

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با تاخیر زمانی در پیش‌بینی سری‌های زمانی مالی

محمدرضا اصغری اسکویی*

تاریخ پذیرش: ۱۳۹۴/۴/۲۴

تایخ دریافت: ۱۳۹۳/۶/۱

چکیده

شبکه عصبی با تاخیر زمانی، یک ابزار مدل‌سازی برگرفته از محاسبات هوشمند است که در کنار روش‌های کلاسیک برای پیش‌بینی سری‌های زمانی مالی بکار گرفته می‌شود. این مدل اغلب در مواردی که از سری زمانی داده‌های فراوان، اما از ساختار مدل اطلاعات محدود وجود دارد، استفاده می‌شود، از این رو انتخاب ساختار و ارزیابی آن خود یک چالش است. در این مقاله یک مدل مبتنی بر شبکه عصبی با تاخیر زمانی برای پیش‌بینی معیارهای مالی بازار سهام ارائه شده و روش پنجره لغزان برای ارزیابی عملکرد پیش‌بینی‌کننده، بکار برده شده است. در این مقاله انتخاب ساختار مناسب شبکه، تعداد عملگرهای تاخیری، تعداد بهینه داده‌های پیشین و پسین و معیار کمی مناسب برای ارزیابی عملکرد پیش‌بینی‌کننده، مورد مطالعه قرار گرفته است. عملکرد این مدل روی شاخص قیمت سهام چهار بنگاه بزرگ اقتصادی در بازار سهام لندن مورد بررسی قرار گرفته است. نتایج به دست آمده، نشان می‌دهد با روش پیشنهادی، می‌توان ساختاری انتخاب کرد که متوسط درصد خطا، متوسط مجذور مربعات خطا و معیار رگرسیون خطی خروجی شبکه در حد قابل توجهی کاهش می‌یابد.

طبقه‌بندی JEL: C6, C8.

کلیدواژه‌ها: پیش‌بینی سری زمانی، شبکه عصبی با تاخیر زمانی، پنجره لغزان، معیار خطا

پیش‌بینی.

* استادیار گروه رایانه، دانشکده علوم ریاضی و رایانه، دانشگاه علامه طباطبائی، پست الکترونیکی:

۱- مقدمه

شناخت ماهیت تغییرات در بازارهای مالی و پیش‌بینی آن همواره برای صاحبان بنگاه‌های اقتصادی، سرمایه‌گذاران، مشتریان و حتی مردم عادی حائز اهمیت ویژه بوده و توجه بسیاری از پژوهشگران را به خود جلب کرده است. داشتن تحلیل درست از روند تغییرات و نیز امکان پیش‌بینی آن، نقش بسیار مهمی در برنامه‌ریزی و تصمیم‌سازی دارد.

مطالعه و مدل‌سازی تغییرات و پیش‌بینی معیارهای مالی با دو رویکرد متفاوت بنیادین (توصیفی) و فنی (تکنیکال) انجام می‌گیرد. رویکرد بنیادین مبتنی بر مفاهیم و نظریه‌های اقتصادی و مالی است. در این رویکرد، رابطه متغیرهای مختلف مورد مطالعه قرار داده و با استخراج مدل ریاضی، معیارهای مالی و یا اقتصادی را پیش‌بینی می‌کنند.

رویکرد فنی (تکنیکال) بدون توجه عمیق به ماهیت معیارهای مالی و فقط با مطالعه روند تغییرات در دوره‌های گذشته، مبادرت به پیش‌بینی تغییرات آتی می‌کند. در این رویکرد، مقادیر معیارهای مالی که در فواصل زمانی یکسان از بازارهای مالی جمع‌آوری و ثبت شده است را به صورت سری زمانی در نظر گرفته و روند تغییرات آن را در بازه زمانی معین مورد مطالعه قرار داده و مقادیر آتی را پیش‌بینی می‌کنند. این رویکرد در سال‌های اخیر با توجه به پیشرفت سامانه‌های محاسباتی و ورود روش‌های محاسبات هوشمند مورد اقبال متخصصان و پژوهشگران قرار گرفته و نتایج بسیار چشمگیری نیز داشته است.

سری زمانی مالی اغلب شامل قیمت اقلام، سهام و برخی معیارهای مرتبط همچون حجم سهام مبادله شده است. از لحاظ ساختاری، سری زمانی مالی می‌تواند متشکل از مولفه‌های روند کلی^۱، دوره‌ای^۲ و تصادفی^۳ باشد. مولفه روند کلی، تغییرات سری زمانی را در دوره طولانی نشان می‌دهد در حالی که مولفه دوره‌ای، تغییرات نوسانی در یک دوره مشخص را نشان می‌دهد. هر دو مولفه الگوی معین^۴ ریاضی دارند و قابل توصیف و

1- Trend

2- Seasonal

3- Stochastic

4- Deterministic

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۷۷

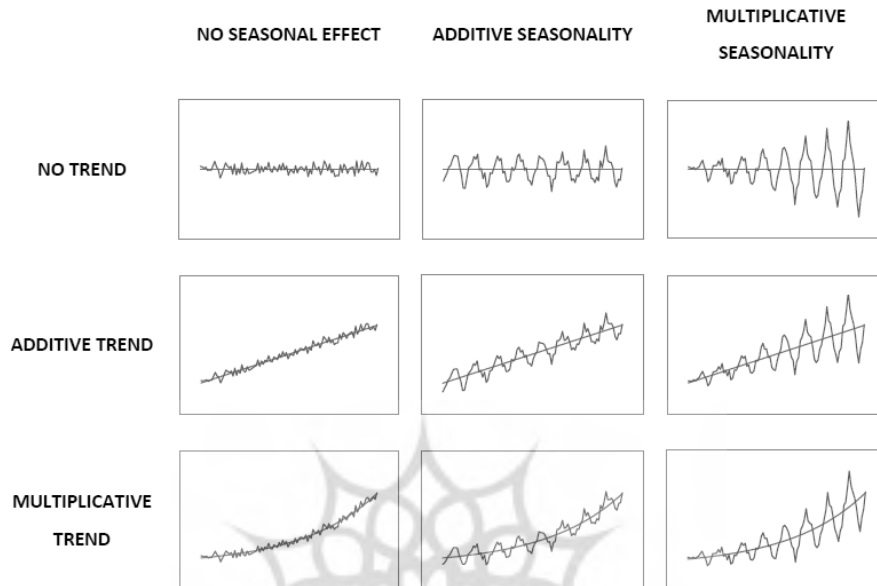
مدلسازی هستند. مولفه‌ای که مقادیر آن قابل پیش‌بینی نباشد، تصادفی تلقی می‌شود و معمولاً توزیع احتمالی این مولفه در مدل ریاضی لحاظ می‌شود. منظور از مولفه تصادفی، عدم وجود اطلاعات کامل و استنباط رابطه دقیق وابستگی داده‌ها در گذشته و آینده است. سه ترکیب ساده، تجمیعی و هندسی از مولفه‌های سری زمانی در شکل (۱) نشان داده شده است. منظور از سری زمانی غیرایستا^۱ این است که توزیع احتمال و مشخصه‌های آماری مولفه سری زمانی نسبت به زمان متغیر است. پیش‌بینی یک سری زمانی مالی که عموماً غیرایستا و همراه با نوفه (نویز) است به لحاظ فنی بسیار چالش برانگیز است. هدف ما در این مقاله ارائه یک مدل مبتنی بر هوش محاسباتی است که قدرت پیش‌بینی این نوع سری زمانی را داشته باشد.

مطالعات فراوانی پیرامون استفاده از هوش محاسباتی در کنار روش‌های کلاسیک، برای مدلسازی و پیش‌بینی سری زمانی انجام گرفته است. از جمله می‌توان به استفاده از شبکه عصبی پیش‌رو، شبکه عصبی بازگشتی، سیستم استنتاج فازی، سیستم خبره، سیستم ترکیبی فازی-عصبی، بردار حائل و سیستم آشوبگونه اشاره کرد.

در این مقاله، مدل شبکه عصبی ارائه شده در رابطه (۱) را با افزودن عملگر تاخیر زمانی و بازخورد (شبکه عصبی بازگشتی نوع اول) توسعه داده و ضمن معرفی روش پنجره لغزان، کاربرد آن برای پیش‌بینی شاخص قیمت چهار شرکت در بازار بورس مورد مطالعه قرار می‌دهیم.

این مقاله از چهار بخش تشکیل شده است؛ بخش اول مقدمه و بخش دوم معرفی مدل ریاضی شبکه عصبی بازگشتی برای پیش‌بینی سری زمانی و بخش سوم مشتمل بر معرفی روش پنجره لغزان و مراحل پیاده‌سازی است. نتایج در بخش چهارم و جمع‌بندی در بخش پنجم ارائه شده است.

شکل (۱) - ترکیب ساده، تجمیعی و هندسی از مولفه‌های اصلی در یک سری زمانی



۲- مدل ریاضی شبکه عصبی برای پیش‌بینی سری زمانی

سری زمانی مجموعه از داده‌های متوالی است که در فواصل زمانی مساوی اندازه‌گیری و ثبت شده‌اند. برای مثال، قیمت یک سهام در روزهای متوالی، ارتفاع سطح آب دریاچه در ساعات مختلف، فروش هفتگی یک بنگاه اقتصادی و جمعیت یک کشور نمونه‌هایی از سری زمانی هستند. منظور از پیش‌بینی متغیر، تعیین مقادیر آتی آن برای یک دوره مشخص است.

پیش‌بینی یک سری زمانی عبارت از ساخت یک مدل از داده‌های موجود برای پیش‌بینی مقادیر آتی است. اساس مدل پیش‌بینی سری زمانی بر این باور نهاده شده است که از مقادیر پیشین یک متغیر می‌توان برای پیشگویی مقادیر آتی استفاده کرد از این رو هدف، شناسایی الگوی حاکم بر مقادیر پیشین و تسری آن به مقادیر پسین است.

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۷۹

روش‌های پیش‌بینی به دو گروه کمی و کیفی تقسیم می‌شوند؛ روش‌های کمی خود به دو مدل بنیادین (توصیفی، علی) و فنی (تکنیکال، سری زمانی) تقسیم می‌شوند. در مدل توصیفی، رابطه متغیر مورد نظر بر اساس یک نظریه و یا رابطه علی با متغیرهای توصیفی دیگر بیان می‌شود. یکی از مدل‌های شناخته شده توصیفی مدل رگرسیون است. انتخاب تابع ریاضی مناسب و تخمین پارامترهای آن، دو مرحله بسیار مهم در فرآیند ساخت مدل رگرسیون است. معمولاً برای تخمین پارامترهای مدل از روش حداقل مربعات^۱ و یا ماکزیمم درست‌نمایی^۲ استفاده می‌شود.

