$$

Y_R: مقدار واقعی

Y_F: مقدار پیش‌بینی

n : تعداد اجزاء خارج نمونه‌ای

سپس بعد از حذف یک مشاهده از بالا، مشاهده ۲۷۴ام واقعی از سری‌های زمانی مربوطه، به منظور برازش مجدد مدل خودرگرسیو میانگین متحرک انباشته و معیارهای ارزیابی مذکور برای مشاهده ۲۷۵ام از حد بالای قیمت، حد پایین قیمت و قمت پایانی به مدل اضافه شده و فرایند فوق برای ۳۰ مرتبه انجام پذیرفته است.

بعد از برازش هر مرتبه از مدل خودرگرسیو میانگین متحرک انباشته (ARIMA)، بر طبق مدل ترکیبی مطرح شده در زیر بخش مدل مفهومی، با استفاده از شبکه‌ی عصبی، الگوی غیرخطیِ پسماندهای حاصل از ARIMA مدل‌سازی گردیده، آن‌گاه پیش‌بینی پسماندها برای یک مرحله پیش‌رو به پیش‌بینی مدل خطی خودرگرسیو میانگین متحرک انباشته (ARIMA) به منظور بررسی دقت پیش‌بینی مدل ترکیبی در حد بالای قیمت، حد پایین قیمت و قیمت پایانی اضافه گردیده است.

به‌منظور بررسی دقت مدل ترکیبی نسبت به دو مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) و شبکه‌ی عصبی بعد از برازش هر مرتبه از مدل ARIMA، بوسیله‌ی بهترین شبکه‌ی عصبی برازش شده، دقت پیش‌بینی مدل شبکه‌ی عصبی پیش‌خور در حد بالای قیمت، حد پایین قیمت و قیمت پایانی مورد ارزیابی قرار داده شده است.

در برازش شبکه‌ی عصبی، چه در برازش خود قیمت‌ها و چه در مدل سازی روابط غیر خطی در پسماندها، از کد نویسی در نرم افزار متلب و تغییر در تعداد وقفه‌های زمانی، لایه‌های پنهان و

www.SID.ir

Arshive of SID

نرون‌های لایه‌های پنهان بهره گرفته شده و شبکه‌ای که در بخش تست (۱۰ درصدرسری‌زمانی) کمترین خطا را دارا می‌باشد مبنای پیش‌بینی یک مرحله پیش‌رو قرار گرفته است.

در نهایت، معناداری تفاوت میانگین بین دو جامعه میانگین مجذور خطا و میانگین قدر مطلق خطا برای بررسی فرضیه‌های پژوهش، با استفاده از آزمون مقایسه زوجی مورد بررسی واقع گردیده است.

اما از آن‌جا که در پیش بینی های سری زمانی برای H مرحله پیش‌رو از داده های روی هم افتاده استفاده می نماییم، خطای پیش بینی ${\varepsilon_{t+H|t}}^1$ و ${\varepsilon_{t+H|t}}^2$ (مربوط به مدل‌های ۱ و ۲) به صورت سریالی همبستگی خواهد داشت. از این رو در آزمون مقایسه میانگین دو جامعه میانگین مجذور خطا علاوه بر آزمون مقایسه زوجی از آزمون دایبولد- ماریانو[۳۴] و آماره تعدیل شده آن توسط هاروی – لیبورن- نیوبولد (برای نمونه‌های کوچک)، برای بررسی معناداری دقت پیش‌بینی مدل ترکیبی نسبت به هریک از مدل‌های ARIMA و شبکه‌ی عصبی، بهره گرفته شده است.

آزمون دایبولد-ماریانو به‌صورت زیر قابل تعریف است:

رابطه (۱۶):

$$d_t = L\left(\varepsilon^1_{t+H|t}\right) - L\left(\varepsilon^2_{t+H|t}\right)$$

رابطه (۱۷):

$$S = \frac{\bar{d}}{\left(\widehat{avar}(\bar{d})\right)^{1/2}} = \frac{\bar{d}}{\left(\widehat{LRV}_{\bar{d}}/T\right)^{1/2}}$$

رابطه (۱۸):

$$LRV_{\bar{d}} = \gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j \ , \qquad \gamma_j = cov(d_t, d_{t-j})$$

$LRV_{\bar{d}}$، تخمین سازگاری از واریانس مجانبی $\sqrt{T}\ \bar{d}$ است. دایبولد و ماریانو (۱۹۹۵) نشان می‌دهند که تحت فرضیه قدرت پیش‌بینی کنندگی یکسان داریم:

$S \sim N(0,1)$

بنابراین فرضیه قدرت پیش بینی یکسان را در سطح ۹۵٪ زمانی رد می کنیم که داشته باشیم:

$|S| > 1/96$

اما این آماره سال‌های بعدها توسط هاروی – لیبورن- نیوبولد برای نمونه‌های کوچک و تابع زیان خاص مجذور خطا به صورت زیر اصلاح گردید:

www.SID.ir

Arshive of SID

رابطه (۱۹):

$$S^* = \frac{\bar{d}}{\left(\widehat{var}(\bar{d})\right)^{1/2}} = \frac{\bar{d}}{\sqrt{\widehat{var}_m}}$$

رابطه (۲۰):

$$\widehat{var}_m = [N + 1 - 2h + (n^{-1} * h(h-1))]^{-1} \times \widehat{\gamma_0} + 2\sum_{j=1}^{h-1} \widehat{\gamma_j}$$

۵- یافته‌های پژوهش

همانطور که در بخش روش پژوهش ذکر شد، به‌منظور بررسی فرضیه‌های تحقیق مبنی بر دقت مدل ترکیبی نسبت به هر یک از مدل‌های ARIMA و شبکه‌ی عصبی در حد بالای قیمت، حد پایین قیمت و قیمت پایانی از آزمون مقایسه زوجی، آزمون دایبولد-ماریانو و آماره‌ی تعدیل شده هاروی – لیبورن- نیوبولد و روش غلتاندن داده‌ها بهره گرفته شده است. در این بخش ابتدا به بررسی آمار توصیفی برای هر یک از سه سری زمانی حد بالای قیمت، حد پایین قیمت و قیمت پایانی پرداخته شده و از آنجایی که یکی از شروط استفاده از مدل ARIMA مانا بودن سری‌های زمانی است نتایج حاصل از آزمون دیکی-فولر برای بازده‌های لگاریتمی سری‌های زمانی حد بالا قیمت، حد پایین قیمت و قیمت پایانی انس طلا به تفصیل ارائه می‌گردد. در نهایت نتایج حاصل از مقایسه سه مدل بر اساس آزمون‌های مذکور و برای هر یک از توابع زیان که در بخش‌های قبل به تفصیل مورد بررسی قرار گرفتند، در قالب جداولی مورد بررسی قرار داده می‌شوند.

جدول شماره ۱ نتایج حاصل از آمار توصیفی برای هر یک از سه سری‌زمانی را به‌صورت مجزا تبین می‌نماید.

اما همانطور که ذکر شد یکی از مهم‌ترین ویژگی در برازش مدل ARIMA مانا بودن سری‌های زمانی است. سری‌های زمانی قیمت در فضای مالی و اقتصادی اکثرا مانا نبوده و دارای گشت تصادفی می‌باشند. در واقع از منظر اقتصاد سنجی سری‌های زمانی مالی و اقتصادی اکثرا هم انباشته از مرتبه اول هستند.از این رو این ایده به ذهن می‌رسد که تفاضل از مرتبه اول می‌تواند منجر به مانایی سری‌های زمانی و در نهایت هم انباشتگی از مرتبه صفر شود. بازده در واقع نوعی تفاضل از مرتبه اول است. بنابراین نتایج حاصل از آزمون دیکی-فولر برای بازده لگاریتمی هر یک سه سری زمانی حد بالا قیمت، حد پایین قیمت و قیمت پایانی در قالب جدول شماره ۲، ۳ و ۴ به تفصیل ارائه شده است.

www.SID.ir

Arshive of SID

جدول ۱- آمار توصیفی

	حد بالا قیمت	حد پایین قیمت	قیمت پایانی
میانگین	۱۰۳۲,۱۵۹	۹۰۳,۴۱۰۳	۱۰۱۲,۹۵۲
میانه	۹۴۲,۶۵۰۰	۹۰۳,۳۵۰۰	۹۲۵,۸۵۰۰
ماکسیمم	۱۹۲۰,۷۸۰	۱۷۹۳,۱۰۰	۱۸۸۵,۳۶۰
مینیمم	۵۵۴,۴۰۰۰	۵۳۳,۵۰۰	۵۴۰,۶۰۰۰
انحراف معیاد	۳۵۸,۲۰۶۵	۳۴۴,۹۵۰۳	۳۵۳,۶۰۱۲
چولگی	۰,۶۶۹۴۸۸	۰,۶۳۶۴۱۳	۰,۶۶۲۸۹۱
کشیدگی	۲,۴۴۵۶۰۱	۲,۲۹۲۲۳۰	۲,۳۹۲۵۰۴
جارک-برا	۲۸,۰۹۰۳	۲۸,۳۶۸۷۳	۲۸,۴۴۵۲۴
سطح معناداری	۰,۰۰۰۰۰۱	۰,۰۰۰۰۰۱	۰,۰۰۰۰۰۱
جمع	۳۳۱۳۲۳,۱	۳۱۶۹۵۸,۷	۳۲۵۱۵۷,۵
تعداد مشاهدات	۳۲۱	۳۲۱	۳۲۱

جدول ۲- نتایج آزمون دیکی فولر برای بازده لگاریتمی حد بالا قیمت

Null Hypothesis: D(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-13.34955	0.0000
Test critical values:	1% level	-3.988840	
	5% level	-3.424825	
	10% level	-3.135494	

*MacKinnon (1996) one-sided p-values.

www.SID.ir

Arshive of SID

جدول ۳- نتایج آزمون دیکی فولر برای بازده لگاریتمی حد پایین قیمت

Null Hypothesis: D(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-15.31724	0.0000
Test critical values:	1% level	-3.988840	
	5% level	-3.424825	
	10% level	-3.135494	

*MacKinnon (1996) one-sided p-values.

جدول ۴- نتایج آزمون دیکی فولر برای بازده لگاریتمی قیمت پایانی

Null Hypothesis: D(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-17.34042	0.0000
Test critical values:	1% level	-3.988840	
	5% level	-3.424825	
	10% level	-3.135494	

*MacKinnon (1996) one-sided p-values.

