

به کارگیری منطق فازی برای بهبود عملکرد شبکه‌های عصبی مصنوعی به منظور پیش‌بینی نرخ ارز

مهدی خاشعی* (دانشجوی پسا دکتری)

مهدی بیجاری (دانشیار)

فریماه مخاطب رفیعی (دانشیار)

دانشکده‌ی مهندسی صنایع و سیستم‌ها، دانشگاه صنعتی اصفهان

مهندسی صنایع و مدیریت شریف (زمستان ۱۳۹۲)
دوره ۱ - ۲۹، شماره ۲، ص. ۶۳-۷۱

روش‌های هوش محاسباتی، همچون شبکه‌های عصبی مصنوعی و منطق فازی، به عنوان ابزاری محبوب به منظور پیش‌بینی بازارهای پیچیده‌ی مالی معرفی شده‌اند. دقت پیش‌بینی‌ها از جمله مهم‌ترین مشخصه‌های مدل‌های پیش‌بینی است و تلاش برای بهبود بخشیدن کارایی مدل‌های سری‌های زمانی هرگز متوقف نشده است. امروزه علی‌رغم روش‌های متعدد پیش‌بینی سری‌های زمانی که در چند دهه‌ی اخیر پیشنهاد شده‌اند، هنوز پیش‌بینی نرخ‌های ارز، کار بسیار دشواری محسوب می‌شود. در این مطالعه، مدل ترکیبی جدیدی از شبکه‌های عصبی مصنوعی براساس مفاهیم پایه‌ی منطق و مجموعه‌های فازی، به منظور حصول نتایج دقیق‌تر در موقعیت‌هایی با دوره‌های کوتاه‌تری از زمان ارائه شده است. نتایج حاصله در پیش‌بینی نرخ ارز بیانگر کارایی روش مذکور در پیش‌بینی نرخ ارز نسبت به مدل‌های تشکیل‌دهنده‌ی خود است.

واژگان کلیدی: شبکه‌های عصبی مصنوعی، رگرسیون فازی، مدل‌های ترکیبی، پیش‌بینی سری‌های زمانی، نرخ ارز.

۱. مقدمه

به کارگیری روش‌های کمی به منظور پیش‌بینی بازارهای مالی، بهبود تصمیم‌گیری‌ها و سرمایه‌گذاری به ضرورتی انکارناپذیر در دنیای امروز تبدیل شده است. پیش‌بینی سری‌های زمانی یکی از مهم‌ترین زمینه‌های پیش‌بینی است که در آن، مشاهدات گذشته‌ی یک متغیر جمع‌آوری و به منظور به دست آوردن روابط اساسی بین مشاهدات و تعیین یک مدل توصیفی، مورد تجزیه و تحلیل قرار می‌گیرد؛ سپس مدل حاصله به منظور برون‌یابی سری‌های زمانی در آینده به کار گرفته می‌شود. این روش مدل‌سازی به‌ویژه زمانی مفید است که در مورد فرایند اساسی تولید داده‌ها دانش کمی در دسترس باشد، یا هیچ مدل توضیحی رضایت‌بخشی که متغیر وابسته را به سایر متغیرهای توضیحی مرتبط سازد وجود نداشته باشد.^[۱]

شبکه‌های عصبی مصنوعی از مهم‌ترین روش‌های غیرخطی پیش‌بینی سری‌هایی هستند که در صورت فراهم‌بودن شرایط مطلوب، نتایج مطلوبی نیز ایجاد خواهند کرد. مزیت اصلی این روش قابلیت مدل‌سازی غیرخطی انعطاف‌پذیر است. در این‌گونه شبکه‌ها دیگر نیازی به تشخیص شکل خاص مدل نیست و مدل براساس اطلاعات موجود در داده‌ها شکل می‌گیرد. این رویکرد مبتنی بر داده برای بسیاری از مجموعه داده‌های تجربی، به‌ویژه زمانی که هیچ اطلاعات نظری برای پیشنهاد فرایند تولید داده‌ی مناسب در دسترس نباشد، بسیار کارآمد است. با وجود تمامی مزیت‌هایی که برای شبکه‌های عصبی مصنوعی در نظر گرفته می‌شود، این شبکه‌ها معایبی نیز