در مدل‌های پیچیده و غیرخطی از روش‌های پیشرفته همچون شبکه بیزی و یا شبکه عصبی استفاده می‌شود. پیش‌بینی سری زمانی نیز می‌تواند در این دسته‌بندی قرار گیرد، تنها با این تفاوت که مقدار آتی متغیر را فقط بر اساس مقادیر قبلی آن و بدون توجه به عوامل تاثیرگذار بیرونی تعیین می‌کند. البته تمام روش‌های پیش‌بینی منحصر به دو گروه کمی و کیفی نمی‌شوند. برخی روش‌ها هم وجه کمی و هم وجه کیفی دارند. از جمله روش مبتنی بر شبیه‌سازی (مثلاً مونت کارلو)، روش مبتنی بر قواعد و روش نظریه بازی را می‌توان نام برد که در رویکرد پیش‌بینی از وجوه کیفی و کمی هر دو برخوردارند.

۲-۱- انواع مدل‌های پیش‌بینی سری زمانی

روش‌های پیش‌بینی سری زمانی به چهار گروه تقسیم می‌شوند که شامل روش‌های پیش‌بینی ساده، پیش‌بینی روند کلی، پیش‌بینی روند کلی- دوره‌ای و سرانجام پیش‌بینی پیشرفته است.

روش‌های پیش‌بینی ساده مبتنی بر تکنیک‌های پایه است. این روش‌ها قابل کاربرد برای هر نوع سری زمانی هستند، اما قادر به ارائه پیش‌بینی دقیق در مورد یک سری زمانی با روند تغییرات کلی و یا دوره‌ای نیست. برای مثال، پیش‌بینی بر اساس آخرین داده، متوسط

1- Least Squared Error

2- Maximum Likelihood

تعدادی از آخرین داده‌ها، متوسط کل داده‌ها و یا متوسط وزن‌دار داده‌ها از جمله این نوع روش‌ها است.

روش‌های پیش‌بینی روند کلی برای سری زمانی بکار می‌رود که تغییرات کلی آن قالب و تعیین‌کننده است. برای مثال، روش خطی تعدیل‌نمایی هولت^۱ در شناسایی روند تغییرات کلی سری زمانی بسیار مناسب عمل می‌کند. در این روش دو معادله تعدیل‌یک‌ی برای تخمین سطح تغییرات و یکی برای روند تغییرات در زمان مورد نظر ارائه می‌شود. تحلیل رگرسیونی نیز که در آن رابطه خطی و یا غیرخطی بین داده‌های سری زمانی (متغیر وابسته) و زمان (متغیر ناپسته) مورد بررسی قرار می‌گیرد، نمونه دیگری از این روش‌ها است.

تحلیل رگرسیونی تخمین مناسبی از روند تغییرات کلی سری زمانی ارائه می‌دهد. شناسایی روند کلی امکان پیش‌بینی مقادیر آتی سری زمانی را در فاصله دور می‌دهد. روش‌های پیش‌بینی روند کلی و دوره‌ای در موارد پیچیده که علاوه بر تغییرات کلی تغییرات دوره‌ای نیز در سری زمانی داریم مورد استفاده قرار می‌گیرد. روش وینتر-هولت^۲ حالت تعمیم‌یافته از روش هولت است که عامل سومی نیز اضافه می‌کند که تغییرات دوره‌ای را تخمین می‌زند.

روش‌های پیش‌بینی پیشرفته برای سری زمانی، هر سه مولفه روند کلی، دوره‌ای و تصادفی را در خود دارند. از جمله روش‌های پیشرفته می‌توان مدل باکس-جنکینز^۳، روش شینسکن، داده کاوی، بیزین و محاسبات هوشمند را نام برد. روش باکس-جنکینز که در سال ۱۹۷۰ ارائه شد، وابستگی داده‌ها در فواصل مشخص را مورد مطالعه قرار داده و مدلی پارامتری برای آنها ارائه می‌دهد.

روش شینسکن سه مولفه اصلی (روند کلی، دوره‌ای و تصادفی) را شناسایی و هر یک را جداگانه تخمین می‌زند. روش داده کاوی با تحلیل آماری حجم زیادی از داده‌ها و

1- Holt Linear Exponential Smoothing

2- Winter-Holt

3- Box-Jenkins Model

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۸۱

کاربرد یادگیری ماشین، الگوی را برای پیش بینی استنباط می کند. روش بیزین یک مدل پارامتری احتمالی بر اساس توزیع داده های پیشین ارائه می دهد.

مزیت روش بیزین این است که تابع توزیع احتمالی پیش بینی را نیز که ناشی از عدم قطعیت پارامترهای احتمالی است، ارائه می دهد. روش های هوشمند می توانند جایگزین کارآمدی برای روش های آماری در پیش بینی سری زمانی طولانی، پیچیده و غیرخطی باشند. از جمله روش های هوشمند می توان شبکه عصبی، سامانه استنتاج فازی، الگوریتم ژنتیک و حالت های ترکیبی آنها را نام برد. در ادامه مبانی نظری مدل شبکه عصبی با تاخیر زمانی - که یکی از روش های هوش محاسباتی است - با مدل باکس-جنکینز که یک روش کلاسیک و شناخته شده است، مقایسه و فرمول های ریاضی آن ارائه می شود.

۲-۲- مدل شبکه عصبی

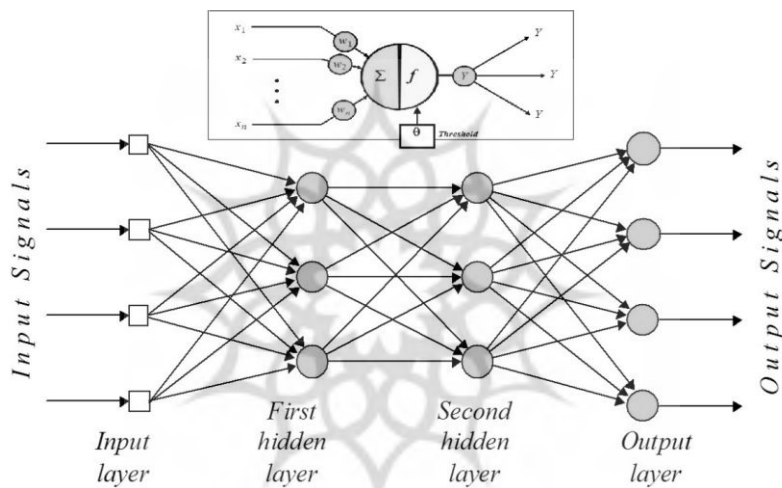
شبکه عصبی مصنوعی ساختاری محاسباتی است که از سیستم عصبی انسان الگو گرفته و مهم ترین آن قدرت یادگیری است. فرآیند یادگیری در شبکه های عصبی از نوع یادگیری با الگو^۱ است. در این نوع یادگیری، ضرائب داخلی شبکه عصبی با کمک مجموعه نمونه هایی که به صورت زوج مرتب (X_i, Y_i) هستند، محاسبه می شوند. شبکه عصبی، رابطه بین داده های ورودی (X_i) و داده های خروجی (Y_i) را که به صورت تابع $Y_i=f(X_i)$ است، تقریب می کند. از این نوع یادگیری در شبکه عصبی برای پیش بینی سری های زمانی استفاده می شود. در این رویکرد، مجموعه داده های ثبت شده از یک سری زمانی به عنوان الگو در آموزش شبکه عصبی بکار گرفته می شود. شبکه عصبی پس از یادگیری روند تغییرات سری زمانی، مقادیر آتی را پیش بینی می کند.

شبکه عصبی مصنوعی متشکل از تعدادی نرون محاسباتی است که در یک ساختار شبکه چند لایه در کنار هم قرار می گیرند. مدل نرون محاسباتی که پرسپترون نامیده

می شود، مانند یک سلول عصبی زیستی، شامل تعدادی ورودی، تابع فعالیت و یک خروجی است.

ساختار پرسپترون و شبکه چندلایه پرسپترون^۱ در شکل (۲) نمایش داده شده است. از لحاظ نظری، شبکه عصبی قدرت تقریب هر تابع غیرخطی را دارد و با وجود نقاط قوت از چند نقطه ضعف نیز برخوردار است، از جمله اینکه تعداد پارامترهای این شبکه بسیار زیاد است، یافتن پاسخ پایدار مشکل است و همواره خطر بیش‌برازش در یادگیری شبکه وجود دارد.

شکل (۲) - ساختار شبکه عصبی پیشرو چند لایه پرسپترون و ساختار درون یک پرسپترون



۲-۳- مدل خطی ARMA

روش باکس-جنکینز، روش آماری و کلاسیک برای مدل سازی و پیش بینی سری زمانی است. در این روش، ابتدا با کمک داده های موجود یک مدل پارامتریک، مبتنی بر معادلات تفاضلی-خطی و یک ورودی تصادفی، ساخته می شود، سپس این مدل برای پیش بینی مقادیر آتی مورد استفاده قرار می گیرد. یکی از مدل های شناخت شده، مدل خود

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۸۳

افزای میانگین متحرک^۱ یا $ARMA$ است. فرمول مدل خطی $ARMA(p, q)$ با فرض اینکه $E(e_t | x_{t-1}, x_{t-2}, \dots) = 0$ است، در رابطه (۱) نشان داده شده است:

$$x_t = c + \sum_{i=1}^p \alpha_i x_{t-i} + \sum_{j=1}^q b_j e_{t-j} + e_t \quad (1)$$

e_t خطایی با میانگین صفر، توزیع یکنواخت و مستقل از مقادیر پیشین x_t و اینکه واریانس خطا همواره کراندار و برابر σ^2 است. برای سادگی x_t را با میانگین صفر ($c=0$) در نظر می‌گیریم، این معادل آن است که x_t را در حالت کلی با $x_t - E(x_t)$ جایگزین کرده باشیم.