نتایج جداول ۲، ۳ و ۴ نشان می‌دهد که بازده لگاریتمی در تمامی سطوح خطا ۱ درصد، ۵ درصد و ۱۰ درصد مانا است. از این رو در این پژوهش از بازده لگاریتمی این سری‌های زمانی استفاده شده است.

اما هدف نهایی این پژوهش مقایسه عملکرد مدل ترکیبی در مقابل هر یک از مدل‌های خطی و غیر خطی به‌صورت مجزا است از این رو جدول ۵ نتایج آزمون مقایسه زوجی با تابع زیان میانگین مجذور خطا را نشان می‌دهد که اندیسup ، Down و Close به ترتیب بیانگر حدبالای قیمت، حد پایین قیمت و قیکت پایانی اونس طلا می‌باشند :

www.SID.ir

Arshive of SID

جدول ۵- آزمون مقایسه‌ی زوجی (میانگین مجذور خطا)

مقایسه زوجی مجذور خطا	Mean	Std. Deviation	t	Df	Sig. (2-tailed)
$Hybrid_{up}$- $ARIMA_{up}$	-540/96	764/08	-3/87	29	.001
$Hybrid_{up} - ANN_{up}$	-708/8۰	1472/86	-2/63	29	.013
$Hybrid_{down} - ARIMA_{down}$	-449/3۰	783/51	-3/14	29	.004
$Hybrid_{down} - ANN_{down}$	-947/5	1976/24	-2/62	29	.014
$Hybrid_{close} - ARIMA_{close}$	-1001/50	1383/79	-3/96	29	.000
$Hybrid_{close}$–ANN_{close}	-571/4	1183/38	-2/64	29	.013

با توجه به جدول فوق بین میانگین دو جامعه‌ی MSE از مدل ترکیبی و مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) و دو جامعه‌ی MSE از مدل ترکیبی و مدل شبکه‌ی عصبی اختلاف معناداری وجود دارد و میانگین MSE برای مدل ترکیبی در نمونه‌ی مورد بررسی از دو مدل دیگر کمتر است. بنابراین، از نظر معیار ارزیابی MSE ، عملکرد مدل ترکیبی به طور معناداری در حد بالای قیمت، حد پایین قیمت و قیمت پایانی از دو مدل خودرگرسیو میانگین متحرک انباشته و شبکه‌ی عصبی (ANN) بهتر است. جدول ۶ نتایج آزمون مقایسه زوجی با تابع زیان میانگین قدر مطلق خطا را نشان می‌دهد.

جدول ۶- آزمون مقایسه‌ی زوجی (میانگین قدر مطلق خطا)

مقایسه ی زوجی قدر مطلق خطا	Mean	Std. Deviation	Std. Error Mean	T	Df	Sig(2-tailed)
$Hybrid_{close}$–ANN_{close}	-10/99	13/95	2/54	-4/3	29	۰/000
$Hybrid_{close}$–$ARIMA_{close}$	-10/72	8/03	1/46713	-7/3	29	۰/000
$Hybrid_{down}$–ANN_{down}	-11/81	13/56	2/47	-4/7	29	۰/000
$Hybrid_{down}$-$ARIMA_{down}$	-5/81	7/47	1/36	-4/2	29	۰/000
$Hybrid_{up} - ANN_{up}$	-10/25	15/67	2/8۶	-3/5	29	۰/001
$Hybrid_{up}$ –$ARIMA_{up}$	-7/21	8/45	1/54	-4/6	29	۰/000

با توجه به جدول فوق بین میانگین دو جامعه‌ی MAE از مدل ترکیبی و مدل خودرگرسیو میانگین متحرک انباشته و دو جامعه‌ی MAE از مدل ترکیبی و مدل شبکه‌ی عصبی اختلاف معنا-داری وجود دارد و میانگین MAE برای مدل ترکیبی در نمونه‌ی مورد بررسی از دو مدل دیگر کمتر است. بنابراین، از نظر معیار ارزیابی MAE، عملکرد مدل ترکیبی به طور معناداری در حد

www.SID.ir

Arshive of SID

بالای قیمت، حد پایین قیمت و قیمت پایانی از دو مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) و شبکه‌ی عصبی (ANN) بهتر است.

جدول ۷ نتایج آزمون دایبولد- ماریانو و آماره‌ی تعدیل شده هاروی – لیبورن- نیوبولد را با تابع زیان میانگین مجذور خطا نشان می‌دهد.

جدول ۷- آزمون دایبولد- ماریانو و آماره‌ی تعدیل شده آن

	Mean	دایبولد- ماریاونو	هاروی-نیوبولد- لیبورن
$Hybrid_{close}–ANN_{close}$	-540.96	-2.6517	-2.6071
$Hybrid_{close}–ARIMA_{close}$	-708.87	-2.3413	-2.3019
$Hybrid_{down}–ANN_{down}$	-449.34	-2.6262	-2.5820
$Hybrid_{down}–ARIMA_{down}$	-947.55	-3.1412	-3.0884
$Hybrid_{up} – ANN_{up}$	-1001.5	-2.6362	-2.5918
$Hybrid_{up} –ARIMA_{up}$	-571.44	-3.8778	-3.8126

با مقایسه آماره‌ی محاسبه شده در جدول۳ با مقادیر بحرانی جداول Z برای آزمون دایبولد- ماریانو و t-student برای آماره‌ی تعدیل شده هاروی – لیبورن- نیوبولد می‌توان نتیجه گرفت که، عملکرد مدل تلفیقی نسبت به هر یک از مدل‌های خودرگرسیو میانگین متحرک انباشته و شبکه‌ی عصبی در حد بالای قیمت، حد پایین قیمت و قیمت پایانی با توجه به معیار MSE دقیق‌تر است.

۶- نتیجه‌گیری و بحث

از آن‌جا که تشخیص الگوی داده‌ها مبنی بر خطی و غیرخطی بودن در واقعیت کمی دشوار است، بنابراین، باوجود انعطاف‌پذیری بالای مدل‌های خودرگرسیو میانگین متحرک انباشته (ARIMA) و شبکه‌ی عصبی در مدل‌سازی مسائل مختلف، نمی‌توان ادعا کرد که هر یک از این دو مدل برای تمام سری‌های زمانی مناسب می‌باشند.

در این پژوهش فرض شد که ترکیب مدل خودرگرسیو میانگین متحرک انباشته و مدل شبکه‌ی عصبی با در نظر گرفتن خصوصیات خطی ARIMA و غیرخطی شبکه‌ی عصبی می‌تواند منجر به افزایش دقت پیش‌بینی گردد. بنابراین در این پژوهش با استفاده حد بالا قیمت، حد پایین قیمت انس طلا در ابتدا به محسابه دو شاخص دامنه تغییرات قیمت و نقطه وسط که در بخش‌های قبل به تفصیل درباره این دو شاخص و علت استفاده از آن‌ها بحث شده است، پرداختیم. سپس سپس با

www.SID.ir

Arshive of SID

بهره‌گیری از ۲۷۳ مشاهده هفتگی به برازش مدل آریمای بهینه و پیش بینی بازده ۲۷۴ ام برای هر یک از سه سری زمانی بازده لگاریتمی دامنه تغییرات، حد وسط و قیمت پایانی پرداخته شد و نتایج به عنوان بازده ناشی از عملکرد مدل خطی لحاظ گردید. سپس با استفاده از شبکه‌های عصبی علاوه بر پیش‌بینی هر یک از سه سری زمانی مذکور، پسماندهای ناشی از مدل آریما بهینه مدل سازی و بازده‌های هر یک از سری زمانی مذکور و پسماند ناشی از مدل آریما برای هفته ۲۷۴ پیش‌بینی گردید. با استفاده از پسماند پیش بینی شده بازده پیش بینی شده برای مدل ترکیبی نیز برای هفته پیش رو محسابه گردید. این فرایند با استفاده از روش غلتاندن برای تمامی داده‌های خارج از نمونه برای هر سه مدل خطی، غیر خطی و ترکیبی تکرار گردید. سپس از طریق بازده های پیش بینی شده و روش تبدیل شاخص‌های دامنه تغییرات و نقطه وسط به حد بالا و پایین قیمت، قیمت‌های خارج از نمونه برای حد بالا ، حد پایین و قیمت پایانی تخمین زده شد. با استفاده از قیمت‌های تخمین زده شده و مقایسه این قیمت‌ها با مقادیر واقعی توابع زیان محاسبه شده و در نهایت با استفاده سه روش آزمون مقایسه زوجی، دایبولد-ماریانو و دایبولد-ماریانوی تعدیل شده به آزمون فرضیه‌های تحقیق دال بر مقایسه عملکرد مدل ترکیبی نسبت بر هر یک از مدل‌های آریما و شبکه عصبی چه در حد بالا قیمت، حد پایین قیمت و قیمت پایانی پرداخته شد.

در پیش‌بینی دامنه‌ی نوسان هفتگی اونس طلا نسبت به هر یک از مدل‌های ARIMA و شبکه‌ی عصبی با توجه معیارهای ارزیابی MSE و MAE نتایج دال بر تائید فرضیه‌های پژوهش مبنی بر عملکرد بهتری مدل ترکیبی نسبت به دو مدل دیگر در حدبالای قیمت، حدپایین قیمت و قیمت پایانی است.