* نویسنده مسئول

تاریخ دریافت: ۱۳۹۰/۱۴/۱۷، اصلاحیه ۱۳۹۰/۵/۱۷، پذیرش: ۱۳۹۰/۱۲/۲۹

khasehi@in.iut.ac.ir
bijari@cc.iut.ac.ir
farimah@cc.iut.ac.ir

دارند که از مهم‌ترین آنها می‌توان نیاز به داده‌های زیاد برای کسب نتایج دقیق اشاره کرد.^[۲]

هرچند قانون تعریف‌شده‌ی برای تعداد داده‌های مورد نیاز شبکه، در مسائل مختلف تعیین نشده‌ا، در حالت کلی تعداد داده‌های مورد نیاز شبکه بستگی دارد به ساختار شبکه، الگوریتم آموزش، پیچیدگی مسئله‌ی مورد نظر و اختلالات موجود در داده‌ها. با آزمایش تأثیرات اندازه‌های متفاوت داده‌های مورد نیاز شبکه مشخص شد که افزایش تعداد داده‌ها، عملکرد شبکه‌ی عصبی مصنوعی را در مسائل پیش‌بینی بهینه می‌کند و نیز، در حالت کلی به دست آوردن سطح معینی از دقت، احتیاج به داده‌های آموزش با مقدار معین دارد.^[۳] با تعداد داده‌های کافی، شبکه‌های عصبی می‌توانند ساختارهای پیچیده‌ی موجود در داده‌ها را مدل کنند. بنابراین برای نمونه‌های بزرگ، شبکه‌های عصبی سودمندتر از سایر مدل‌های آماری خواهد بود. البته باید به این نکته نیز توجه داشت که شبکه‌های عصبی مصنوعی به منظور رسیدن به عملکرد مناسب، در مقایسه با مدل‌های خطی نیازمند تعداد داده‌های بیشتر نیستند. مطالعات به عمل آمده نشان داد که شبکه‌های عصبی مصنوعی پیش‌بینی‌های کاملاً خوبی با داده‌های کم‌تر از ۵۰ خواهند داشت، در حالی که مدل‌های باکس-جکینز برای پیش‌بینی‌های موفق دست کم به ۵۰ داده نیاز دارند.^[۴]

روش‌های پیش‌بینی فازی -- همچون رگرسیون فازی -- به دلیل استفاده از اعداد فازی به جای اعداد قطعی، نسبت به سایر مدل‌های پیش‌بینی احتیاج به داده‌های کم‌تری دارند، اما عملکرد این‌گونه روش‌ها عموماً چندان رضایت‌بخش

نیست. به منظور جلوگیری از خطای مدل‌سازی، رگرسیون فازی که اساساً یک مدل پیش‌بینی فاصله‌ی است پیشنهاد شده است، اما این مدل نیز معایبی دارد که از مهم‌ترین آنها می‌توان به وسع شدن بیش از حد بازه‌ی پیش‌بینی، به علت وجود برخی از مقادیر پرت، اشاره کرد. سری‌های زمانی فازی براساس معادلات فازی و منطق تقریبی مدل‌سازی و مطرح شدند.^[۷] همچنین یک روش سری زمانی براساس سری‌های زمانی و مفاهیم سانگ پیشنهاد شده است.^[۸]

استفاده از مدل‌های ترکیبی یا ترکیب مدل‌های مختلف روشی معمول برای بهبود دقت پیش‌بینی‌هاست. ادبیات موضوع مربوط به مدل‌های ترکیبی بسیار گسترده بوده و از زمان اولین تحقیقات انجام شده در این زمینه، تاکنون مطالعات فراوانی صورت گرفته است.^[۹-۱۱] ایده‌ی اصلی ترکیب چندین مدل در پیش‌بینی، استفاده از هر یک از مدل‌های تکی برای مدل‌سازی الگوهای متفاوت موجود در داده‌هاست. یافته‌های تجربی و نظری نشان می‌دهند که ترکیب مدل‌های متفاوت شیوه‌ی مؤثر و کارآ برای بهبود دقت پیش‌بینی‌هاست.^[۱۲، ۱۳] در تحقیقات مربوط به پیش‌بینی با شبکه‌های عصبی مصنوعی مدل‌های ترکیبی متعددی ارائه شده است.^[۱۴]