پیش‌بینی بهینه سری زمانی با مدل $ARMA(p, q)$ بر اساس معیار حداقل میانگین مربعات خطا و با فرض داشتن مقادیر نامتناهی از داده‌های پیشین سری زمانی، برابر با امید شرطی رابطه (۲) خواهد بود.

$$\hat{x}_t = E(e_t | x_{t-1}, x_{t-2}, \dots) \quad (2)$$

در واقعیت، ما همواره تعداد متناهی از مشاهدات پیشین داریم، بنابراین پیش‌بینی با یک تخمین همراه است. اثبات شده است که تخمین پیش‌بینی بهینه مدل $ARMA(p, q)$ بر اساس حداقل میانگین مربعات خطا به روش بازگشتی رابطه (۳) قابل محاسبه است.

$$\hat{x}_t = \sum_{i=1}^p a_i x_{t-i} + \sum_{j=1}^q b_j \hat{e}_{t-j} + e_t \quad (3)$$

البته با فرض اینکه خطای $e_{t,j}$ و شرایط آغازین به صورت رابطه (۴) تعریف شوند.

$$\forall k \leq 0 \quad \hat{x}_k = \hat{e}_k = 0 \quad \text{و} \quad \hat{e}_{t-j} = x_{t-j} - \hat{x}_{t-j} \quad j = 1, 2, \dots, q \quad (4)$$

حالت خاص از مدل کلی فوق، مربوط به مدل خودافزا $AR(p)$ می‌باشد که در رابطه (۵) آمده است:

1- Auto Regressive Moving Average (ARMA)

$$x_t = \sum_{i=1}^p a_i x_{t-i} + e_t \quad (5)$$

تخمین پیش‌بینی بهینه مدل خودافزا رابطه (۵) عبارت است از: $\hat{x}_t = \sum_{i=1}^p a_i x_{t-i}$

۲-۴- مدل شبکه عصبی پیش‌رو معادل مدل غیرخطی AR

یک شبکه عصبی پیش‌رو معادل مدل AR غیرخطی است. به بیان دیگر، شبکه عصبی پیش‌رو می‌تواند مدل غیرخطی AR را به لحاظ محاسباتی تقریب کند. مدل غیرخطی $AR(p)$ را که $NAR(p)$ می‌نامیم، مطابق رابطه (۶) است.

$$x_t = h(x_{t-1}, x_{t-2}, \dots, x_{t-p}) + e_t \quad (6)$$

البته با فرض اینکه $E(e_t | x_{t-1}, x_{t-2}, \dots) = 0$ و واریانس خطا کراندار و برابر 2 باشد. در این مدل تخمین پیش‌بینی بهینه x_t با فرض داشتن p مقدار پیشین به صورت رابطه (۷) است. حداکثر خطای این پیش‌بینی برابر 2 است.

$$\hat{x}_t = E(x_t | x_{t-1}, x_{t-2}, \dots, x_{t-p}) = h(x_t | x_{t-1}, x_{t-2}, \dots, x_{t-p}), t \geq p+1 \quad (7)$$

چنانچه اشاره شد، شبکه عصبی چندلایه پرسپترون (MLP) می‌تواند تابع $h(\cdot)$ را تقریب کند، از این رو می‌توان رابطه (۸) را نوشت:

$$\hat{x}_t = h(x_{t-1}, x_{t-2}, \dots, x_{t-p}) = \sum_{i=1}^I W_i f\left(\sum_{j=1}^p W_{ij} x_{t-j} + \theta_i\right) \quad (8)$$

یک شبکه MLP با یک پرسپترون خطی در لایه خروجی، تعداد I پرسپترون با تابع فعالیت $f(\cdot)$ در لایه پنهان و تعداد p ورودی، معادله (۸) را محقق می‌کند. تابع فعالیت $f(\cdot)$ یک تابع کراندار، مشتق‌پذیر و مونوتونیک است که اغلب، تابع سیگموئید برای این منظور استفاده می‌شود. شکل (۳) ساختار یک شبکه MLP با یک خروجی را نمایش می‌دهد.

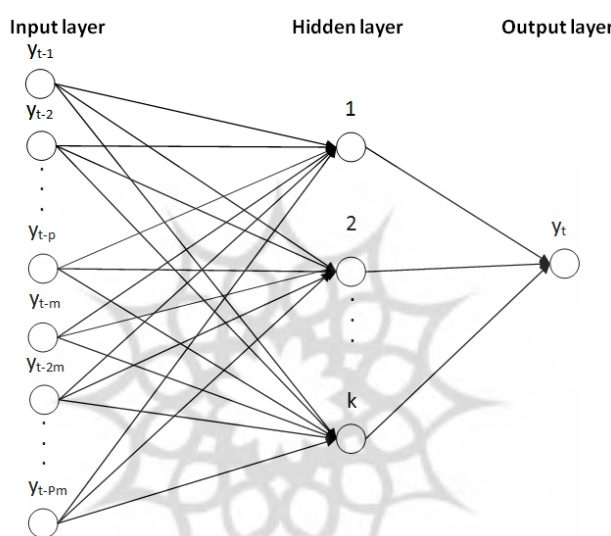
ضرایب وزنی W_i و w_{ij} به ترتیب مربوط به پرسپترون لایه خروجی و پرسپترون های لایه پنهان هستند که در فرآیند آموزش شبکه عصبی با کمک مجموعه داده‌های آموزشی

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۸۵

محاسبه می‌شوند. محاسبه ضرایب وزنی، بر اساس یافتن حداقل میانگین مربعات خطا (MSE) با کمک الگوریتم پس انتشار خطا انجام می‌گیرد. میانگین مربعات

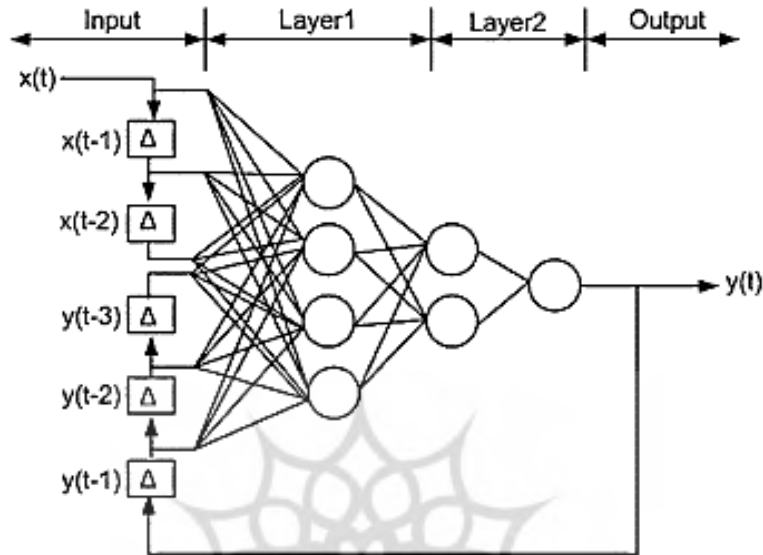
$$\text{خطا عبارت از } \frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_t)^2 \text{ است.}$$

شکل (۳) - ساختار یک شبکه عصبی MLP برای پیش‌بینی سری زمانی با مدل NAR



با توجه به اینکه ورودی شبکه MLP ، مقادیر متوالی سری زمانی است، می‌توان با قراردادن عملگر تاخیری در ورودی شبکه، ساختار شبکه عصبی با تاخیر زمانی^۱ ($TDNN$) را که برای اولین بار توسط ویبل و همکاران معرفی شده است، برای مدل NAR بکار برد. شکل (۴) ساختار شبکه عصبی با تاخیر زمانی را نشان می‌دهد. این مدل برای تقریب مدل غیرخطی خودافزا (NAR) است، از این رو، آن را مدل $NAR-TDNN$ می‌نامیم.

شکل (۴) - ساختار شبکه عصبی پیشرو چند لایه پرسپترون با تاخیر زمانی



۲-۵- مدل شبکه عصبی بازگشتی معادل مدل غیرخطی ARMA

حالت کلی مدل باکس-جنکینز مدل غیرخطی ARMA است که اصطلاحاً NARMA نامیده می‌شود. مدل $NARMA(p, q)$ به صورت رابطه (۹) است.

$$x_t = h(x_{t-1}, x_{t-2}, \dots, x_{t-p}, e_{t-1}, e_{t-2}, \dots, e_{t-q}) + e_t \quad (9)$$

با فرض اینکه، تابع $h(\cdot)$ یک تابع نامشخص و هموار، امید خطا $E(e_t | x_{t-1}, x_{t-2}, \dots) = 0$ و اینکه واریانس خطای e_t کراندار و برابر با σ^2 باشد. امید شرطی پیش‌بینی سری زمانی بر اساس تعداد نامتناهی از مشاهدات پیشین به صورت رابطه (۱۰) تعریف می‌شود:

$$\hat{x}_t = E(h(x_{t-1}, x_{t-2}, \dots, x_{t-p}, e_{t-1}, e_{t-2}, \dots, e_{t-q}) | x_{t-1}, x_{t-2}, \dots) \quad (10)$$

با فرض اینکه مدل NARMA مدلی معکوس‌پذیر است تابع $g(\cdot)$ وجود دارد که به شکل رابطه (۱۱) است.

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۸۷

$$x_t = g(x_{t-1}, x_{t-2}, \dots) + e_t \quad (11)$$

با داشتن مشاهدات قبلی، امکان محاسبه مقدار e_{t-j} همواره وجود دارد. بنابراین امید شرطی پیش‌بینی مدل $NARMA$ به صورت رابطه (۱۲) به دست می‌آید و خطای تخمین این پیش‌بینی برابر σ^2 خواهد بود.

$$\hat{x}_t = h(x_{t-1}, x_{t-2}, \dots, x_{t-p}, e_{t-1}, e_{t-2}, \dots, e_{t-q}) \quad (12)$$

در عمل چون فقط تعداد متناهی از مشاهدات قبلی وجود دارد، تخمین پیش‌بینی مدل $NARMA$ بر اساس مقادیر متناهی و با کمک محاسبات بازگشتی به صورت روابط (۱۳) و (۱۴) به دست می‌آید.

$$\hat{x}_t = h(x_{t-1}, x_{t-2}, \dots, x_{t-p}, \hat{e}_{t-1}, \hat{e}_{t-2}, \dots, \hat{e}_{t-q}) \quad (13)$$

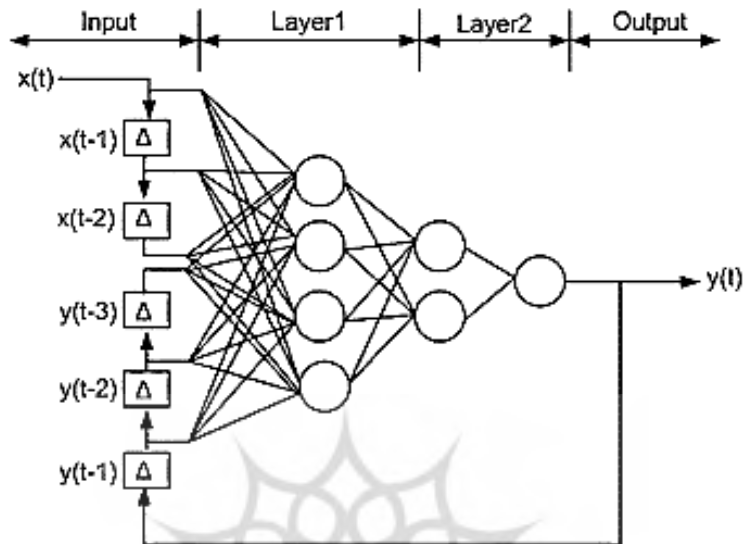
$$\forall k \leq 0 \quad \hat{x}_k = \hat{e}_k = 0 \quad \text{و} \quad \hat{e}_{t-j} = x_{t-j} - \hat{x}_{t-j} \quad h = j = 1, 2, \dots, q \quad (14)$$

با استفاده از شبکه عصبی بازگشتی^۱ می‌توان تخمین پیش‌بینی مدل $NARMA$ را که در معادلات (۱۳) و (۱۴) نمایش داده شده است، محاسبه کرد. خروجی شبکه عصبی بازگشتی در رابطه (۱۵) نشان داده می‌شود.

$$\hat{x}_t = \sum_{i=1}^I w_i f\left(\sum_{j=1}^p w_{ij} x_{t-j} + \sum_{j=1}^q w'_{ij} (x_{t-j} - \hat{x}_{t-j}) + \theta_i\right) \quad (15)$$

یک نمونه از ساختار شبکه عصبی بازگشتی در شکل (۵) نمایش داده شده است. ضرایب وزنی w_i, w_{ij}, w'_{ij} همانند شبکه MLP پیشرو با روش حداقل مربعات و یا هدف کمینه شدن معیار $\sum (x_t - \hat{x}_t)^2$ محاسبه می‌شوند. در ساختار توپولوژیک شبکه عصبی بازگشتی، علاوه بر اتصالات پیشرو (از ورودی به سمت خروجی)، اتصالات بازخور (از خروجی به سمت ورودی) نیز وجود دارد. این نوع شبکه در برخی کتب به شبکه حلقه بسته یا بسته شناخته می‌شود.