اما برای پژوهش‌های آتی، با توجه در این پژوهش به منظور مدل‌سازی الگوی غیرخطی پسماند ها در مدل ترکیبی از شبکه‌ی عصبی پیش‌خور استفاده شده است، از این رو می‌توان به‌جای استفاده از شبکه‌ی عصبی از SVM برای مدل‌سازی بخش غیرخطی استفاده کرد و در نهایت عملکرد این مدل ترکیبی را با هر یک از مدل ARIMA و SVM مقایسه نمود.

www.SID.ir

Arshive of SID

فهرست منابع

* احمدی علی، ذوالفقاری مهدی، نژاد مهربانی آیدین، (۱۳۸۸)، مطالعه‌ی تطبیقی روش‌های ARIMA و شبکه‌ی عصبی مصنوعی در پیش‌بینی نیازهای داخلی برق کشور، فصل نامه ی پژوهش های اقتصادی ایران،۱۳(۴۱)، ۱۰۷-۱۲۱
* راعی رضا، (۱۳۷۷)، طراحی مدل سرمایه‌گذاری مناسب در سبد سهام با استفاده از شبکه‌ی عصبی مصنوعی، رساله ی دکتری، دانشکده ی مدیریت
* روشن رضا، (۱۳۸۵)، پیش‌بینی تورم ایران به کمک مدل‌های،ARIMA، ARCH/GARACH، شبکه‌های مصنوعی و مقایسه‌ی کارآیی مدل‌های مذکور. پایا نامه ی کار شناسی ارشد رشته ی علوم اقتصادی،دانشگاه سیستان و بلوچستان
* صادقی حسین، ذوالفقاری مهدی، الهامی نژاد مجتبی، (۱۳۹۰)، مقایسه‌ی عملکرد شبکه‌های عصبی و مدل ARIMA در مدل‌سازی و پیش‌بینی کوتاه‌مدت قیمت سبد نفت خام اوپک(با تاکید بر انتطارات تطبیقی)، فصل نامه ی مطالعات اقتصاد انرژی، ۸ (۲۸)، ۲۵-۴۷
* متوسلی محمود، طالب کاشفی بیژن، (۱۳۸۵)، بررسی مقایسه‌ای توان شبکه‌ی عصبی با ورودی شاخص‌های تحلیل تکنیکی برای پیش‌بینی قیمت سهام، نامه مفید، ۱۲(۵۴)، ۵۷-۸۲
* Brooks, c. (2008). Introductory Econometrics for Finance.
* Cadenas, E & River, W. (2010). Wind speed forecasting in three different region of mexico,using a hybrid ARIMA-ANN model. Renewable Enegy, 35(12), 2732-2738.
* Chen, K & Wang, C. (2007). A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. Expert System with Applications, 32(1), 254-264.
* Enders, W. (2009) Applied Econometric Time Series. Wiley
* Faruk, D. (2010) Ahybrid neural network and ARIMA model for water quality time series prediction. Engineering Application of Artificial Intligence, 23(4), 586-594
* Foster, W & Collopy, F & Ungar, L. (1992). Neural network forecasting of short, noisy time series. Computers and Chemical Engineering, 16(4), 293-297
* Indro, D & Jiang, C & Patuwo, B & Zhang, G. (1999). Predicting mutual fund performance using artificial neural networks. Omega The International Journal of Management Science, 27(3), 373-380.
* Kaastra, I & Boyd, M. (1996). Designing a neural network for forecasting financing and economic time-series. Neurocomputing, 10(3), 215-236.
* Klein, B & Rossin, D. (1999). Data quality in neural network models: effect of error rate and magnitude of error on predictive accuracy. Omega, The International Journal of ManagementScience, 27(5), 569-582.
* Lefebvre, M . (2007) Applied Stochastic processes.

www.SID.ir

Arshive of SID

* Markham, I & Rakes, T. (1998). The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression. ComputOper Res, 25(4), 251-263.
* Pai, P & Liu, C. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. Omega, 33(6), 497-505.
* Sharda, R & Patill, R. (1990) Neural networks as forecasting exports: an empirical test. Proceedings of IJCNN Meeting, 2(1990), 491-494.
* Tang, Z & Fishwich, P. (1993). feed-forward neural nets as models for time series forcasting. ORSA Journal on Computing, 5(4), 374-385.
* Tseng, F & Yu, H & Tzeng, G. (2002). Combining neural network model with seasonal time series ARIMA model. Technol Forecast Soc Change , 69(1), 71-87.
* Wedding, D & Cios, K. (1996). Time series forecasting by combining RBF networks certainty factors and the Box-Jenkins model. Neurocomputing , 10(2), 149-168.
* Zhang, G & Patuwo, E & Hu, M.(1998) Forecasting with artifcial neural networks: the state of the art. Int J Forecasting, 14(1), 35-62.
* Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network Model. Neurocomputing , 50, 159-175

یادداشت‌ها

[1]. Interval Analysis
[2]. Box-Jenkins
[3]. Universal
[4]. Artificial Neural Network
[5]. Noise
[6]. Structural models
[7]. Autoregressive processes
[8]. Lag operator
[9]. Moving average processes
[10]. Autoregressive moving average
[11]. Invariant
[12]. Strict- Sense Stationary (SSS)
[13]. Random wall with drift
[14]. trend- stationary processes
[15]. stochastic non- stationary
[16]. Deterministic non- stationary
[17]. Diffrence
[18]. Autoregressive Integrated moving average
[19]. Identification
[20]. Estimation
[21]. Diagnostic checking
[22]. Dendrite
[23]. Soma or Cell Body
[24]. Axon
[25]. Synapse
[26]. Multi- layer Feed Forward
[27]. pure line
[28]. Seosonal autoregressive integrated moving average

www.SID.ir

Arshive of SID

[29]. Support Vector Machines
[30]. Buyok Menderes
[31]. Interval value
[32]. Range of Interval
[33]. Mid point
[34]. Diebold-Mariano

www.SID.ir

Arshive of SID

۱- مقدمه

پیش‌بینی قیمت همواره یکی از چالش برانگیزترین مسائل پیش‌روی سفته‌بازان، معامله‌گران، کارگزاران و ... می‌باشد. از طرفی در تحلیل فاصله‌ای[۱] فرض می‌شود که مشاهدات و تخمین‌ها در دنیای واقعی معمولاً کامل و مطمئن نمی‌باشند، بنابراین برای افزایش دقت، باید داده‌ها را به صورت فاصله‌هایی که کمیت‌های واقعی را در بر می‌گیرند بیان نمود، از این رو، در این پژوهش علاوه بر پیش‌بینی قیمت پایانی به پیش‌بینی دامنه‌ی نوسان قیمت (بالاترین و پایین‌ترین قیمت) پرداخته شده است.

خودرگرسیو میانگین متحرک انباشته (ARIMA) که خلاصه شده‌ی (Autoregressive Integrated Moving Average) می‌باشد و تحت عنوان روش باکس و جنکینز[۲] شناخته می‌شود، یکی از پرکاربردترین مدل‌ها در پیش‌بینی سری‌های زمانی در طول سه دهه‌ی گذشته بوده است، اما پیش‌فرض اصلی آن این است که رابطه‌ی خطی میان ارزش‌های سری بر قرار است بنابراین رابطه‌های غیرخطی بوسیله‌ی مدل خودرگرسیو میانگین متحرک انباشته نمی‌توانند خوب توضیح داده شوند.

یکی دیگر از روش‌های مورد استفاده در پیش‌بینی سری‌های زمانی شبکه‌ی عصبی است که توان تخمین روابط غیرخطی مختلفی را دارا می‌باشد (اصطلاحاً به شبکه‌ی عصبی تخمین زننده‌ی همگانی[۳] می‌گویند). اما استفاده از شبکه، طبق ادبیات پژوهش در روابط خطی نتایج پیچیده‌ای در بر داشته است. دانستن الگوی داده‌ها، مبنی بر خطی و غیرخطی بودن در دنیای واقعی کمی دشوار است و به ندرت سری‌های زمانی به طور خالص خطی و غیرخطی می‌باشند و اغلب از هر دو الگو تبعیت می‌کنند. بنابراین مسأله اینجاست که چگونه می‌توانیم قیمت پایانی و دامنه‌ی نوسان قیمت را با خطای کمتری پیش‌بینی کنیم؟

۲- مبانی نظری و مروری بر پیشینه پژوهش

در زمینه مدل‌سازی سرهای زمانی، روش‌های متفاوتی مورد استفاده قرار می‌گیرند. مدل‌های سنتی مانند میانگین متحرک، هموارسازی نمایی و خودرگرسیو میانگین متحرک انباشته پیش‌بینی آینده را به روابط خطی از گذشته محدود می‌نمایند، از این مدل‌ها به دلیل سادگی در فهم و کاربرد در دهه‌های اخیر بسیار استفاده شده است.

به دلیل مشاهده الگوهای غیرخطی در دنیای واقعی، یک سری از مدل‌های غیرخطی مانند ناهمسانی واریانس شرطی و ناهمسانی واریانس شرطی تعمیم یافته مطرح گردیدند. همگی این مدل‌ها، الگوهای غیرخطی بخصوصی را توضیح می‌دهند، اما شبکه عصبی مصنوعی (ANN[۴]) توان و

www.SID.ir

Arshive of SID

قدرت پیش‌بینی روابط غیرخطی متفاوتی را داراست و کاملاً انعطاف پذیر عمل می‌نماید، ولی آنچه از ادبیات پژوهش بر می‌آید وجود نتایج متفاوت در استفاده از شبکه‌ی عصبی برای روابط خطی است. برای مثال مارخام و راکز (۱۹۹۸) اذعان داشتند، عملکرد شبکه‌ی عصبی برای مسأله‌های رگرسیون خطی وابسته به اندازه نمونه و سطح شوک[۵] می‌باشد[16]. بنابراین چون دانستن خصوصیت داده‌ها مبنی بر خطی و یا غیرخطی بودن دشوار است و از طرفی دیگر به ندرت روابط کاملاً خطی و یا کاملاً غیرخطی مشاهده می‌شوند (اغلب ترکیبی از دو الگو را دارا می‌باشند)، بنابراین ترکیبی از مدل‌های خودرگرسیو میانگین متحرک انباشته و شبکه‌ی عصبی می‌تواند به افزایش دقت پیش‌بینی منجر گردد. مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) مناسب، مدلی است که پسماندهای آن همبستگی خطی نداشته باشند، اما تحلیل‌های پسماند قادر به شناسایی الگوهای غیرخطی نمی‌باشند؛ بنابراین وجود روابط غیرخطی در پسماندها می‌تواند محدودیتی برای مدل خودرگرسیو میانگین متحرک انباشته به حساب آید.

در مدل ترکیبی سعی بر این است که روابط غیرخطی در پسماندها با استفاده از شبکه‌ی عصبی شناسایی شده و پیش‌بینی پسماندهای حاصل از شبکه به مدل خودرگرسیو میانگین متحرک انباشته اضافه گردد. با این ترکیب، بخش خطی بوسیله‌ی مدل خودرگرسیو میانگین متحرک انباشته و بخش غیرخطی بوسیله‌ی شبکه پیش‌بینی خواهد شد.