به منظور پیش‌بینی سری‌های زمانی، روشی برای ترکیب شبکه‌های عصبی مصنوعی و روش میانگین متحرک خودرگرسیون انباشته ارائه شد.^[۱۵] به منظور رفع محدودیت‌های موجود در روش‌های ترکیبی خطی و غیر خطی سنتی نیز، یک روش ترکیبی با استفاده از شبکه‌های عصبی چندلایه‌ی پیش‌خور و مدل‌های باکس-جنکینز ارائه شده است.^[۱۶] در مطالعه‌ی دیگر، رویکردی مرکب از ماشین‌های بردار پشتیبان و مدل‌های خودرگرسیون، میانگین متحرک انباشته را به منظور پیش‌بینی مورد استفاده قرار داده است.^[۱۷] برخی از محققین نیز یک روش ترکیبی با استفاده از مدل سری‌های زمانی فصلی با شبکه‌های عصبی مصنوعی پیشنهاد کرده‌اند.^[۱۸] در پژوهش‌های دیگر نیز ترکیب چند شبکه‌ی عصبی پیش‌رو به منظور بهبود دقت پیش‌بینی سری‌های زمانی پیشنهاد شده است.^[۱۹]

به منظور پیش‌بینی سری‌های زمانی فصلی، یک مدل ترکیبی، مرکب از مدل‌های فصلی خودرگرسیون میانگین متحرک انباشته و ماشین‌های بردارهای پشتیبان ساخته شد.^[۲۰] با به کارگیری شیوه‌ی الگوریتم ژنتیک نیز روشی ترکیبی از شبکه‌های عصبی مصنوعی ارائه و در بازارهای بورس اوراق بهادار به کار گرفته شد.^[۲۱] در تحقیقی مشابه، یک مدل ترکیبی با استفاده از شبکه‌های عصبی مصنوعی و الگوریتم ژنتیک به منظور پیش‌بینی الگوهای موقتی در بازارهای بورس اوراق بهادار ارائه شد.^[۲۲] به منظور پیش‌بینی بازارهای مالی، مدلی مرکب از شبکه‌های عصبی مصنوعی با استفاده از مدل‌های خودرگرسیون خطی تعمیم یافته ارائه شد.^[۲۳]

با ترکیب شبکه‌های عصبی احتمالی و مدل‌های بهبود یافته‌ی شبکه‌های چندلایه‌ی پیش‌خور، مدل ترکیبی جدیدی به منظور پیش‌بینی سری‌های زمانی ارائه شد و در بازار ارز مورد استفاده قرار گرفت.^[۲۴] یک مدل ترکیبی جدید از شبکه‌های پرسپترون چندلایه با استفاده از نگاشت‌های خودسازمان‌ده کوهنن ارائه شد و در پیش‌بینی قیمت الکتریسته مورد استفاده قرار گرفت.^[۲۵] به منظور پیش‌بینی سری‌های زمانی، از مدلی مرکب از شبکه‌های عصبی برگردنده‌ی المان با مدل‌های خودرگرسیون میانگین متحرک انباشته استفاده شد.^[۲۶]

به منظور پیش‌بینی قیمت الکتریسته نیز، مدلی مرکب از شبکه‌های عصبی شعاعی محور و مدل‌های خودرگرسیون میانگین متحرک انباشته ارائه و به کار گرفته شد.^[۲۷] به منظور پیش‌بینی قیمت محصولات نیز یک مدل ترکیبی ارائه شد که در آن شبکه‌های عصبی خودسازمان‌ده با الگوریتم ژنتیک و مدل‌های فازی با یکدیگر ترکیب شده‌اند.^[۲۸] محققین به منظور ارائه‌ی یک مدل ترکیبی پیش‌بینی سیستم‌های فازی را با رگرسیون غیرخطی ترکیب کرده و در پیش‌بینی سری‌های زمانی به کار گرفته‌اند.^[۲۹]