شکل (۵) - ساختار شبکه عصبی بازگشتی با تاخیر زمانی



در یک تقسیم‌بندی کلی، دو نوع شبکه بازگشتی وجود دارد؛ ساده و استاندارد. عملکرد شبکه ساده به این ترتیب است که به ازای هر تغییر یا تحریک در ورودی، فضای حالت شبکه از یک نقطه شروع و پس از مدتی در یک نقطه تعادل پایان می‌یابد. این ویژگی در شبکه عصبی برای کاربردهایی همچون طبقه‌بندی کننده‌ها ایده‌آل است.

در حالت کلی، همواره این امکان وجود دارد که شبکه هرگز به نقطه تعادل نرسد یا به عبارتی همگرا نباشد که این حالت برای مدلسازی مطلوب نیست. با اعمال محدودیت‌ها به اتصالات پس‌خور می‌توان همگرایی شبکه به نقطه تعادل را تضمین کرد.

شبکه عصبی بازگشتی استاندارد برخلاف نوع ساده که تنها به دنبال رسیدن به نقطه تعادل است، سعی می‌کند تا یک مسیر حرکت مشخصی را در فضای حالت تعقیب و مدل آن را پیدا کند. این ویژگی در پیش‌بینی سری زمانی بسیار مهم است، چراکه تغییرات در ورودی شبکه لزوماً همواره به یک نقطه تعادل ثابت ختم نمی‌شود. در عمل شبکه‌های

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۸۹

بازگشتی استاندارد برای پیش بینی بکار برده می شود، چراکه امکان مدل سازی و تقریب متغیرها در محدوده پیوسته را فراهم می کنند.

شبکه عصبی بازگشتی از لحاظ توپولوژیک نیز به سه گروه تقسیم می شود؛ در نوع اول، اتصالات بازخور فقط از خروجی شبکه به ورودی آن، در نوع دوم بازخورها از خروجی پرسپترون به ورودی آن و در نوع سوم اتصالات بازخور از خروجی پرسپترون به ورودی هر پرسپترون دیگر است.

مدلی که در معادله (۱۵) توصیف شده، یک شبکه بازگشتی استاندارد نوع اول است که آن را مدل بازگشتی *NARMA-RNN* می نامیم. این مدل حالت خاصی از شبکه عصبی بازگشتی کامل^۱ (*NARMA-FRNN*) است. شبکه عصبی بازگشتی کامل از نوع سوم بوده و مدل ریاضی آن در روابط (۱۶) و (۱۷) تعریف شده است.

$$\hat{x}_t = \sum_{i=1}^I w_i g_i(t) \quad (16)$$

$$\sum_{j=1}^{\max(p,q)} \tilde{w}_{ij} x_{t-j} + \left(\sum_{k=1}^q \sum_{l=1}^I \tilde{w}_{ilk} g_l(t-k) \right) + \theta_i \quad (17)$$

مدل بازگشتی *NARMA-RNN* حالت خاصی از مدل بازگشتی کامل *NARMA-FRNN* است که در آن روابط (۱۸) و (۱۹) برقرار است. با توجه به مشکل پایداری در شبکه بازگشتی کامل در این مقاله از مدل بازگشتی *NARMA-RNN* استفاده می شود.

$$\tilde{w}_{ilk} = w_{ik} W_l, \quad 1 \leq i, l \leq I, \quad 1 \leq k \leq q \quad (18)$$

$$\tilde{w}_{ij} = \begin{cases} w_{ij} + w'_{ij} & 1 \leq j \leq q \\ w_{ij} & q+1 \leq j \leq p \end{cases}, \quad p < q \quad (19)$$

1- Fully Connected Recurrent Neural Network (FRNN)

۶-۲- خطای پیش‌بینی سری زمانی

اطلاع از دقت پیش‌بینی برای یک سازمان به اندازه خود پیش‌بینی حائز اهمیت است. تحلیلگر نه تنها باید پیش‌بینی خوب و به موقع داشته باشد، بلکه باید از میزان حداکثر خطا نیز تخمین صحیح داشته باشد. داشتن تخمین صحیح از خطا کمک می‌کند تا در هر تصمیم انعطاف لازم برای تفاوت احتمالی بین مقدار پیش‌بینی شده و واقعی وجود داشته باشد.

عملکرد مدل پیش‌بینی از مقایسه مقادیر تولید شده توسط مدل با مقادیر واقعی محاسبه می‌شود. عملکرد مدل از دو جنبه مورد مطالعه قرار می‌گیرد؛ اول به کمک داده‌های پیشین که در ساخت مدل مورد استفاده بودند و دوم ارزیابی به کمک داده‌های پسین که در ساخت مدل مورد استفاده نبودند. در فرآیند یادگیری ماشین، اولی داده‌های آموزش و دومی داده‌های آزمون نامیده می‌شوند.

از دیدگاه سری زمانی، داده‌های آموزش مجموعه داده‌های متوالی مشاهده شده از ابتدای سری زمانی تا لحظه پیش‌بینی و داده‌های آزمون مجموعه داده‌های متوالی بعد از لحظه پیش‌بینی هستند. ارزیابی عملکرد مدل پیش‌بینی با مجموعه داده‌های آموزش معیاری از برازش^۱ مدل و ارزیابی عملکرد مدل پیش‌بینی با مجموعه داده‌های آزمون معیار تعمیم‌پذیری^۲ مدل هستند.

بیش‌برازش^۳ یک پدیده نامطلوب در یادگیری ماشین است که در آن درجه آزادی مدل بسیار بیشتر از درجه آزادی واقعی، انتخاب شده و در نتیجه اگرچه مدل روی داده‌های آموزشی بسیار خوب نتیجه می‌دهد، اما روی داده‌های آزمون دارای خطای زیاد است. به عبارت دیگر، بیش‌برازش یک مدل، معیار برازش مناسب، اما معیار تعمیم‌پذیری نامناسب می‌شود. انتخاب درجه آزادی مناسب به کمک واریانس اعتبار^۴ و تنظیم کردن^۵ مدل از

1- Fitness

2- Generalization

3- Over Fitting

4- Cross-validation

5- Regularization

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۹۱

راه‌های مقابله با این پدیده ناخواسته است. هدف اصلی در پیش‌بینی، یافتن مدلی با حد برازش و تعمیم‌پذیری مطلوب است.

در یک مدل مبتنی بر شبکه عصبی، ابتدا ساختار مناسب شبکه انتخاب می‌شود و سپس ضرایب وزنی شبکه با استفاده از الگوریتم یادگیری پس‌انتشار^۱ خطا به صورت محاسبات عددی و تکراری روی داده‌های آموزشی محاسبه می‌شوند. انتخاب ساختار مناسب چالشی است که می‌تواند منجر به پدیده بیش‌برازش شود.

در انتخاب ساختار باید نقطه تعادل بین خطای برازش و خطای تعمیم‌پذیری شبکه یافته شود. خطای پیش‌بینی از تفاضل خروجی مدل و مقدار واقعی سری زمانی محاسبه می‌شود. معیار خطا در داده‌های آموزش و داده‌های آزمون هر چه کوچک‌تر باشد به ترتیب بیانگر میزان مناسب برازش و تعمیم‌پذیری مدل است. خطای برازش به عنوان معیاری برای سنجش کیفیت مدل، پایان دادن به یادگیری مدل و رفتن به مرحله آزمون مدل استفاده می‌شود. خطای e_t به ازای مقدار واقعی Y_t و پیش‌بینی F_t در دوره t عبارت است از:

$$e_t = Y_t - F_t \quad (20)$$

هایدمن و کوهرلر معیارهای آماری برای اندازه‌گیری خطای پیش‌بینی را در پنج گروه دسته‌بندی می‌کند؛ شاخص وابسته به مقیاس، شاخص درصد خطا، شاخص متقارن خطا، شاخص نسبی خطا و شاخص مقیاس خطا.

در این مقاله از سه معیار برای اندازه‌گیری عملکرد پیش‌بینی استفاده می‌شود. معیار اول مجذور مربعات خطا است که از شاخص‌های وابسته به مقیاس است. این معیار برای مقایسه روش‌های مختلف پیش‌بینی روی یک مجموعه داده مناسب است، اما برای مقایسه نتیجه یک روش بر مجموعه داده‌های مختلف مناسب نیست. معیار دوم، متوسط قدر مطلق درصد خطا است که از نوع شاخص درصد خطا است. این معیار مستقل از مقیاس داده‌ها بوده و می‌توان آن را برای مقایسه نتایج حاصل از مجموعه داده‌های مختلف بکار برد. این نوع معیارها با وجود متداول بودن معیایی نیز دارند. از جمله اینکه به ازای مقادیر داده‌های واقعی صفر ($Y_t=0$) خطا بینهایت و به ازای مقادیر نزدیک به صفر خطا توزیع اریبی شدیدی دارد.

ارزیابی عملکرد مدل پیش‌بینی بر اساس میزان انحراف از مقادیر واقعی مهم‌ترین معیار قضاوت در مورد روش‌های پیش‌بینی سری زمانی است، اما تنها معیار نیست. معیار سوم در این مقاله با کمک رگرسیون محاسبه می‌شود. با استفاده از رگرسیون خطی مقادیر واقعی سری زمانی و خروجی پیش‌بینی کننده، معیاری از عملکرد پیش‌بینی کننده محاسبه می‌شود. معیار R رگرسیون، شیب و مقدار ثابت معادله خطی سه معیار رگرسیون هستند. خلاصه معیارهای مورد استفاده در این مقاله و نحوه محاسبه آنها در جدول (۱) ارائه شده است.

جدول (۱) - معیارهای اندازه‌گیری خطای پیش‌بینی مورد استفاده در این مقاله

$e_t = Y_t - F_t$ محاسبه خطای پیش‌بینی	
Root Mean Squared Error مجذور متوسط مربعات خطا	$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N e_t^2}$
Mean Absolute Percentage Error متوسط قدرمطلق درصد خطا	$MAPE = \frac{1}{N} \sum_{t=1}^n \left \frac{e_t}{Y_t} \times 100 \right $
$Y_t - F_t$ Linear Regression: R-value, Slope, Offset	$SLOP = \frac{\sum_{i=1}^n (Y_i - \bar{Y}_i)(F_i - \bar{F}_i)}{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2}$
ضریب همبستگی، شیب و ثابت معادله خطی رگرسیون	$R = \frac{\sum_{i=1}^n (Y_i - \bar{Y}_i)(F_i - \bar{F}_i)}{\sum_{i=1}^n (Y_i - \bar{Y}_i)^2}$
	$OFFSET = F_t - SLOP \times Y_t$

۳- طراحی مدل شبکه عصبی برای پیش‌بینی داده‌های سری زمانی

۳-۱- مجموعه داده‌ها

داده‌های مورد استفاده در این تحقیق مربوط به تغییرات قیمت و حجم معاملات سهام چهار بنگاه بزرگ اقتصادی است که در بازار بورس لندن (FTSE) خرید و فروش شده است. بنگاه‌های اقتصادی عبارتند از: شرکت اپل (Apple)، شرکت مایکروسافت (MS)، شرکت نفت بریتانیا (BP) و شرکت نفت شل (Shell).