از آن‌جا که افزایش روز افزون قیمت جهانی طلا بیانگر رشد تقاضا برای این فلز گرانبها و تمایل جهانی برای در اختیار داشتن آن به عنوان یک کالای سرمایه‌ای است، در این پژوهش با استفاده مدل ترکیبی، علاوه بر پیش‌بینی قیمت پایانی، به پیش‌بینی بالاترین و پایین‌ترین قیمت هفتگی اونس طلا پرداخته شده است. در واقع علاوه بر پیش‌بینی خود قیمت، دامنه‌ی نوسان قیمت نیز پیش‌بینی گردیده و فرضیه‌های زیر را مورد آزمون قرار گرفته است:

۱) پیش‌بینی مدل ترکیبی در حد بالای قیمت طلا نسبت به پیش‌بینی مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری می‌باشد.
۲) پیش‌بینی مدل ترکیبی در حد بالای قیمت طلا نسبت به پیش‌بینی شبکه دارای خطای کمتری می‌باشد.
۳) پیش‌بینی مدل ترکیبی در حد پایین قیمت طلا نسبت به پیش‌بینی مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری می‌باشد.
۴) پیش‌بینی مدل ترکیبی در حد پایین قیمت طلا نسبت به پیش‌بینی شبکه دارای خطای کمتری می‌باشد.

www.SID.ir

Arshive of SID

۵) پیش‌بینی مدل ترکیبی در قیمت پایانی هفتگی نسبت به پیش‌بینی مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری است.
۶) پیش‌بینی مدل ترکیبی در قیمت پایانی هفتگی نسبت به پیش‌بینی مدل شبکه‌ی عصبی پیش خور دارای خطای کمتری است.

پیشینه پژوهش

در این بخش بعد از مروری کوتاه بر مبانی نظری مدل‌های سری‌زمانی تک‌متغیره و شبکه‌ی عصبی به بررسی پیشینه تجربی پرداخته می‌شود و در نهایت مدل ترکیبی به‌کار رفته در پژوهش ارائه می‌گردد.

مدل‌های سری زمانی تک متغیره مدل‌هایی هستند که برای پیش‌بینی ارزش یک متغیر تنها از اطلاعات گذشته‌ی همان متغیر و اطلاعات حال و گذشته‌ی اجزای اخلال بهره می‌گیرند.

مدل‌های سری‌زمانی در مقابل مدل‌های ساختاری[۶] مطرح می‌گردند. اما این مدل‌ها نسبت به مدل‌های ساختاری دارای دو مزیت عمده به قرار زیر می‌باشند:

۱) مدل‌های ساختاری یک پیش‌فرض تئوریک را مبنای برقراری رابطه‌ی علت و معلول بین متغیرهای مستقل و وابسته قرار می‌دهند، در حالی که مدل‌های سری زمانی نیازی به این پیش فرض ندارند.
۲) مدل‌های ساختاری در پیش‌بینی خارج از نمونه‌ای چندان خوب عمل نمی‌کنند.

فرآیندهای خودرگرسیو[۷] (AR)

فرآیندهای خودرگرسیو فرآیندهایی هستند که در آن ارزش جاری متغیر X_t تنها به ارزش-های گذشته خود متغیر و جزء اخلال، وابسته می‌باشد. یک فرآیند AR از مرتبه P را به صورت های زیر می‌توان نمایش داد:

$$X_t = \mu + Q_1 X_{t-1} + Q_2 X_{t-2} + \cdots + Q_p X_{t-p} + U_t \qquad AR(p) \qquad \text{رابطه(۱)}:$$

$$X_t = \mu + \sum_{i=1}^{p} Q_i L^i X_t + U_t$$

$$\begin{cases} Q(L)X_t = \mu + U_t \\ Q(L) = 1 - Q_1 L^1 - Q_2 L^2 - \cdots - Q_p L^p \end{cases}$$

www.SID.ir

Arshive of SID

L: عملگر وقفه[8]

U_t: یک فرآیند نوفه سفید

یک فرآیند خودرگرسیو زمانی مانا است که $|Q|<1$ باشد.[6]

فرآیندهای میانگین متحرک[9] (MA)

فرآیندهای میانگین متحرک، ترکیب خطی ساده‌ای از فرآیندهای نوفه سفید می‌باشند که در آن ارزش جاری X_t به ارزش‌های جاری و گذشته اجزاء اخلال، وابسته است.[6]

فرآیندهای میانگین متحرک از مرتبه ی q را به صورت های زیر می‌توان نمایش داد:

رابطه (۲): $\mu A(q)$

$$X_t = \mu + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \ldots + \theta_q U_{t-q} + U_t$$

$$X_t = \mu + \sum_{j=1}^{q} \theta_j L^j U_t + U_t$$

$$\begin{cases} X_t = \mu + \theta(L)U_t \\ \theta(L) = 1 + \theta_1 L^1 + \theta_2 L^2 + \cdots + \theta_p L^p \end{cases}$$

فرآیندهای خودرگرسیو میانگین متحرک[10] (ARMA)

مدل‌های ARMA(p,q) از ترکیب مدل‌های خودرگرسیو از مرتبه‌ی p و میانگین متحرک از مرتبه‌ی q حاصل می‌شوند. در این مدل‌ها ارزش جاری سری زمانی X_t به صورت خطی به ارزش-های گذشته‌ی خودش و ارزش‌های جاری و گذشته‌ی اجزاء اخلال، وابسته است.

مدل‌ها خودرگرسیو میانگین متحرک را به صورت‌های زیر می‌توان نمایش داد:

رابطه (۳):

$$X_t = \mu + Q_1 x_{t-1} + Q_2 x_{t-2} + .. + Q_p x_{t-p} + \theta_1 U_{t-1} + \ldots + \theta_q U_{t-q} + U_t$$

$$X_t = \mu + \sum_{i=1}^{p} Q_i L^i X_t + \sum_{j=1}^{q} \theta_j L^j U_t + U_t$$

$$\begin{cases} Q(L)X_t = \mu + \theta(L)U_t \\ Q(L) = 1 - Q_1 L^1 - Q_2 L^2 - \cdots - Q_p L^p \\ \theta(L) = 1 + \theta_1 L^1 + \theta_2 L^2 + \cdots + \theta_p L^p \end{cases}$$

www.SID.ir

Arshive of SID

مدل‌های رگرسیو میانگین متحرک، هم خصوصیات مدل‌های خودرگسیو (AR) و هم خصوصیات مدل‌های میانگین متحرک (MA) را در بر می‌گیرند. در مدل‌های خودرگسیو، تابع خودهمبستگی به صورت نمایی کاهش پیدا می‌کند در حالی که تابع خودهمبستگی جزئی تعداد وقفه‌های مورد نیاز برای مدل را معین می‌سازد، اما در مدل‌های میانگین متحرک، این دو تابع برعکس مدل‌های خودرگسیو، عمل می‌نمایند.[6]

مانایی

اگر تابع توزیع یک سری زمانی مانند $X(t)$ از مرتبه ی n ام نسبت به تغییر مبدأ زمانی ناوردا[11] باشد، آن گاه سری زمانی $X(t)$ را یک سری مانای قوی[12] می نامیم.[15]

رابطه (۴): $$F(x_1, x_2, ...x_n; t_1, t_2, ...t_n) = F(x_1, x_2, ...x_n; t_1 + S, t_2 + S, ...t_n + S)$$

اما از آن جایی که در عمل به سختی می‌توانیم نشان دهیم که یک سری زمانی مانای قوی است (به جزء در توزیع نرمال)، بنابراین به جای در نظر گرفتن تابع توزیع مشترک n مرتبه‌ای، تابع توزیع یک یا دو مرتبه‌ای را لحاظ می‌کنیم. بنابراین اگر تابع توزیع یک سری مانند $X(t)$ از مرتبه‌ی اول و یا دوم نسبت به تغییر مبدأ زمانی ناوردا باشد، داریم:

رابطه(۵): $$F(x;t) = F(x;t+S) \quad \forall t, S$$

$$F(x_1, x_2; t_1, t_2) = F(x_1, x_2; t_1 + S, t_2 + S) \quad \forall t_1, t_2, S$$

از معادلات بالا می‌توان نتیجه گرفت که میانگین سری زمانی $X(t)$ در طول زمان ثابت و کوواریانس بین $X(t)$ و $X(t+S)$ تنها به فاصله‌ی زمانی یعنی S وابسته است.

انواع نامانایی

اغلب دو مدل به منظور مشخص کردن نامانایی به کار گرفته می‌شود:

۱. مدل گشت تصادفی با رانش[13]

رابطه (۶): $$X_t = \mu + x_{t-1} + U_t$$

۲. فرآیند روند- مانا[14] (اطراف خط روند مانا می باشند)

www.SID.ir

Arshive of SID

رابطه (۷):

$$X_t = \alpha + \beta T + U_t$$

که مدل اول را نامانای تصادفی[۱۵] و مدل دوم را نامانای معین[۱۶] می نامند.
برای مانا کردن مدل‌های گشت تصادفی، معمولاً از تفاضل[۱۷] و برای مانا کردن مدل‌های روند- مانا معمولاً از روند زدایی استفاده می‌گردد.[6]

مدل‌های خودرگرسیو میانگین متحرک انباشته (ARIMA)[۱۸]

شرط بهره‌گیری از مدل‌های خودرگرسیو میانگین متحرک (ARMA)، مانا بودن سری‌های زمانی می‌باشد. برای مانا کردن سری‌های زمانی مالی و اقتصادی اغلب از روش تفاضل استفاده می-شود. جزء یکپارچگی در مدل خودرگرسیو میانگین متحرک انباشته بیان‌گر میزان تفاضل مورد نیاز برای مانا کردن سری زمانی است.[9]

مدل‌های خودرگرسیو میانگین متحرک انباشته را معمولاً به صورت ARIMA(p,d,q) نمایش می‌دهند که در آن p مرتبه‌ی مورد نیاز برای خودرگرسیو، d میزان تفاضل مورد نیاز برای مانا کردن سری زمانی و q میزان مرتبه‌ی مورد نیاز برای میانگین متحرک است.