در این مقاله به منظور پیش‌بینی‌های بهتر، کسب نتایج دقیق‌تر، و همچنین نیاز به داده‌های کم‌تر در شبکه‌های عصبی مصنوعی، از مفاهیم پایه‌ی منطق فازی به منظور پیشنهاد روشی ترکیبی استفاده شده است. در روش پیشنهادی پارامترهای شبکه‌های عصبی به جای اعداد قطعی به صورت اعداد فازی در نظر گرفته می‌شوند. با اضافه کردن ابهام به ساختار شبکه‌های عصبی در قالب اعداد فازی، و استفاده از مزایای منحصر به فرد محاسبات نرم مجموعه‌های فازی، تعداد داده‌های مورد نیاز شبکه برای آموزش و جست‌وجوی فضاهای محتمل در مورد مقادیر بهینه‌ی پارامترها کاهش می‌یابد. بدین ترتیب شبکه‌ی عصبی پارامترهای خود را، در شرایطی که داده‌های قابل حصول برای شبکه کم‌تر از مقدار معمول آن است، به صورتی کارا تر تنظیم می‌کند.

به منظور نشان دادن کارایی مدل پیشنهادی در پیش‌بینی سری‌های زمانی در بازارهای مالی و در شرایطی که داده‌های قابل حصول اندک هستند، مدل پیشنهادی در پیش‌بینی نرخ ارز (دلار آمریکا در مقابل ریال ایران) به کار گرفته شده و نتایج آن با سایر مدل‌ها مقایسه شده است. سایر قسمت‌های این مطالعه چنین مرتب شده است: در بخش دوم شبکه‌های عصبی مصنوعی در مدل‌سازی و پیش‌بینی سری‌های زمانی شرح داده شده‌اند. مدل رگرسیون فازی در بخش سوم توضیح داده شده است و در بخش چهارم مدل ترکیبی شبکه‌های عصبی مصنوعی با رگرسیون فازی تشریح شده است. استفاده از مدل ترکیبی در پیش‌بینی نرخ ارز (دلار آمریکا در مقابل ریال ایران) و مقایسه‌ی عملکرد آن با مدل شبکه‌های عصبی مصنوعی و سایر مدل‌ها در بخش پنجم آورده شده است. در پایان، نتایج حاصله مورد تجزیه و تحلیل قرار گرفته و نتیجه‌گیری نهایی ارائه شده است.

۲. شبکه‌ی عصبی مصنوعی

شبکه‌ی عصبی مصنوعی از جمله روش‌های قادر به تخمین موارد غیرخطی متعدد در داده‌هاست و یک چارچوب محاسبه‌ی اعطاف پذیر برای دامنه‌ی وسیعی از مسائل غیرخطی به شمار می‌رود. یکی از مزیت‌های بارز این گونه مدل‌ها، نسبت به مدل‌های غیرخطی دیگر، این است که شبکه‌ی عصبی مصنوعی یک تقریب‌زننده‌ی جهانی هستند و می‌توانند هر نوع تابعی را با دقت دلخواه تقریب بزنند.^[۳۰] نقطه‌ی قوت شبکه‌های عصبی پردازش موازی اطلاعات داده‌هاست؛ این گونه شبکه‌ها به هیچ‌گونه پیش‌فرضی در مورد شکل مدل در فرایند مدل‌سازی نیاز ندارند و به طور کلی یک مدل مبتنی بر داده‌اند. شبکه‌های عصبی پیش‌رو با یک لایه‌ی مخفی از پرکاربردترین مدل‌های شبکه‌های عصبی به منظور پیش‌بینی سری‌های زمانی هستند.^[۳۱] این گونه مدل‌ها از سه لایه پردازش ساده‌ی اطلاعات متصل به هم تشکیل شده‌اند. رابطه‌ی بین خروجی (y_t) و ورودی‌ها $(y_{t-1}, \dots, y_{t-p})$ در این گونه شبکه‌ها چنین است:

$$y_t = f(b_0 + \sum_{j=1}^q w_{j,t} \cdot g(b_{0,j} + \sum_{i=1}^p w_{i,j} \cdot y_{t-i})) + \varepsilon_t \quad (۱)$$