مجموعه داده‌ها مشتمل بر قیمت سهام، حجم سهام معامله شده و شاخص قیمت به صورت روزانه و مربوط به معاملات انجام شده از ابتدای ژانویه ۲۰۰۰ تا سه ماه اول ۲۰۱۳

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۹۳

است. قیمت سهام هر روز به صورت جداگانه در ابتدا، پایان، بالاترین و کمترین قیمت روز ثبت شده است. به طور متوسط داده‌های بیش از ۳۴۰۰ روز کاری برای هر شرکت در این تحقیق مورد استفاده قرار گرفته است.

در جدول (۲) به عنوان نمونه، بخشی از داده‌های خام یکی از شرکت‌ها نشان داده شده است. در این تحقیق شاخص قیمت روزانه مورد تحلیل و پیش‌بینی است.

جدول (۲) - نمونه‌ای از داده‌های سری زمانی گرفته شده از FTSE

Date	Open	High	Low	Close	Volume	Adj. Close
زمان	قیمت ابتدا	بالاترین	پایین‌ترین	قیمت انتها	حجم معاملات	
۳-Jan-۲۰۰۰	۱۰۴/۸۷	۱۱۲/۵	۱۰۱/۶۹	۱۱۱/۹۴	۱۹۱۴۴۴۰۰	۲۷/۲۲
۴-Jan-۲۰۰۰	۱۰۸/۲۵	۱۱۰/۶۲	۱۰۱/۱۹	۱۰۲/۵	۱۸۳۱۰۰۰۰	۲۴/۹۲
۵-Jan-۲۰۰۰	۱۰۳/۷۵	۱۱۰/۵۶	۱۰۳	۱۰۴	۲۷۸۱۸۰۰۰	۲۵/۲۹
۶-Jan-۲۰۰۰	۱۰۶/۱۲	۱۰۷	۹۵	۹۵	۲۷۴۴۳۲۰۰	۲۳/۱
۷-Jan-۲۰۰۰	۹۶/۵	۱۰۱	۹۵/۵	۹۹/۵	۱۶۴۶۳۲۰۰	۲۴/۱۹
۵-Jul-۲۰۱۳	۴۲۰/۳۹	۴۲۳/۲۹	۴۱۵/۳۵	۴۱۷/۴۲	۹۷۸۶۶۰۰	۴۱۴/۶۸
۸-Jul-۲۰۱۳	۴۲۰/۱۱	۴۲۱	۴۱۰/۶۵	۴۱۵/۰۵	۱۰۶۴۷۸۰۰	۴۱۲/۳۳
۹-Jul-۲۰۱۳	۴۱۳/۶	۴۲۳/۵	۴۱۰/۳۸	۴۲۲/۳۵	۱۲۵۹۲۳۰۰	۴۱۹/۵۸
۱۰-Jul-۲۰۱۳	۴۱۹/۶	۴۲۴/۸	۴۱۸/۲۵	۴۲۰/۷۳	۱۰۰۵۰۲۰۰	۴۱۷/۹۷

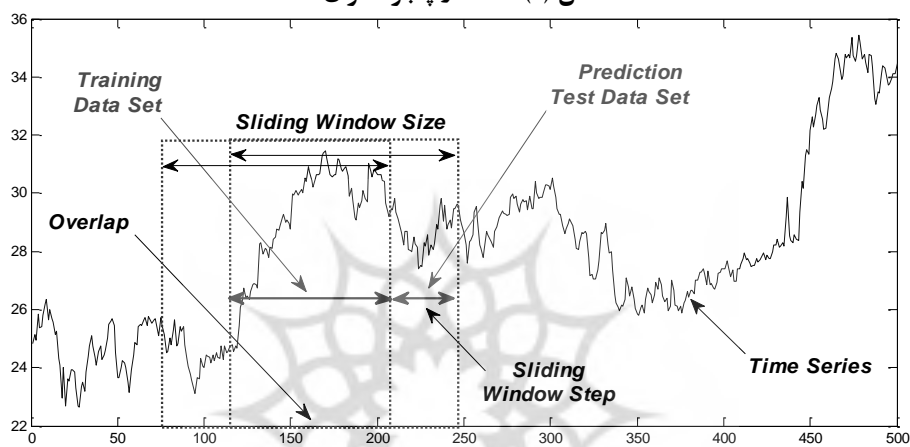
۲-۳- ساختار پنجره لغزان

ارزیابی عملکرد شبکه عصبی در پیش‌بینی سری زمانی با استفاده از روش پنجره لغزان^۱ انجام می‌گیرد. در این روش، پیش‌بینی به صورت مکرر روی زیرمجموعه‌های متوالی از سری زمانی صورت می‌گیرد. زیرمجموعه متناهی و متوالی از سری زمانی، پنجره نامیده می‌شود. اندازه پنجره^۲ عبارت از تعداد داده‌های درون آن است که عملاً معرف یک دوره زمانی است. پنجره با فاصله یا قدم^۳ ثابت روی کل داده‌های سری زمانی حرکت داده می‌شود. پنجره‌های

- 1- Sliding Windows
- 2- Sliding Window Size
- 3- Sliding Window Step

متوالی قسمت‌های مشترک دارند که به اصطلاح همپوشانی^۱ پنجره‌ها نامیده می‌شود. فرآیند پیش‌بینی سری زمانی به صورت مکرر و مستقل از هم روی هر پنجره انجام می‌گیرد. اندازه و قدم پنجره دو پارامتر تعیین‌کننده در مدیریت داده‌های سری زمانی و تعداد تکرار انجام پیش‌بینی هستند. شکل (۶) ساختار پنجره لغزان را با نمایش دو پنجره متوالی نشان می‌دهد.

شکل (۶) - ساختار پنجره لغزان



اندازه پنجره با تعداد نمونه داده‌های درون آن و قدم پنجره با تعداد نمونه بین انتهای دو پنجره متوالی مشخص می‌شود. به ازای هر پنجره لغزان، فرآیند پیش‌بینی سری زمانی انجام می‌گیرد، از این رو شبیه‌سازی فرآیند پیش‌بینی سری زمانی در طول زمان امکانپذیر می‌شود. جدول (۳) تعداد حالت‌های مختلف از پنجره لغزان را نمایش می‌دهد که در این تحقیق مورد استفاده قرار گرفته است.

در هر دوره پیش‌بینی (پنجره لغزان)، بخشی بزرگی از داده‌های متوالی ابتدای هر پنجره برای آموزش شبکه عصبی اختصاص داده می‌شود که به اصطلاح مجموعه داده آموزش

1- Overlap

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۹۵

نامیده می‌شود. بخش کوچک‌تر باقیمانده برای آزمون شبکه عصبی و محاسبه خطای پیش‌بینی اختصاص داده می‌شود. این بخش مجموعه داده آزمون نامیده می‌شود. به عبارت دیگر، در هر دوره پیش‌بینی با استفاده از داده‌های آموزشی ضرایب وزنی محاسبه و مدل شبکه عصبی ساخته می‌شود و سپس با استفاده از داده‌های آزمون، عملکرد مدل ارزیابی می‌شود. این روند در هر پنجره لغزان تکرار و خطای پیش‌بینی در هر پنجره محاسبه و ثبت می‌شود. چنانچه در جدول (۳) نشان داده شده، تعداد داده‌های آزمون همواره مساوی با قدم پنجره لغزان است.

جدول (۳) - ساختارهای پنجره لغزان مورد استفاده در این گزارش

تعداد نمونه‌ها	اندازه پنجره	قدم پنجره	تعداد پنجره
۵۰۰	۵۰	۳	۱۵۰
۵۰۰	۵۰	۲	۲۲۵
۱۰۰۰	۳۰	۳	۳۲۹
۱۰۰۰	۱۵	۲	۴۹۳
۱۰۰۰	۵۰	۲	۴۷۵
۳۴۰۰	۱۰۰	۱۰	۳۴۰

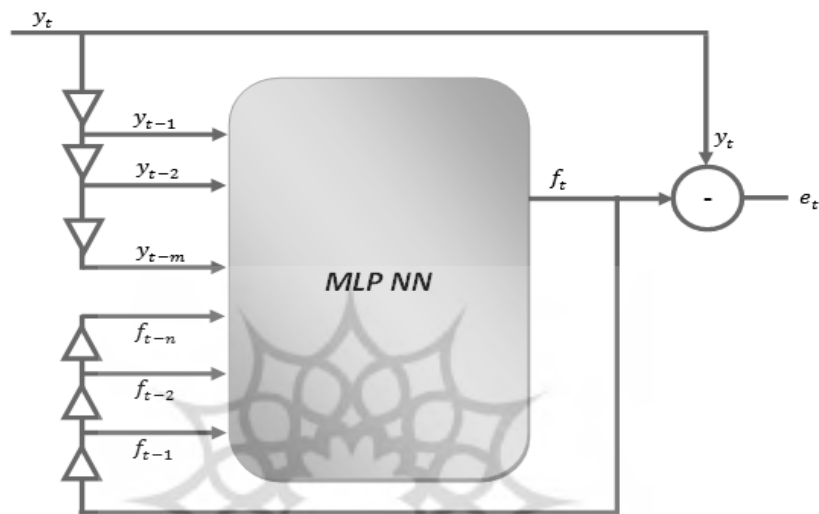
به عنوان مثال، انتخاب اندازه پنجره معادل ۵۰ به معنی بررسی عملکرد پیش‌بینی کننده در یک دوره ۵۰ روزه است. انتخاب قدم پنجره معادل «۳» به معنی استفاده از داده‌های ۴۷ روز ابتدایی برای آموزش شبکه عصبی و آزمون عملکرد آن در پیش‌بینی داده‌های «۳» روز پایانی هر دوره است. این فرآیند برای یک سری زمانی با تعداد کل ۵۰۰ نمونه، ۱۵۰ بار تکرار می‌شود. در هر تکرار متوسط خطای پیش‌بینی سه روز محاسبه و ثبت می‌شود. ارزیابی عملکرد پیش‌بینی کننده با محاسبه میانگین خطا در ۱۵۰ تکرار به دست می‌آید.

۳-۳- ساختار شبکه عصبی بازگشتی با تاخیر زمانی

ساختار شبکه عصبی مورد مطالعه در این مقاله در شکل (۷) نمایش داده شده است. ساختار شبکه عصبی بازگشتی با تاخیر زمانی به گونه‌ای است که هر نمونه از سری زمانی، بر اساس

m نمونه اصلی و n نمونه پیش‌بینی، تخمین زده می‌شود. شبکه عصبی یک خروجی دارد و خطای خروجی حاصل تفاضل خروجی و مقدار واقعی سری زمانی است. مدل ریاضی این شبکه در بخش قبل با نام NARMA-RNN ارائه شد.