مراحل ساخت مدل های ARIMA

۱) مانا کردن سری زمانی: با استفاده از روش تفاضل، به منظور استفاده از مدل‌های خودرگرسیو میانگین متحرک انباشته ، سری زمانی را مانا می‌کنیم.
۲) شناسایی مدل[۱۹]: در این مرحله مرتبه‌های مورد نیاز برای ساخت مدل تعیین می‌گردد. برای این منظور می‌توان از توابع خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) بهره گرفت.
۳) تخمین مدل[۲۰]: تخمین ضرایب با استفاده از تکنیک‌هایی مانند حداقل مربعات، حداکثر درست نمایی و از این قبیل.
۴) بررسی مدل[۲۱]: در این مرحله با استفاده از توابع خودهمبستگی و خودهمبستگی جزئی و آماره‌ی باکس و الژانگ، مستقل بودن سری پسماندها، مورد آزمون قرار می‌گیرد. مدل خودرگرسیو میانگین متحرک انباشته بهینه مدلی است که پسماندهای آن فاقد همبستگی خطی باشند.
۵) پیش‌بینی: با استفاده از مدل خودرگرسیو میانگین متحرک انباشته بهینه، ارزش متغیر در آینده را پیش‌بینی می‌کنیم.

www.SID.ir

Arshive of SID

شبکه های عصبی مصنوعی

شبکه‌های عصبی مصنوعی از عناصر عملیاتی ساده‌ای ساخته می‌شوند که به صورت موازی در کنار یکدیگر عمل می‌کنند. این عناصر که از سیستم‌های عصبی زیستی الهام گرفته شده‌اند، در تلاش‌اند که به صورت ناپارامتریک، مغز انسان را شبیه سازی نمایند. [2]

در شبکه‌ی عصبی مصنوعی به این عناصر پردازشی، نرون‌های مصنوعی می‌گویند که این نرون‌ها مشابه نرون‌های بیولوژیکی مغز انسان می‌باشند.

مغز انسان دارای 10^{11} نرون است که هر کدام از آن‌ها دارای هزاران ارتباط با یکدیگر می باشند. هر نرون از ۳ بخش اصلی تشکیل می شود که عبارتند از: دندریت[۲۲]، هسته[۲۳] و آکسون[۲۴] نرون‌ها از طریق دندریت‌ها سیگنال‌ها را از سایر نرون‌ها دریافت کرده و از طریق آکسون‌ها به دیگر نرون‌ها ارسال می‌کنند. به ارتباط میان آکسون و دندریت، سیناپس[۲۵] گفته می‌شود. در واقع، ورودی یک شبکه‌ی عصبی مصنوعی، مشابه ایمپلاسهای الکتروشیمیایی است که توسط دندریت-های نرون‌های بیولوژیکی دریافت می‌گردد و خروجی نرون‌های مصنوعی مشابه سیگنال‌های خروجی از آکسون یک نرون بیولوژیکی است.

شبکه‌های پس‌انتشار

شبکه پس‌انتشار، که به اختصار آن را با BP نمایش می‌دهیم، یک شبکه چند لایه با تابع انتقال غیرخطی و قاعده یادگیری Widrow-Hoff می‌باشد. از بردار ورودی و هدف در راستای آموزش این نوع شبکه، برای تقریب زدن یک تابع، یافتن رابطه‌ی بین ورودی و خروجی و دسته‌بندی ورودی‌ها استفاده می‌گردد.

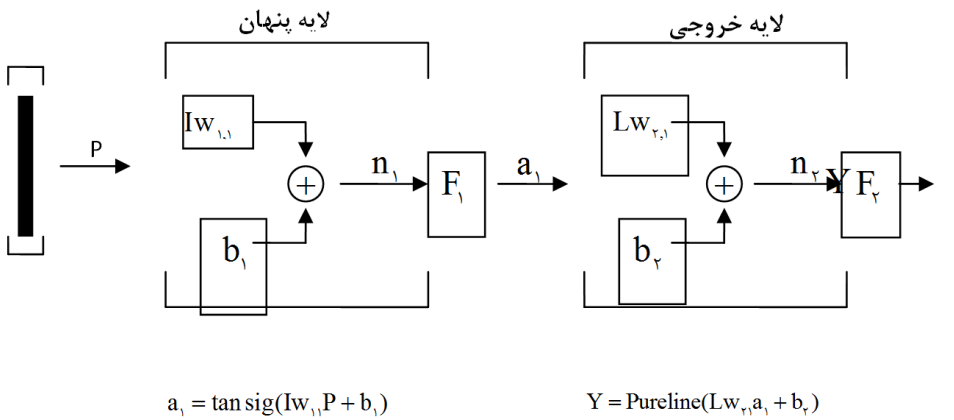
یکی از مرسوم‌ترین معماری‌های شبکه‌های BP، شبکه‌های چند لایه پیش‌خور می‌باشند[۲۶] که به منظور تخمین مدل، طبقه‌بندی و مدل‌سازی به طور گسترده مورد استفاده قرار می‌گیرد. [13]

شبکه‌های پیش‌خور اغلب دارای یک یا چند لایه‌ی پنهان از نرون‌های sigmoid بوده و در لایه‌ی خروجی خود از تابع فعال سازی خطی[۲۷] استفاده می‌کنند.

شبکه‌ی پیش‌خور (ff) را می‌توان به صورت $I\times H\times O$ بیان نمود به طوری که I میزان نرون‌ها در لایه‌ی ورودی و H میزان نرون‌ها در لایه‌ی پنهان و O میزان نرون‌ها در لایه‌ی خروجی را نشان می‌دهد. شکل زیر یک شبکه‌ی پیش‌خور را به نمایش می‌کشد.

www.SID.ir

Arshive of SID



شکل ۱- شبکه‌ی عصبی پیش‌خور با تابع فعال سازی تانژانت هیپربولیک

مطالعات زیادی در زمینه‌ی مقایسه‌ی روش‌های خطی از قبیل مدل خودرگرسیو میانگین متحرک انباشته و غیرخطی از قبیل شبکه‌های عصبی انجام پذیرفته است. آن‌چه از این مطالعات بر می‌آید، وجود نتایج متفاوت در استفاده از شبکه‌های عصبی است برای مثال:

شاردا و پاتیل (۱۹۹۰) در مقاله‌ای از ۷۵ سری زمانی برگرفته از مقاله‌ی به منظور مقایسه عملکرد شبکه‌ی عصبی و خودرگرسیو میانگین متحرک انباشته (ARIMA) بهره جستند که در ۳۹ سری زمانی عملکرد شبکه، و در ۳۶ سری زمانی، عملکرد ARIMA دقیق‌تر بود.[18]

فاستر و همکاران (۱۹۹۲) به مقایسه‌ی عملکرد شبکه‌ی عصبی با مدل رگرسیون خطی و میانگین شش مدل هموارسازی نمایی در ۳۸۴ سری زمانی اقتصادی و جمعیت‌شناختی پرداختند. نتایج نشان داد که استفاده از رگرسیون خطی به طور معناداری از شبکه‌ی عصبی به عنوان تخمین زننده، در این سری های تصادفی، بهتر است. [11]

تانگ و فیش ویچ (۱۹۹۳) در مقاله‌ی خود اذعان داشتند، عملکرد شبکه‌ی به ماهیت داده‌ها وابسته می‌باشد و برای سری‌های زمانی بی‌قاعده و پیش‌بینی چند مرحله‌ای، این مدل نسبت به مدل خودرگرسیو میانگین متحرک انباشته دارای خطای کمتری است.[19]

ایندرو و همکاران (۱۹۹۹) اذعان داشتند که، عملکرد شبکه عصبی نسبت به رگرسیون خطی، در صندوق‌های مشترک سرمایه‌گذاری به نوع صندوق وابسته است.[12]

www.SID.ir

Arshive of SID

کلین و راسین (۱۹۹۹) بیان داشتند که، عملکرد شبکه‌ی عصبی وابسته به ماهیت داده‌ها می-باشد و کاهش نرخ خطا و مقدار خطا می‌تواند منجر به بهبود عملکرد شبکه‌ی عصبی در پیش‌بینی خارج از نمونه‌ای گردد.[14]

رضا روشن (۱۳۸۵) در پایان‌نامه‌ی خود توان پیش‌بینی تورم مدل‌های خودرگرسیو میانگین متحرک انباشته، ناهمسانی واریانس شرطی و شبکه‌ی عصبی را با هم مقایسه کرده و نتایج بر کارآتر بودن مدل ناهمسانی واریانس شرطی نسبت به مدل خودرگرسیو میانگین متحرک انباشته و مدل شبکه‌های عصبی مصنوعی نسبت به رقبای خود دلالت دارد.[3]

متوسلی و طالب (۱۳۸۵) نتیجه گرفتند که عملکرد مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) به طور معناداری نسبت به شبکه‌های عصبی بهتر است و در واقع مدل خطی خودرگرسیو میانگین متحرک انباشته بهتر از مدل‌های غیرخطی (شبکه‌های عصبی) می‌تواند پیچیدگی سری زمانی قیمت سهام را تجزیه و تحلیل نماید.[5]

بنابراین با توجه به عملکرد متفاوت مدل‌های خطی و غیر خطی در سری‌های زمانی مختلف در بسیاری از مطالعات به بررسی توان پیش‌بینی مدل‌های ترکیبی پرداخته شده است برای نمونه ژانگ (۲۰۰۳) در مقاله‌ای خود به تلفیق ARIMA و شبکه‌ی عصبی پرداخت که نتایج این پژوهش دال بر موفقیت مدل ترکیبی نسبت به مدل‌های ARIMA و شبکه‌ی عصبی می‌باشد. [23]

ودینگ و سیوس (۱۹۹۶) نتیجه گرفتند، اگر پیش‌بینی سری زمانی و فاکتورهای اطمینان که توسط شبکه‌ی عصبی RBF تولید شده است را با مدل باکس و جنکینز ترکیب نماییم، این امر موجب افزایش دقت پیش‌بینی می‌شود.[21]

تی سنگ و همکاران (۲۰۰۲) در مقاله‌ای سعی در ترکیب شبکه‌ی عصبی پس‌انتشار و مدل-های سری زمانی فصلی (SARIMA[۲۸]) نمودند و آن را با (SARIMABP) نمایش دادند. نتایج پژوهش نشان داد که عملکرد مدل SARIMABP نسبت به مدل خودرگرسیو میانگین متحرک انباشته‌ی فصلی و شبکه‌های عصبی با داده‌های تفاضلی و فصل زدا شده به طور معناداری بهتر است.[20]

لیو و پای (۲۰۰۵) برای سری زمانی ده سهم بورس نیویورک، سعی در ترکیب مدل خطی خودرگرسیو میانگین متحرک انباشته و مدل غیر خطی ماشین بردار پشتیبان[۲۹] نمودند و نتایج نشان داد که عملکرد مدل ترکیبی به طور معناداری از سایر مدل‌ها بهتر است.[17]