به طوری که $w_{i,j}$ ، $w_{j,t}$ پارامترهای مدل هستند که غالباً وزن‌های اتصال نامیده می‌شوند؛ p تعداد گره‌های ورودی و q تعداد گره‌های مخفی‌اند. f نیز به ترتیب توابع فعال‌سازی لایه‌ی میانی و لایه‌ی خروجی‌اند. توابع سیگموئیدی (رابطه‌ی ۲) و خطی از جمله توابعی هستند که اغلب به عنوان توابع فعال‌سازی برای لایه‌های مخفی و خروجی مورد استفاده قرار می‌گیرند.

$$Sig(x) = \frac{1}{1 + \exp(-x)} \quad (۲)$$

توزیع در ارتباط با مشاهدات حقیقی به کار گرفته می‌شوند. یک مدل رگرسیون خطی فازی در حالت کلی عبارت است از:

$$\tilde{Y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \dots + \tilde{\beta}_n x_n = \sum_{i=1}^n \tilde{\beta}_i x_i = X' \tilde{\beta} \quad (4)$$

به طوری که X بردار متغیرهای مستقل، علامت پریم ($'$) عملگر ترانزپوز، n تعداد متغیرها و $\tilde{\beta}_i$ مجموعه‌های فازی بیانگر نامین پارامتر مدل هستند. این اعداد فازی (پارامترهای $\tilde{\beta}_i$) به شکل اعداد فازی نوع ال «دابیوس و پریس»^[۲۵] $(\alpha_i, c_i)_L$ با توزیع احتمال زیر هستند:

$$\mu_{\tilde{\beta}_i}(\beta_i) = L\{(\alpha_i - \beta_i)/c\} \quad (5)$$

که در آن L یک تابع است. پارامترهای فازی نیز به شکل اعداد فازی مثلثی متقارن به کار گرفته شده‌اند:

$$\mu_{\tilde{\beta}_i}(\beta_i) = \begin{cases} 1 - \frac{|\alpha_i - \beta_i|}{c_i} & \alpha_i - c_i \leq \beta_i \leq \alpha_i + c_i, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

به طوری که $\mu_{\tilde{\beta}_i}(\beta_i)$ تابع عضویت مجموعه‌ی فازی بیانگر پارامترهای β_i است. α_i نیز مرکز عدد فازی و c_i گسترش حول مرکز است. حال با توجه به اصل گسترش، تابع عضویت عدد فازی $\tilde{y}_i = X'_t \tilde{\beta}$ را می‌توان چنین تعریف کرد:

$$\mu_{\tilde{y}_i}(y_t) = \begin{cases} 1 - \frac{|y_t - x_t \alpha|}{c' |x_t|} & \text{for } x_t \neq 0, \\ 1 & \text{for } x_t = 0, y_t = 0, \\ 0 & \text{for } x_t = 0, y_t \neq 0, \end{cases} \quad (7)$$

به طوری که α و c به ترتیب بردار مقادیر مربوط به پارامترها و گسترش‌های آنها حول مرکز هستند. به طور کلی مدل از کمیته‌سازی کل ابهامات (برابر با مجموع گسترش‌های تک‌ی و مربوط به هر یک از پارامترهای فازی مدل) استفاده می‌کند.

$$\text{Minimize } S = \sum_{t=1}^k c' |x_t| \quad (8)$$

این روش به طور همزمان شرایطی را در نظر می‌گیرد که مقدار عضویت به ازای هر مشاهده‌ی y_t بزرگ‌تر از حد آستانه‌ی تعیین شده در سطح h است ($h \in [0, 1]$). این معیار بیانگر این حقیقت است که خروجی فازی مدل باید برای تمامی نقاط داده‌ی y_1, y_2, \dots, y_k بیشتر از مقدار انتخابی سطح h باشد. انتخاب مقدار سطح h برگسترش‌های پارامترهای فازی مدل مؤثر است.