شکل (۷) - ساختار شبکه عصبی بازگشتی با تاخیر زمانی



با توجه به اصول یادگیری ماشین، ابتدا بخشی از داده‌ها برای انتخاب ساختار بهینه و نیز یادگیری شبکه عصبی بکار برده می‌شود. انتخاب ساختار بهینه شامل تعداد ورودی تاخیری (m)، تعداد بازخورد تاخیری (n)، تعداد لایه‌های پنهان، تعداد پرسپترون و تابع فعالیت آن در لایه‌های پنهان است. لایه خروجی یک پرسپترون خطی است. پس از انتخاب ساختار مناسب، ضرایب وزنی شبکه با استفاده از الگوریتم پس انتشار خطا محاسبه می‌شوند.

ضرایب وزنی به گونه‌ای محاسبه می‌شوند که متوسط مربعات خطای خروجی شبکه در مجموعه داده‌های آموزشی کمتر از ۰/۰۰۱ شود. پس از پایان یادگیری شبکه، خروجی آن به ازای داده‌های مجموعه آزمون محاسبه و خطای پیش‌بینی ثبت می‌شود.

عملکرد شبکه عصبی در هر پنجره لغزان به صورت مستقل و جداگانه ارزیابی می‌شود. بنابراین در هر پنجره، ابتدا ساختار مناسب انتخاب و فرآیند یادگیری با استفاده از داده‌های آموزشی آن پنجره انجام می‌شود و سپس خطای پیش‌بینی به ازای داده‌های آزمون آن

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۹۷

پنجره محاسبه می شود. این به دلیل شبیه سازی فرآیند استفاده از شبکه عصبی در طول زمان است. انتخاب ساختار برای شبکه عصبی با هدف رسیدن به حداقل خطا در داده های آموزشی است. در انتخاب ساختار، فرض کردیم تعداد تاخیرها در ورودی و بازخورد یکسان و عددی بین ۲ و ۱۰ باشد. این به معنی پیش بینی هر نمونه سری زمانی بر اساس نمونه های ۲ تا ۱۰ روز گذشته است.

تنوع نسبتاً کاملی از تعداد لایه و پرسپترون برای شبکه عصبی نیز در نظر گرفته شده است که در جدول (۴) نمایش داده می شود. انتخاب ساختار بهینه به صورت تکرار از ساختار ساده به سمت پیچیده است. با توجه به انتخاب تصادفی ضرایب وزنی، هر ساختار سه بار با مقادیر متفاوت اولیه مورد ارزیابی قرار می گیرد.

جدول (۴) - گزینه ها در انتخاب ساختار مناسب شبکه عصبی

<i>Inputs or Delay Taps:</i> ۲, ۳, ۵, ۷, ۱۰	لایه ورودی ها (عملگرهای تاخیری)
<i>Hidden Layers:</i> ۲, ۳, ۵, ۷, ۱۰, ۱۲, ۱۵, ۳۰, ۵۰,	لایه پنهان و پرسپترون ها
۲-۳, ۳-۵, ۵-۵, ۵-۱۰, ۳-۳ (Sigmoid function)	
<i>Output:</i> ۱, Linear function	لایه خروجی
<i>Training Performance:</i> $MSE = 0.001$	معیار خطا آموزش
<i>Max Epoch = 500</i>	حداکثر تکرار

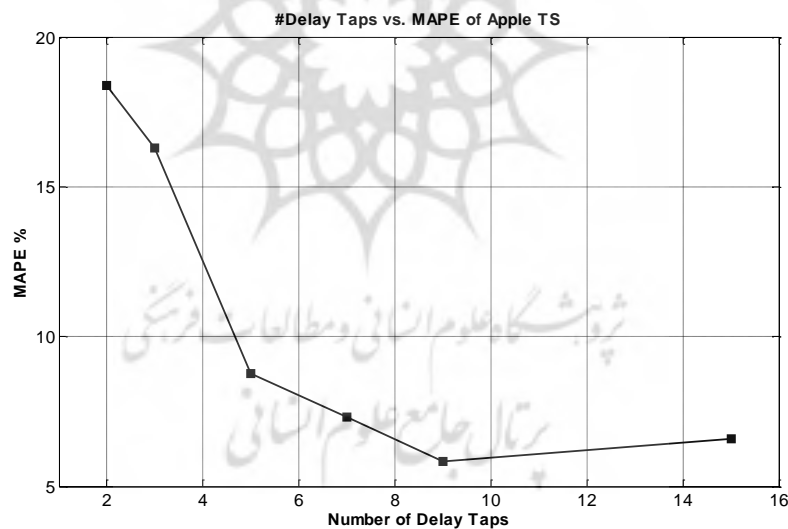
۴- نتایج حاصل از ارزیابی عملکرد شبکه عصبی

در این قسمت نتایج تحقیق ارائه می شود؛ در بخش اول نحوه انتخاب ساختار بهینه و نتایج آن و در بخش بعدی نحوه مدیریت داده ها و انتخاب اندازه و قدم مناسب برای پنجره لغزان مطرح و نتایج آن ارائه می شود. در ادامه نتایج به دست آمده از پیش بینی سری زمانی با استفاده از شبکه عصبی و انتخاب ساختار بهینه مورد بررسی و تحلیل قرار می گیرد.

۴-۱- بررسی تعداد عملگرهای تاخیری در پیش‌بینی سری زمانی

یکی از چالش‌های پیش‌بینی سری زمانی، انتخاب بهینه تعداد مشاهدات مورد نیاز از نمونه‌های پیشین برای پیش‌بینی نمونه آتی است. این تعداد بستگی شدید به ماهیت و پیچیدگی سری زمانی، مقطع زمانی پیش‌بینی و مدل پیش‌بینی‌کننده دارد. در مدل شبکه عصبی تاخیر زمانی (TDNN) تعداد عملگرهای تاخیر (پارامتر m)، معادل پارامتر p در مدل $NARMA(p,q)$ ، مشخص‌کننده تعداد نمونه‌های پیشین و موثر در تخمین نمونه آتی است. تعداد زیاد عملگرهای تاخیری باعث افزایش بار محاسباتی، پیچیدگی ساختار شبکه عصبی، احتمال فرابرازش و نهایتاً افت کیفیت پیش‌بینی می‌شود. تعداد کم عملگرها نیز سبب نقصان مدل شبکه عصبی و افزایش خطای پیش‌بینی می‌شود. در این بخش تاثیر تعداد عملگرهای تاخیری بر معیار خطای پیش‌بینی مطالعه شده است.

شکل (۸) - تاثیر تعداد عملگرهای تاخیری بر متوسط درصد خطای پیش‌بینی



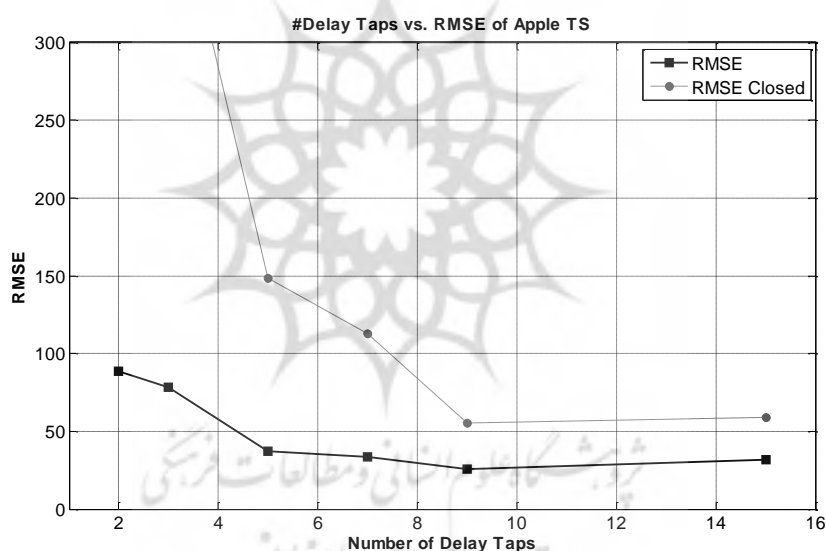
نمودارهای دو شکل (۸) و (۹) تاثیر تعداد عملگرهای تاخیری را در عملکرد شبکه عصبی برای پیش‌بینی سری زمانی شاخص قیمت Apple نمایش می‌دهد. شکل (۸)

کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۹۹

نمودار متوسط درصد خطا (MAPE) و شکل (۹) نمودار متوسط مربعات خطا (RMSE) را برای تعداد ۲، ۳، ۵، ۷، ۹ و ۱۵ عملگر تأخیری نشان می‌دهد. متوسط خطای پیش‌بینی با توجه به روش پنجره لغزان به ازای هر تعداد عملگر با استفاده از بیش از ۳۰۰ تکرار محاسبه شده است.

در شکل (۹) نتایج حاصل از دو شبکه پیش‌رو و بازگشتی (Closed loop) مقایسه شده است. نتایج نشان می‌دهد افزایش تعداد عملگرها باعث افت خطای پیش‌بینی می‌شود. تعداد بهینه برای عملگرهای تأخیری ۷ و ۹ است. به عبارت دیگر، برای پیش‌بینی هر نمونه در سری زمانی، استفاده از ۹ نمونه قبلی توصیه می‌شود.

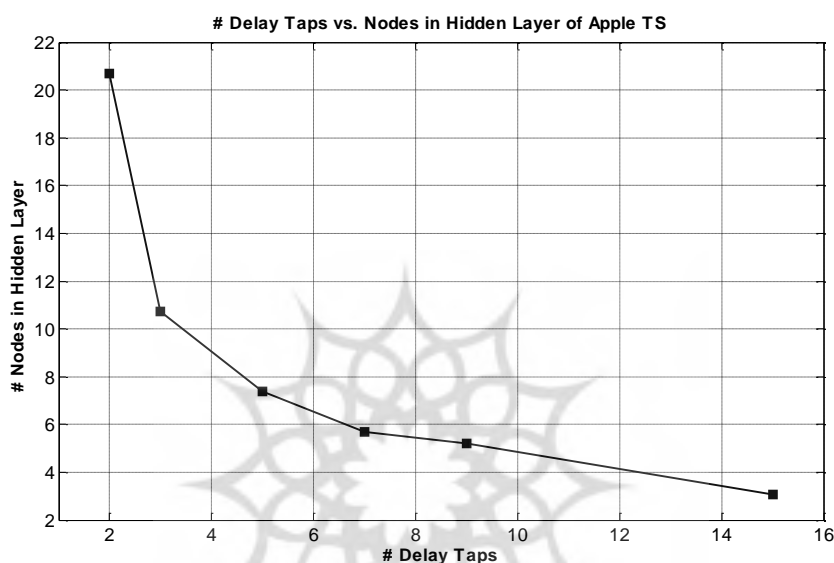
شکل (۹) - تاثیر تعداد عملگرهای تأخیری بر متوسط مربعات خطای پیش‌بینی برای شبکه پیش‌رو و بازگشتی



نکته بعدی در بررسی تعداد عملگرهای تأخیری، ساختار بهینه شبکه است. شکل (۱۰) تعداد نرون‌ها در لایه پنهان شبکه عصبی (محور عمودی) را در مقایسه با تعداد عملگرهای تأخیری (محور افقی) نشان می‌دهد. چنانچه ملاحظه می‌شود، تعداد نرون‌ها با افزایش ورودی‌ها کاهش می‌یابد.