فاروق (۲۰۱۰) در مقاله خود از سه سری زمانی دمای آب، ماده‌ی بورون (Boron) و اکسیژن حل شده در طول سال‌های ۲۰۰۴-۱۹۹۶ در رودخانه‌ی بویاک مندرز[۳۰] ترکیه برای مقایسه توان پیش‌بینی ترکیب مدل خطی خودرگرسیو میانگین متحرک انباشته‌ی فصلی و شبکه‌ی عصبی

www.SID.ir

Arshive of SID

نسبت به هر یک از مدل‌ها مذکور بهره جست و نتایج دال بر موفقیت مدل ترکیبی نسبت به دو مدل دیگر می‌باشد.[10]

۳- مدل مفهومی

مدل ترکیبی به کار گرفته شده در این پژوهش به این صورت است ابتدا هر سری زمانی مانند Y_t به دو بخش خطی و غیرخطی به صورت زیر تقسیم می‌گردد:

رابطه(۸):

$$Y_t = L_t + N_t$$

L_t: جزء خطی

N_t: جزء غیرخطی

آنگاه با استفاده از مدل خودرگرسیو میانگین متحرک انباشته بخش خطی سری زمانی Y_t پیش-بینی گردیده و پسماندهای حاصل از این پیش‌بینی به قرار زیر محاسبه می‌شود:

رابطه(۹):

$$Y_t - \hat{L}_t = e_t$$

e_t: پسماندهای حاصل از پیش‌بینی مدل ARIMA

$\hat{L}_t$: پیش‌بینی مدل ARIMA

سپس با استفاده از شبکه‌ی عصبی، بخش غیرخطی به صورت زیر محاسبه می‌گردد:

رابطه(۱۰):

$$e_t = f(e_{t-1}, e_{t-2}, ..., e_{t-n}) + \varepsilon_t$$

f: تابع غیرخطی برازش شده توسط شبکه‌ی عصبی

ε_t: جزء اخلال

در نهایت با افزودن پیش‌بینی شبکه‌ی عصبی به پیش‌بینی مدل خطی خودرگرسیو میانگین متحرک انباشته، پیش‌بینی مدل ترکیبی برای سری زمانی Y_t حاصل خواهد شد. در واقع داریم:

رابطه(۱۱):

$$\hat{e}_t = \hat{N}_t$$

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t$$

$\hat{Y}_t$: پیش‌بینی مدل ترکیبی

www.SID.ir

Arshive of SID

$\hat{N}_t$: پیش‌بینی پسماندها توسط شبکه

۴- روش‌شناسی پژوهش

در این پژوهش به‌منظور مقایسه عملکرد مدل ترکیبی ارائه شده در قسمت قبل با هر یک از مدل‌های ARIMA و شبکه‌ی عصبی از سری‌های زمانی حدبالای قیمت هفتگی اونس طلا ، حد پایین قیمت هفتگی اونس طلا و قیمت پایانی هفتگی اونس طلا از ابتدای سال ۲۰۰۶ تا اکتبر ۲۰۱۱ بهره گرفته شده است.

از آن‌جایی که یکی از روش‌های مورد استفاده در محاسبه‌ی ارزش فاصله‌ای[۳۱]، بهره‌گیری از دامنه‌ی تغییرات[۳۲] و نقطه وسط[۳۳] می‌باشد؛ بنابراین، از بالاترین و پایین‌ترین قیمت طلا در طول یک هفته استفاده شده و دو سری زمانی دامنه‌ی تغییرات و نقطه‌ی وسط برای هر هفته به قرار زیر محاسبه گردیده است:

رابطه(۱۲):

$$x_t^c = \frac{x_t^u + x_t^l}{۲}$$

رابطه(۱۳):

$$x_t^r = \frac{x_t^u - x_t^l}{۲}$$

x_t^c: نقطه وسط

x_t^r: دامنه‌ی تغییرات

x_t^u: بالاترین قیمت هفتگی (حد بالا)

X_t^l: پایین‌ترین قیمت هفتگی (حد پایین)

همان‌طور که مشخص است، جمع سری‌های زمانی دامنه‌ی تغییرات و نقطه‌ی وسط بیان‌گر حد بالای قیمت (بالاترین قیمت) در طول یک هفته و تفاضل این دو سری زمانی بیان‌گر حد پایین قیمت (پایین‌ترین قیمت) در طول یک هفته می‌باشد. بنابراین از سری‌های زمانی دامنه‌ی تغییرات و نقطه‌ی وسط به عنوان متغیرهای اصلی برای پیش‌بینی دامنه‌ی نوسان طلا و از سری زمانی قیمت پایانی به عنوان متغیر اصلی برای پیش‌بینی قیمت پایانی در هر یک از مدل‌های خودرگرسیو میانگین متحرک انباشته (ARIMA)، شبکه‌ی عصبی و مدل ترکیبی استفاده شده است.

بعد از مانا کردن سری‌های زمانی نقطه‌ی وسط، دامنه‌ی تغییرات و قیمت پایانی (از طریق محاسبه‌ی بازده‌های لگاریتمی)، ابتدا با ۲۷۳ مشاهده از سری‌های مربوطه مدل خودرگرسیو

www.SID.ir

Arshive of SID

میانگین متحرک انباشته برازش شده و معیارهای میانگین مجذور خطا و میانگین قدر مطلق خطا برای مشاهده ۲۷۴ام از حد بالای قیمت، حد پایین قیمت و قیمت پایانی به روش زیر محاسبه گردیده است:

رابطه(۱۴):

$$MSE = \frac{\sum_{i=1}^{n}(Y_{R_i} - Y_{F_i})^2}{n}$$

رابطه(۱۵):

$$MAE = \frac{\sum_{i=1}^{n}\left|Y_{R_i} - Y_{F_i}\right|}{n}$$

Y_R: مقدار واقعی

Y_F: مقدار پیش‌بینی

n : تعداد اجزاء خارج نمونه‌ای

سپس بعد از حذف یک مشاهده از بالا، مشاهده ۲۷۴ام واقعی از سری‌های زمانی مربوطه، به منظور برازش مجدد مدل خودرگرسیو میانگین متحرک انباشته و معیارهای ارزیابی مذکور برای مشاهده ۲۷۵ام از حد بالای قیمت، حد پایین قیمت و قمت پایانی به مدل اضافه شده و فرایند فوق برای ۳۰ مرتبه انجام پذیرفته است.

بعد از برازش هر مرتبه از مدل خودرگرسیو میانگین متحرک انباشته (ARIMA)، بر طبق مدل ترکیبی مطرح شده در زیر بخش مدل مفهومی، با استفاده از شبکه‌ی عصبی، الگوی غیرخطیِ پسماندهای حاصل از ARIMA مدل‌سازی گردیده، آن‌گاه پیش‌بینی پسماندها برای یک مرحله پیش‌رو به پیش‌بینی مدل خطی خودرگرسیو میانگین متحرک انباشته (ARIMA) به منظور بررسی دقت پیش‌بینی مدل ترکیبی در حد بالای قیمت، حد پایین قیمت و قیمت پایانی اضافه گردیده است.

به‌منظور بررسی دقت مدل ترکیبی نسبت به دو مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) و شبکه‌ی عصبی بعد از برازش هر مرتبه از مدل ARIMA، بوسیله‌ی بهترین شبکه‌ی عصبی برازش شده، دقت پیش‌بینی مدل شبکه‌ی عصبی پیش‌خور در حد بالای قیمت، حد پایین قیمت و قیمت پایانی مورد ارزیابی قرار داده شده است.

در برازش شبکه‌ی عصبی، چه در برازش خود قیمت‌ها و چه در مدل سازی روابط غیر خطی در پسماندها، از کد نویسی در نرم افزار متلب و تغییر در تعداد وقفه‌های زمانی، لایه‌های پنهان و

www.SID.ir

Arshive of SID

نرون‌های لایه‌های پنهان بهره گرفته شده و شبکه‌ای که در بخش تست (۱۰ درصدرسری‌زمانی) کمترین خطا را دارا می‌باشد مبنای پیش‌بینی یک مرحله پیش‌رو قرار گرفته است.

در نهایت، معناداری تفاوت میانگین بین دو جامعه میانگین مجذور خطا و میانگین قدر مطلق خطا برای بررسی فرضیه‌های پژوهش، با استفاده از آزمون مقایسه زوجی مورد بررسی واقع گردیده است.

اما از آن‌جا که در پیش بینی های سری زمانی برای H مرحله پیش‌رو از داده های روی هم افتاده استفاده می نماییم، خطای پیش بینی ${\varepsilon_{t+H|t}}^1$ و ${\varepsilon_{t+H|t}}^2$ (مربوط به مدل‌های ۱ و ۲) به صورت سریالی همبستگی خواهد داشت. از این رو در آزمون مقایسه میانگین دو جامعه میانگین مجذور خطا علاوه بر آزمون مقایسه زوجی از آزمون دایبولد- ماریانو[۳۴] و آماره تعدیل شده آن توسط هاروی – لیبورن- نیوبولد (برای نمونه‌های کوچک)، برای بررسی معناداری دقت پیش‌بینی مدل ترکیبی نسبت به هریک از مدل‌های ARIMA و شبکه‌ی عصبی، بهره گرفته شده است.