به عبارت دیگر، تعداد ضرایب وزنی شبکه در لایه ورودی به ازای هر سری زمانی تقریباً ثابت است. با توجه به اینکه تعداد ضرایب وزنی کاملاً مرتبط با مدل بهینه شبکه برای تقریب سری زمانی است، این نتیجه با مبانی نظری کاملاً مطابقت دارد.

شکل (۱۰) - مقایسه تعداد عملگرهای تاخیری و تعداد پرسپترون در لایه پنهان شبکه عصبی



در جدول (۵) ساختارهای بهینه شبکه عصبی که برای پیش‌بینی سری زمانی استفاده شده است، همراه با تعداد ضرایب وزنی ارائه می‌شود.

جدول (۵) - ساختارهای بهینه برای پیش‌بینی سری زمانی

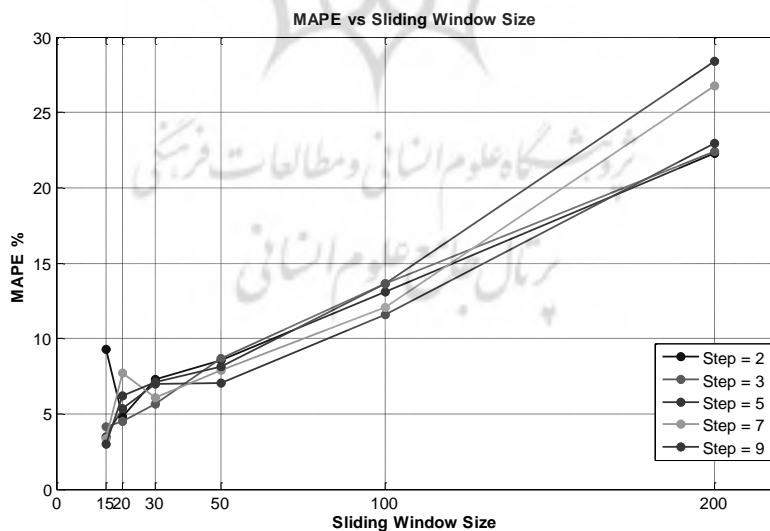
۲-۲۱-۱	۴۲ + ۲۱
۳-۱۱-۱	۳۳ + ۱۱
۵-۷-۱	۳۵ + ۷
۷-۵-۱	۳۵ + ۵
۹-۴-۱	۳۶ + ۴
۱۵-۲-۱	۳۰ + ۲

۴-۲- بررسی اندازه پنجره لغزان در پیش‌بینی سری زمانی

چالش دیگر در طراحی پیش‌بینی کننده شبکه عصبی، انتخاب صحیح اندازه پنجره است. اندازه پنجره لغزان، تعداد نمونه‌های بکار رفته در هر مرحله از پیش‌بینی است. این تعداد مشتمل بر دو مجموعه آموزش و آزمون است. قدم پنجره لغزان دوره پیش‌بینی و دفعات پیش‌بینی را مشخص می‌کند. قدم‌های کوچک به معنی پیش‌بینی تعداد کمی نمونه در هر مرحله و در نتیجه تعداد زیاد دفعات پیش‌بینی و قدم‌های بزرگ به معنی پیش‌بینی نمونه‌های زیاد در هر مرحله و در نتیجه تعداد دفعات کم پیش‌بینی است. در هر پنجره به تعداد قدم از نمونه‌های آخر پنجره به عنوان مجموعه آزمون و مابقی به عنوان مجموعه آموزش استفاده می‌شود.

نمودار دو شکل (۱۱) و (۱۲) به ترتیب تاثیر اندازه و قدم پنجره لغزان بر متوسط خطای پیش‌بینی را نشان می‌دهد. چنانچه در شکل (۱۱) نمایان است، درصد خطای پیش‌بینی با کاهش اندازه پنجره، کاهش چشمگیری دارد. در این نمودار اندازه پنجره معادل ۲۰۰، ۱۰۰، ۵۰، ۳۰ و ۲۰ نمونه مورد مطالعه قرار گرفته است. نمودار شکل (۱۲) نشان می‌دهد اندازه قدم پنجره لغزان تاثیر معناداری بر عملکرد پیش‌بینی کننده ندارد.

شکل (۱۱) - تاثیر اندازه پنجره لغزان بر متوسط درصد خطای پیش‌بینی



کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۱۰۳

جدول (۶) - متوسط درصد خطا به ازای اندازه و قدم پنجره لغزان

Size/Step	۱۵	۲۰	۳۰	۵۰	۱۰۰	۲۰۰
۲	۹/۲۶	۴/۷۹	۷/۲۸	۸/۵۹	۱۳/۱۱	۲۲/۳۲
۳	۴/۱۴	۴/۴۸	۵/۶۴	۸/۶۶	۱۳/۶۶	۲۲/۳۹
۵	۳/۴۶	۵/۳۸	۶/۹۶	۷/۰۴	۱۱/۵۷	۲۲/۹۵
۷	۳/۳۶	۷/۷۳	۶/۰۶	۷/۹۱	۱۲/۰۶	۲۶/۷۷
۹	۲/۹۹	۶/۲۱	۷/۱۱	۸/۱۲	۱۳/۶۵	۲۸/۳۹

جدول (۷) - تعداد نرون‌های ساختار بهینه شبکه به ازاء اندازه و قدم پنجره لغزان

Size/Step	۱۵	۲۰	۳۰	۵۰	۱۰۰	۲۰۰
۲	۲	۳	۵	۸	۱۶	۳۴
۳	۲	۳	۵	۷	۱۵	۳۴
۵	۲	۳	۴	۷	۱۶	۳۳
۷	۲	۲	۴	۷	۱۵	۳۳
۹	۲	۲	۴	۷	۱۵	۳۲

۴-۳- نتایج حاصل از پیش‌بینی سری زمانی با کمک شبکه عصبی تاخیری

با توجه به نتایج به دست آمده از مطالعه تاثیر تعداد عملگرهای تاخیری و اندازه پنجره، ساختارهای بهینه برای پیش‌بینی سری زمانی در جدول (۵) مشخص شد. در این قسمت نتایج به دست آمده از پیش‌بینی سری زمانی شاخص قیمت چهار بنگاه اقتصادی ارائه می‌شود.

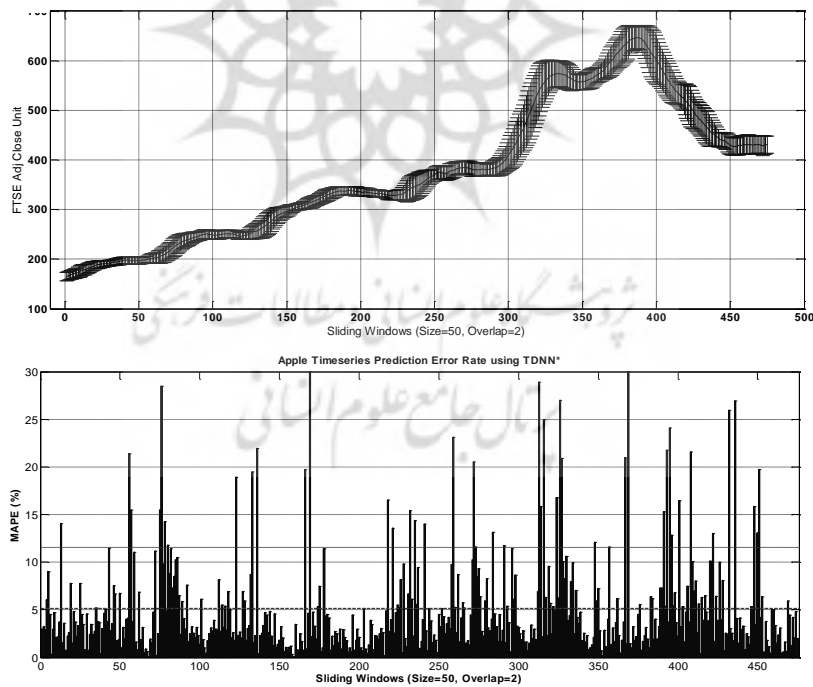
در جدول (۸) نتایج پیش‌بینی برای یک هزار نمونه در ۱۰ ستون نشان داده می‌شود. تعداد پیش‌بینی که تابعی از اندازه و قدم پنجره لغزان است در ستون M و متوسط انحراف معیار کل پنجره‌ها در ستون STD نمایش داده می‌شود. ستون چهارم از سمت چپ اندازه و قدم پنجره لغزان را نشان می‌دهد و در ستون $TDNN$ ساختار بهینه مشتمل بر تعداد نرون‌های ورودی، پنهان و خروجی نمایش داده می‌شود. متوسط مولفه‌های محاسبه شده برای رگرسیون خطی خروجی شبکه و خروجی واقعی در ستون $Regression$ و متوسط درصد خطا و متوسط مربعات خطای پیش‌بینی در ستون‌های $MAPE$ و $RMSE$ نمایش داده شده است. همچنین متوسط درصد خطا و متوسط مربعات خطای پیش‌بینی به صورت بازگشتی شبکه در ستون‌های $MAPE\ Closed$ و $RMSE\ Closed$ نمایش داده شده است.

در اشکال (۱۴)، (۱۵)، (۱۶) و (۱۷) نتایج پیش‌بینی برای چهار شرکت به ترتیب $Apple$ ، MS ، BP و $Shell$ ارائه شده است. نمودار بالایی تغییرات میانگین و انحراف معیار در هر پنجره و نمودار زیرین تغییرات متوسط درصد خطا در هر پنجره را نمایش می‌دهد.

جدول (۸) - نتایج تکرار پیش‌بینی برای سری زمانی شاخص قیمت چهار بنگاه اقتصادی

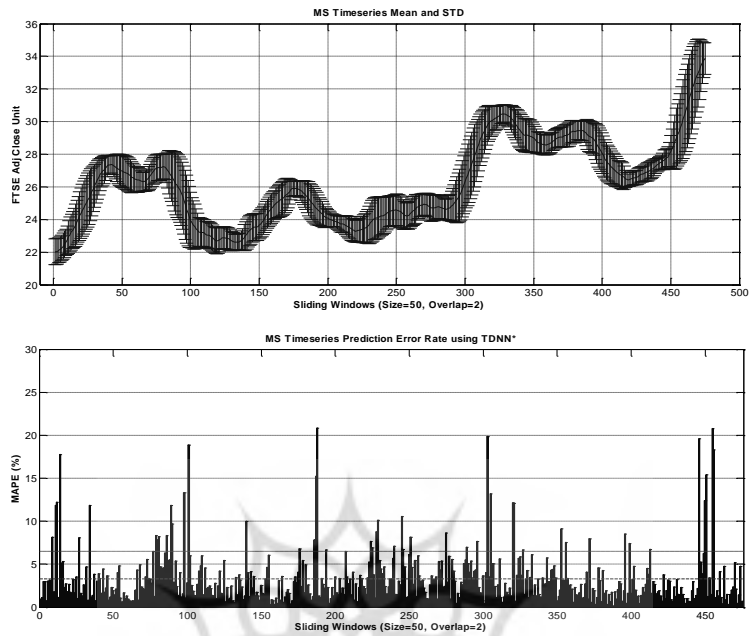
TS (1000)	M	STD	Sliding window	TDNN	Regression R-Slope-Offset	MAPE%	RMSE	MAPE% Closed	RMSE Closed
Apple	۳۲۴	۱۴/۲	۳۰/۳	۵-۵-۱	۰/۸۲-۳/۴-۹۰۰	۴۲	۲۰۳	۲۹	۱۷۰
MS	۳۲۴	۰/۷۴	۳۰/۳	۵-۵-۱	۰/۸۱-۰/۹-۲/۴	۴	۱/۲	۵۹	۲۵
BP	۳۲۴	۱۴/۴	۳۰/۳	۵-۵-۱	۰/۸-۰/۹-۲۹	۵۰/۴	۲۷/۷	۱۶/۷	۷۷
Shell	۳۲۴	۴۵	۳۰/۳	۵-۵-۱	۰/۸-۰/۹-۱۵۰	۴/۵	۹۳	۸/۵	۱۷۶
BP	۴۹۳	۱۰	۱۵/۲	۵-۳-۱	۰/۷-۰/۷-۱۱۲	۳/۳	۱۵	۳/۲	۱۴/۶
Apple	۴۹۳	۱۰/۲	۱۵/۲	۵-۳-۱	۰/۷-۰/۸۱-۷۰	۴/۱	۱۷/۱	۳/۹	۱۶/۵
Apple	۴۷۵	۱۸/۶	۵۰/۲	۹-۵-۱	۰/۹۴-۰/۹۹-۵/۳	۵/۲	۲۲/۷	۱۳/۶	۶۲
MS	۴۷۵	۰/۹۷	۵۰/۲	۹-۵-۱	۰/۹۵-۰/۹۷-۰/۷	۳/۳	۰/۹۶	۹/۸	۲/۸
BP	۴۷۵	۱۹	۵۰/۲	۹-۵-۱	۰/۹۳-۰/۹۸-۹/۹	۵/۵	۲۵	۱۲/۳	۵۴
Shell	۴۷۵	۵۶	۵۰/۲	۹-۵-۱	۰/۹۳-۰/۹۸-۴۵/۴	۳/۹	۷۸/۷	۹/۹	۱۹۸

شکل (۱۴) - نمودار تغییرات مشخصه‌های آماری و نتایج پیش‌بینی برای شرکت Apple

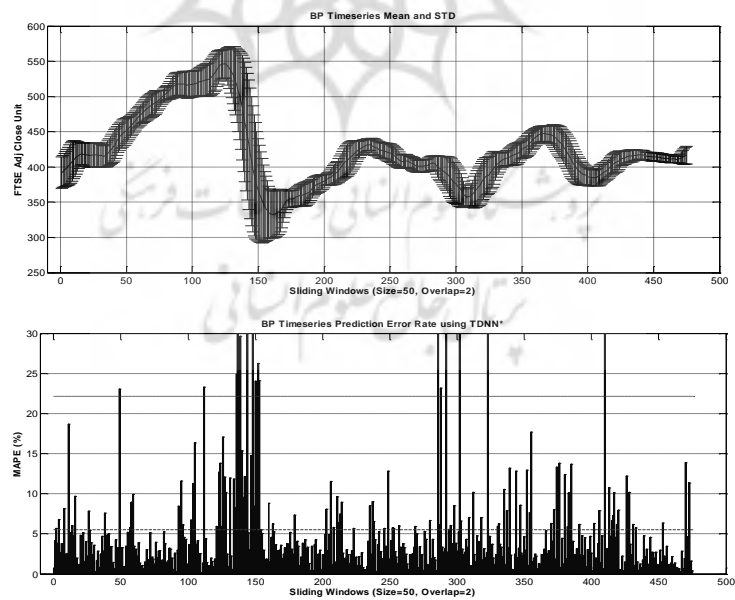


کاربرد روش پنجره لغزان برای انتخاب ساختار شبکه عصبی با ... ۱۰۵

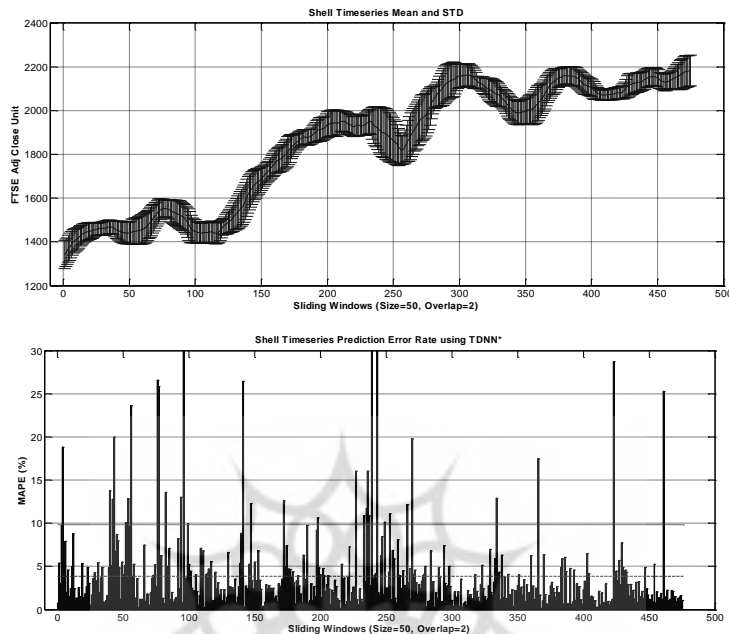
شکل (۱۵) - نمودار تغییرات مشخصه‌های آماری و نتایج پیش‌بینی برای شرکت MS



شکل (۱۶) - نمودار تغییرات مشخصه‌های آماری و نتایج پیش‌بینی برای شرکت BP



شکل (۱۷) - نمودار تغییرات مشخصه‌های آماری و نتایج پیش‌بینی برای شرکت Shell



۵- جمع‌بندی و نتیجه‌گیری

در این مقاله یک مدل مبتنی بر شبکه عصبی با تاخیر زمانی برای پیش‌بینی معیارهای مالی بازار سهام ارائه می‌شود. روش پنجره لغزان برای انتخاب ساختار مناسب و ارزیابی عملکرد پیش‌بینی بکار برده شده است. در این مقاله انتخاب ساختار مناسب شبکه، تعداد عملگرهای تاخیری، تعداد بهینه داده‌های پیشین و پسین و معیار کمی مناسب برای ارزیابی عملکرد پیش‌بینی‌کننده، مورد مطالعه قرار گرفته است. عملکرد این مدل پیش‌بینی‌کننده روی شاخص قیمت سهام چهار بنگاه بزرگ اقتصادی در بازار سهام لندن مورد بررسی قرار گرفته است. نتایج به دست آمده، نشان می‌دهد با روش پیشنهادی، ساختاری انتخاب شد که متوسط درصد خطا، متوسط مجذور مربعات خطا و معیار رگرسیون خطی خروجی شبکه در مقایسه با ساختارهای اولیه با در حد قابل توجهی کاهش دارد.

منابع

الف - فارسی

- اصغری اسکویی، محمدرضا (۱۳۸۳)، «پیش بینی سری های زمانی با کمک شبکه عصبی»، پژوهش های اقتصادی ایران، فصلنامه علمی-پژوهشی مرکز تحقیقات اقتصاد ایران، سال شانزدهم، شماره ۴۷، تابستان، صفحات ۱۸۳-۱۶۳.
- اصغری اسکویی، محمدرضا (۱۳۹۲)، «هوش محاسباتی»، کرسی های علمی-ترویجی، دانشکده اقتصاد، دانشگاه علامه طباطبایی

ب - انگلیسی

- Castillo, Oscar and Patricia Melin (2002), Hybrid Intelligent Systems for Time Series Prediction Using Neural Networks, Fuzzy Logic and Fractal Theory , *Neural Networks, IEEE Transactions on*, Vol.13, No.6, pp.1395,1408.
- Chng, E.S., S. Chen and B. Mulgrew (1996) Gradient Radial Basis Function Networks for Nonlinear and Nonstationary Time Series Prediction , *Neural Networks, IEEE Transactions on*, Vol.7, No.1, 190-194.
- Connor, J.T., R.D. Martin and L.E. Atlas (1994), Recurrent Neural Networks and Robust Time Series Prediction , *Neural Networks, IEEE Transactions on* , Vol.5, No.2, pp. 240-254.
- Daniel Graves, Witold Pedrycz (2009), Fuzzy Prediction Architecture Using Recurrent Neural Networks , *Neurocomputing*, Vol. 72, No. 7-9, pp. 1668-1678.
- De Gooijer, G. J. and R. J. Hyndman (2006), 25 Years of Time Series Forecasting , *International Journal of Forecasting*, Vol. 22, No 3, pp. 443- 473.
- Frank, R. J., N. Davey and S. P. Hunt (2001), Time Series Prediction and Neural Networks , *Journal of Intelligent and Robotic Systems*, Vol. 31, No. 1-3, pp. 91-103.

- Hamzaçebi, Co kun, Diyar Akay and Fevzi Kutay (2009), Comparison of Direct and Iterative Artificial Neural Network Forecast Approaches in Multi-periodic Time Series Forecasting , *Expert Systems with Applications*, Vol. 36, No. 2, pp. 3839-3844.
- Lu, Chi-Jie, Tian-Shyug Lee and Chih-Chou Chiu (2009), Financial Time Series Forecasting Using Independent Component Analysis and Support Vector Regression, *Decision Support Systems*, Vol. 47, No. 2, pp. 115-125.
- Knerr, C. (2004), Time Series Prediction Using Neural Networks , PhD Thesis, Texas Tech University.
- Panagiotopoulos, A. (2012), Optimizing Time Series Forecast Through Linear Programming , PhD Thesis, Nottingham University.
- Sapankevych, N. I. and Ravi Sankar (2009), Time Series Prediction Using Support Vector Machines: A Survey , *Computational Intelligence Magazine, IEEE* , Vol.4, No.2, pp. 24-38.
- Van Gestel, T. and others (2001), Financial Time Series Prediction Using Least Squares Support Vector Machines within the Evidence Framework , *Neural Networks, IEEE Transactions on* , Vol.12, No.4, pp.809-821.
- Zhang, Jun, H.S.H.Chung and Wai-Lun Lo (2008), Chaotic Time Series Prediction Using a Neuro-Fuzzy System with Time-Delay Coordinates , *Knowledge and Data Engineering, IEEE Transactions on*, Vol.20, No.7, pp. 956,964.