آزمون دایبولد-ماریانو به‌صورت زیر قابل تعریف است:

رابطه (۱۶):

$$d_t = L\left(\varepsilon^1_{t+H|t}\right) - L\left(\varepsilon^2_{t+H|t}\right)$$

رابطه (۱۷):

$$S = \frac{\bar{d}}{\left(\widehat{avar}(\bar{d})\right)^{1/2}} = \frac{\bar{d}}{\left(\widehat{LRV}_{\bar{d}}/T\right)^{1/2}}$$

رابطه (۱۸):

$$LRV_{\bar{d}} = \gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j \ , \qquad \gamma_j = cov(d_t, d_{t-j})$$

$LRV_{\bar{d}}$، تخمین سازگاری از واریانس مجانبی $\sqrt{T}\ \bar{d}$ است. دایبولد و ماریانو (۱۹۹۵) نشان می‌دهند که تحت فرضیه قدرت پیش‌بینی کنندگی یکسان داریم:

$S \sim N(0,1)$

بنابراین فرضیه قدرت پیش بینی یکسان را در سطح ۹۵٪ زمانی رد می کنیم که داشته باشیم:

$|S| > 1/96$

اما این آماره سال‌های بعدها توسط هاروی – لیبورن- نیوبولد برای نمونه‌های کوچک و تابع زیان خاص مجذور خطا به صورت زیر اصلاح گردید:

www.SID.ir

Arshive of SID

رابطه (۱۹):

$$S^* = \frac{\bar{d}}{\left(\widehat{avar}(\bar{d})\right)^{1/2}} = \frac{\bar{d}}{\sqrt{\widehat{var}_m}}$$

رابطه (۲۰):

$$\widehat{var}_m = [N + 1 - 2h + (n^{-1} * h(h-1))]^{-1} \times \widehat{\gamma_0} + 2\sum_{j=1}^{h-1} \widehat{\gamma_j}$$

۵- یافته‌های پژوهش

همانطور که در بخش روش پژوهش ذکر شد، به‌منظور بررسی فرضیه‌های تحقیق مبنی بر دقت مدل ترکیبی نسبت به هر یک از مدل‌های ARIMA و شبکه‌ی عصبی در حد بالای قیمت، حد پایین قیمت و قیمت پایانی از آزمون مقایسه زوجی، آزمون دایبولد-ماریانو و آماره‌ی تعدیل شده هاروی – لیبورن- نیوبولد و روش غلتاندن داده‌ها بهره گرفته شده است. در این بخش ابتدا به بررسی آمار توصیفی برای هر یک از سه سری زمانی حد بالای قیمت، حد پایین قیمت و قیمت پایانی پرداخته شده و از آنجایی که یکی از شروط استفاده از مدل ARIMA مانا بودن سری‌های زمانی است نتایج حاصل از آزمون دیکی-فولر برای بازده‌های لگاریتمی سری‌های زمانی حد بالا قیمت، حد پایین قیمت و قیمت پایانی انس طلا به تفصیل ارائه می‌گردد. در نهایت نتایج حاصل از مقایسه سه مدل بر اساس آزمون‌های مذکور و برای هر یک از توابع زیان که در بخش‌های قبل به تفصیل مورد بررسی قرار گرفتند، در قالب جداولی مورد بررسی قرار داده می‌شوند.

جدول شماره ۱ نتایج حاصل از آمار توصیفی برای هر یک از سه سری‌زمانی را به‌صورت مجزا تبین می‌نماید.

اما همانطور که ذکر شد یکی از مهم‌ترین ویژگی در برازش مدل ARIMA مانا بودن سری‌های زمانی است. سری‌های زمانی قیمت در فضای مالی و اقتصادی اکثرا مانا نبوده و دارای گشت تصادفی می‌باشند. در واقع از منظر اقتصاد سنجی سری‌های زمانی مالی و اقتصادی اکثرا هم انباشته از مرتبه اول هستند.از این رو این ایده به ذهن می‌رسد که تفاضل از مرتبه اول می‌تواند منجر به مانایی سری‌های زمانی و در نهایت هم انباشتگی از مرتبه صفر شود. بازده در واقع نوعی تفاضل از مرتبه اول است. بنابراین نتایج حاصل از آزمون دیکی-فولر برای بازده لگاریتمی هر یک سه سری زمانی حد بالا قیمت، حد پایین قیمت و قیمت پایانی در قالب جدول شماره ۲، ۳ و ۴ به تفصیل ارائه شده است.

www.SID.ir

Arshive of SID

جدول ۱- آمار توصیفی

	حد بالا قیمت	حد پایین قیمت	قیمت پایانی
میانگین	۱۰۳۲,۱۵۹	۹۰۳,۴۱۰۳	۱۰۱۲,۹۵۲
میانه	۹۴۲,۶۵۰۰	۹۰۳,۳۵۰۰	۹۲۵,۸۵۰۰
ماکسیمم	۱۹۲۰,۷۸۰	۱۷۹۳,۱۰۰	۱۸۸۵,۳۶۰
مینیمم	۵۵۴,۴۰۰۰	۵۳۳,۵۰۰	۵۴۰,۶۰۰۰
انحراف معیاد	۳۵۸,۲۰۶۵	۳۴۴,۹۵۰۳	۳۵۳,۶۰۱۲
چولگی	۰,۶۶۹۴۸۸	۰,۶۳۶۴۱۳	۰,۶۶۲۸۹۱
کشیدگی	۲,۴۴۵۶۰۱	۲,۲۹۲۲۳۰	۲,۳۹۲۵۰۴
جارک-برا	۲۸,۰۹۰۳	۲۸,۳۶۸۷۳	۲۸,۴۴۵۲۴
سطح معناداری	۰,۰۰۰۰۰۱	۰,۰۰۰۰۰۱	۰,۰۰۰۰۰۱
جمع	۳۳۱۳۲۳,۱	۳۱۶۹۵۸,۷	۳۲۵۱۵۷,۵
تعداد مشاهدات	۳۲۱	۳۲۱	۳۲۱

جدول ۲- نتایج آزمون دیکی فولر برای بازده لگاریتمی حد بالا قیمت

Null Hypothesis: D(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-13.34955	0.0000
Test critical values:	1% level	-3.988840	
	5% level	-3.424825	
	10% level	-3.135494	

*MacKinnon (1996) one-sided p-values.

www.SID.ir

Arshive of SID

جدول ۳- نتایج آزمون دیکی فولر برای بازده لگاریتمی حد پایین قیمت

Null Hypothesis: D(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-15.31724	0.0000
Test critical values:	1% level	-3.988840	
	5% level	-3.424825	
	10% level	-3.135494	

*MacKinnon (1996) one-sided p-values.

جدول ۴- نتایج آزمون دیکی فولر برای بازده لگاریتمی قیمت پایانی

Null Hypothesis: D(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-17.34042	0.0000
Test critical values:	1% level	-3.988840	
	5% level	-3.424825	
	10% level	-3.135494	

*MacKinnon (1996) one-sided p-values.

نتایج جداول ۲، ۳ و ۴ نشان می‌دهد که بازده لگاریتمی در تمامی سطوح خطا ۱ درصد، ۵ درصد و ۱۰ درصد مانا است. از این رو در این پژوهش از بازده لگاریتمی این سری‌های زمانی استفاده شده است.

اما هدف نهایی این پژوهش مقایسه عملکرد مدل ترکیبی در مقابل هر یک از مدل‌های خطی و غیر خطی به‌صورت مجزا است از این رو جدول ۵ نتایج آزمون مقایسه زوجی با تابع زیان میانگین مجذور خطا را نشان می‌دهد که اندیسup ، Down و Close به ترتیب بیانگر حدبالای قیمت، حد پایین قیمت و قیکت پایانی اونس طلا می‌باشند :

www.SID.ir

Arshive of SID

جدول ۵- آزمون مقایسه‌ی زوجی (میانگین مجذور خطا)

مقایسه زوجی مجذور خطا	Mean	Std. Deviation	t	Df	Sig. (2-tailed)
$Hybrid_{up}$- $ARIMA_{up}$	-540/96	764/08	-3/87	29	.001
$Hybrid_{up} - ANN_{up}$	-708/8۰	1472/86	-2/63	29	.013
$Hybrid_{down} - ARIMA_{down}$	-449/3۰	783/51	-3/14	29	.004
$Hybrid_{down} - ANN_{down}$	-947/5	1976/24	-2/62	29	.014
$Hybrid_{close} - ARIMA_{close}$	-1001/50	1383/79	-3/96	29	.000
$Hybrid_{close}$–ANN_{close}	-571/4	1183/38	-2/64	29	.013

با توجه به جدول فوق بین میانگین دو جامعه‌ی MSE از مدل ترکیبی و مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) و دو جامعه‌ی MSE از مدل ترکیبی و مدل شبکه‌ی عصبی اختلاف معناداری وجود دارد و میانگین MSE برای مدل ترکیبی در نمونه‌ی مورد بررسی از دو مدل دیگر کمتر است. بنابراین، از نظر معیار ارزیابی MSE ، عملکرد مدل ترکیبی به طور معناداری در حد بالای قیمت، حد پایین قیمت و قیمت پایانی از دو مدل خودرگرسیو میانگین متحرک انباشته و شبکه‌ی عصبی (ANN) بهتر است. جدول ۶ نتایج آزمون مقایسه زوجی با تابع زیان میانگین قدر مطلق خطا را نشان می‌دهد.

جدول ۶- آزمون مقایسه‌ی زوجی (میانگین قدر مطلق خطا)

مقایسه ی زوجی قدر مطلق خطا	Mean	Std. Deviation	Std. Error Mean	T	Df	Sig(2-tailed)
$Hybrid_{close}$–ANN_{close}	-10/99	13/95	2/54	-4/3	29	۰/000
$Hybrid_{close}$–$ARIMA_{close}$	-10/72	8/03	1/46713	-7/3	29	۰/000
$Hybrid_{down}$–ANN_{down}	-11/81	13/56	2/47	-4/7	29	۰/000
$Hybrid_{down}$-$ARIMA_{down}$	-5/81	7/47	1/36	-4/2	29	۰/000
$Hybrid_{up} - ANN_{up}$	-10/25	15/67	2/8۶	-3/5	29	۰/001
$Hybrid_{up}$ –$ARIMA_{up}$	-7/21	8/45	1/54	-4/6	29	۰/000

با توجه به جدول فوق بین میانگین دو جامعه‌ی MAE از مدل ترکیبی و مدل خودرگرسیو میانگین متحرک انباشته و دو جامعه‌ی MAE از مدل ترکیبی و مدل شبکه‌ی عصبی اختلاف معنا-داری وجود دارد و میانگین MAE برای مدل ترکیبی در نمونه‌ی مورد بررسی از دو مدل دیگر کمتر است. بنابراین، از نظر معیار ارزیابی MAE، عملکرد مدل ترکیبی به طور معناداری در حد

www.SID.ir

Arshive of SID

بالای قیمت، حد پایین قیمت و قیمت پایانی از دو مدل خودرگرسیو میانگین متحرک انباشته (ARIMA) و شبکه‌ی عصبی (ANN) بهتر است.

جدول ۷ نتایج آزمون دایبولد- ماریانو و آماره‌ی تعدیل شده هاروی – لیبورن- نیوبولد را با تابع زیان میانگین مجذور خطا نشان می‌دهد.

جدول ۷- آزمون دایبولد- ماریانو و آماره‌ی تعدیل شده آن

	Mean	دایبولد- ماریاونو	هاروی-نیوبولد- لیبورن
$Hybrid_{close}–ANN_{close}$	-540.96	-2.6517	-2.6071
$Hybrid_{close}–ARIMA_{close}$	-708.87	-2.3413	-2.3019
$Hybrid_{down}–ANN_{down}$	-449.34	-2.6262	-2.5820
$Hybrid_{down}–ARIMA_{down}$	-947.55	-3.1412	-3.0884
$Hybrid_{up} – ANN_{up}$	-1001.5	-2.6362	-2.5918
$Hybrid_{up} –ARIMA_{up}$	-571.44	-3.8778	-3.8126

با مقایسه آماره‌ی محاسبه شده در جدول۳ با مقادیر بحرانی جداول Z برای آزمون دایبولد- ماریانو و t-student برای آماره‌ی تعدیل شده هاروی – لیبورن- نیوبولد می‌توان نتیجه گرفت که، عملکرد مدل تلفیقی نسبت به هر یک از مدل‌های خودرگرسیو میانگین متحرک انباشته و شبکه‌ی عصبی در حد بالای قیمت، حد پایین قیمت و قیمت پایانی با توجه به معیار MSE دقیق‌تر است.

۶- نتیجه‌گیری و بحث

از آن‌جا که تشخیص الگوی داده‌ها مبنی بر خطی و غیرخطی بودن در واقعیت کمی دشوار است، بنابراین، باوجود انعطاف‌پذیری بالای مدل‌های خودرگرسیو میانگین متحرک انباشته (ARIMA) و شبکه‌ی عصبی در مدل‌سازی مسائل مختلف، نمی‌توان ادعا کرد که هر یک از این دو مدل برای تمام سری‌های زمانی مناسب می‌باشند.

در این پژوهش فرض شد که ترکیب مدل خودرگرسیو میانگین متحرک انباشته و مدل شبکه‌ی عصبی با در نظر گرفتن خصوصیات خطی ARIMA و غیرخطی شبکه‌ی عصبی می‌تواند منجر به افزایش دقت پیش‌بینی گردد. بنابراین در این پژوهش با استفاده حد بالا قیمت، حد پایین قیمت انس طلا در ابتدا به محسابه دو شاخص دامنه تغییرات قیمت و نقطه وسط که در بخش‌های قبل به تفصیل درباره این دو شاخص و علت استفاده از آن‌ها بحث شده است، پرداختیم. سپس سپس با

www.SID.ir

Arshive of SID

بهره‌گیری از ۲۷۳ مشاهده هفتگی به برازش مدل آریمای بهینه و پیش بینی بازده ۲۷۴ ام برای هر یک از سه سری زمانی بازده لگاریتمی دامنه تغییرات، حد وسط و قیمت پایانی پرداخته شد و نتایج به عنوان بازده ناشی از عملکرد مدل خطی لحاظ گردید. سپس با استفاده از شبکه‌های عصبی علاوه بر پیش‌بینی هر یک از سه سری زمانی مذکور، پسماندهای ناشی از مدل آریما بهینه مدل سازی و بازده‌های هر یک از سری زمانی مذکور و پسماند ناشی از مدل آریما برای هفته ۲۷۴ پیش‌بینی گردید. با استفاده از پسماند پیش بینی شده بازده پیش بینی شده برای مدل ترکیبی نیز برای هفته پیش رو محسابه گردید. این فرایند با استفاده از روش غلتاندن برای تمامی داده‌های خارج از نمونه برای هر سه مدل خطی، غیر خطی و ترکیبی تکرار گردید. سپس از طریق بازده های پیش بینی شده و روش تبدیل شاخص‌های دامنه تغییرات و نقطه وسط به حد بالا و پایین قیمت، قیمت‌های خارج از نمونه برای حد بالا ، حد پایین و قیمت پایانی تخمین زده شد. با استفاده از قیمت‌های تخمین زده شده و مقایسه این قیمت‌ها با مقادیر واقعی توابع زیان محاسبه شده و در نهایت با استفاده سه روش آزمون مقایسه زوجی، دایبولد-ماریانو و دایبولد-ماریانوی تعدیل شده به آزمون فرضیه‌های تحقیق دال بر مقایسه عملکرد مدل ترکیبی نسبت بر هر یک از مدل‌های آریما و شبکه عصبی چه در حد بالا قیمت، حد پایین قیمت و قیمت پایانی پرداخته شد.

در پیش‌بینی دامنه‌ی نوسان هفتگی اونس طلا نسبت به هر یک از مدل‌های ARIMA و شبکه‌ی عصبی با توجه معیارهای ارزیابی MSE و MAE نتایج دال بر تائید فرضیه‌های پژوهش مبنی بر عملکرد بهتری مدل ترکیبی نسبت به دو مدل دیگر در حدبالای قیمت، حدپایین قیمت و قیمت پایانی است.

اما برای پژوهش‌های آتی، با توجه در این پژوهش به منظور مدل‌سازی الگوی غیرخطی پسماند ها در مدل ترکیبی از شبکه‌ی عصبی پیش‌خور استفاده شده است، از این رو می‌توان به‌جای استفاده از شبکه‌ی عصبی از SVM برای مدل‌سازی بخش غیرخطی استفاده کرد و در نهایت عملکرد این مدل ترکیبی را با هر یک از مدل ARIMA و SVM مقایسه نمود.

www.SID.ir

Arshive of SID

فهرست منابع

* احمدی علی، ذوالفقاری مهدی، نژاد مهربانی آیدین، (۱۳۸۸)، مطالعه‌ی تطبیقی روش‌های ARIMA و شبکه‌ی عصبی مصنوعی در پیش‌بینی نیازهای داخلی برق کشور، فصل نامه ی پژوهش های اقتصادی ایران،۱۳(۴۱)، ۱۰۷-۱۲۱
* راعی رضا، (۱۳۷۷)، طراحی مدل سرمایه‌گذاری مناسب در سبد سهام با استفاده از شبکه‌ی عصبی مصنوعی، رساله ی دکتری، دانشکده ی مدیریت
* روشن رضا، (۱۳۸۵)، پیش‌بینی تورم ایران به کمک مدل‌های،ARIMA، ARCH/GARACH، شبکه‌های مصنوعی و مقایسه‌ی کارآیی مدل‌های مذکور. پایا نامه ی کار شناسی ارشد رشته ی علوم اقتصادی،دانشگاه سیستان و بلوچستان
* صادقی حسین، ذوالفقاری مهدی، الهامی نژاد مجتبی، (۱۳۹۰)، مقایسه‌ی عملکرد شبکه‌های عصبی و مدل ARIMA در مدل‌سازی و پیش‌بینی کوتاه‌مدت قیمت سبد نفت خام اوپک(با تاکید بر انتطارات تطبیقی)، فصل نامه ی مطالعات اقتصاد انرژی، ۸ (۲۸)، ۲۵-۴۷
* متوسلی محمود، طالب کاشفی بیژن، (۱۳۸۵)، بررسی مقایسه‌ای توان شبکه‌ی عصبی با ورودی شاخص‌های تحلیل تکنیکی برای پیش‌بینی قیمت سهام، نامه مفید، ۱۲(۵۴)، ۵۷-۸۲
* Brooks, c. (2008). Introductory Econometrics for Finance.
* Cadenas, E & River, W. (2010). Wind speed forecasting in three different region of mexico,using a hybrid ARIMA-ANN model. Renewable Enegy, 35(12), 2732-2738.
* Chen, K & Wang, C. (2007). A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. Expert System with Applications, 32(1), 254-264.
* Enders, W. (2009) Applied Econometric Time Series. Wiley
* Faruk, D. (2010) Ahybrid neural network and ARIMA model for water quality time series prediction. Engineering Application of Artificial Intligence, 23(4), 586-594
* Foster, W & Collopy, F & Ungar, L. (1992). Neural network forecasting of short, noisy time series. Computers and Chemical Engineering, 16(4), 293-297
* Indro, D & Jiang, C & Patuwo, B & Zhang, G. (1999). Predicting mutual fund performance using artificial neural networks. Omega The International Journal of Management Science, 27(3), 373-380.
* Kaastra, I & Boyd, M. (1996). Designing a neural network for forecasting financing and economic time-series. Neurocomputing, 10(3), 215-236.
* Klein, B & Rossin, D. (1999). Data quality in neural network models: effect of error rate and magnitude of error on predictive accuracy. Omega, The International Journal of ManagementScience, 27(5), 569-582.
* Lefebvre, M . (2007) Applied Stochastic processes.

www.SID.ir

Arshive of SID

* Markham, I & Rakes, T. (1998). The effect of sample size and variability of data on the comparative performance of artificial neural networks and regression. ComputOper Res, 25(4), 251-263.
* Pai, P & Liu, C. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. Omega, 33(6), 497-505.
* Sharda, R & Patill, R. (1990) Neural networks as forecasting exports: an empirical test. Proceedings of IJCNN Meeting, 2(1990), 491-494.
* Tang, Z & Fishwich, P. (1993). feed-forward neural nets as models for time series forcasting. ORSA Journal on Computing, 5(4), 374-385.
* Tseng, F & Yu, H & Tzeng, G. (2002). Combining neural network model with seasonal time series ARIMA model. Technol Forecast Soc Change , 69(1), 71-87.
* Wedding, D & Cios, K. (1996). Time series forecasting by combining RBF networks certainty factors and the Box-Jenkins model. Neurocomputing , 10(2), 149-168.
* Zhang, G & Patuwo, E & Hu, M.(1998) Forecasting with artifcial neural networks: the state of the art. Int J Forecasting, 14(1), 35-62.
* Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network Model. Neurocomputing , 50, 159-175

یادداشت‌ها

[1]. Interval Analysis
[2]. Box-Jenkins
[3]. Universal
[4]. Artificial Neural Network
[5]. Noise
[6]. Structural models
[7]. Autoregressive processes
[8]. Lag operator
[9]. Moving average processes
[10]. Autoregressive moving average
[11]. Invariant
[12]. Strict- Sense Stationary (SSS)
[13]. Random wall with drift
[14]. trend- stationary processes
[15]. stochastic non- stationary
[16]. Deterministic non- stationary
[17]. Diffrence
[18]. Autoregressive Integrated moving average
[19]. Identification
[20]. Estimation
[21]. Diagnostic checking
[22]. Dendrite
[23]. Soma or Cell Body
[24]. Axon
[25]. Synapse
[26]. Multi- layer Feed Forward
[27]. pure line
[28]. Seosonal autoregressive integrated moving average

www.SID.ir

Arshive of SID

[29]. Support Vector Machines
[30]. Buyok Menderes
[31]. Interval value
[32]. Range of Interval
[33]. Mid point
[34]. Diebold-Mariano

www.SID.ir