$$\mu_{\tilde{y}_i}(y_t) \geq h \quad \text{for } t = 1, 2, \dots, k \quad (9)$$

شاخص t به تعداد داده‌های غیرفازی به کار گرفته شده در ساخت مدل برمی‌گردد. مسئله‌ی یافتن پارامترهای رگرسیون فازی توسط تاناکا به صورت یک برنامه‌ریزی خطی فرموله شده است.^[۵]

$$\begin{aligned} \text{Minimize } S &= \sum_{t=1}^k c' |x_t| \\ &x'_t \alpha + (1-h)c' |x_t| \geq y_t \quad t = 1, 2, \dots, k \\ &x'_t \alpha - (1-h)c' |x_t| \leq y_t \quad t = 1, 2, \dots, k \\ &c \geq 0 \end{aligned} \quad (10)$$

از این رو مدل شبکه‌های عصبی مصنوعی (رابطه‌ی ۱) در حقیقت به عنوان یک نگاشت غیرخطی از مشاهدات گذشته به مقدار آینده خواهد بود:

$$y_t = f(y_{t-1}, \dots, y_{t-p}, w) + \varepsilon_t \quad (3)$$

به طوری که w بردار همه‌ی پارامترها و f تابعی است که توسط ساختار شبکه‌ی عصبی و وزن‌های انصالی تعیین می‌شود. لذا شبکه‌ی عصبی معادل یک مدل خودرگرسیون غیرخطی است. رابطه‌ی ۳ همچنین بیان می‌دارد که یک نرون (خروجی) در لایه‌ی خروجی برای پیش‌بینی یک مرحله‌ی مورد استفاده قرار می‌گیرد.

مدل ساده‌ی بیان‌شده در رابطه‌ی ۳ دارای قدرت تقریب بسیار زیادی است، چرا که اگر تعداد گره‌های مخفی (q) به اندازه‌ی کافی بزرگ باشند، قادر است هر تابع دلخواه را تقریب بزند. ساختار شبکه‌ی عصبی ساده، که تعداد نرون کم‌تری در لایه‌ی مخفی دارد، در پیش‌بینی‌های خارج از نمونه غالباً بهتر عمل می‌کند که علت آن دوری جستن از مسئله‌ی برازش بیش از حد است که عمدتاً در فرایند مدل‌سازی شبکه‌های عصبی رخ می‌دهد.^[۲۲]

یک مدل بیش از حد برازش‌شده، برای داده‌هایی که برای ساخت مدل (داده‌های آموزش) استفاده شده‌اند، برازش خوبی نشان می‌دهد اما توانایی تعمیم آن، برای داده‌های خارج از نمونه بسیار ضعیف است.^[۲۳] انتخاب تعداد نرون‌های لایه‌ی مخفی وابسته به داده‌هاست و هیچ قانون سیستماتیکی به منظور تعیین این پارامتر در طراحی شبکه‌های عصبی وجود ندارد. علاوه بر این، مرحله‌ی مهم دیگر مدل‌سازی شبکه‌های عصبی، انتخاب تعداد وقعه‌های موجود در مشاهدات به عنوان بردار ورودی شبکه است که شاید مهم‌ترین پارامتر در طراحی مدل‌های شبکه‌های عصبی مصنوعی باشد، چرا که این پارامتر در تعیین ساختار خودهمبسته‌ی غیرخطی سری‌های زمانی نقش اساسی دارند.

برای انتخاب p هیچ نظریه‌ی خاصی وجود ندارد؛ از این رو تجربه و آزمایشات مشخص‌کننده‌ی مقدار مناسب p و q خواهد بود. وقتی ساختار شبکه‌ی p و q مشخص شد شبکه‌ی آماده‌ی آموزش و تخمین پارامترهاست. همانند مدل‌سازی میانگین متحرک خودرگرسیون انباشته‌ی پارامترها طوری تخمین زده می‌شود که معیار دقت مانند میانگین مربعات خطاها کمیته شود. فرایند آموزش مذکور توسط الگوریتم‌های بهینه‌سازی غیرخطی همچون پس‌انتشار خطا انجام می‌گیرد. مدل برآورد شده معمولاً توسط نمونه‌ی جداگانه‌ی p که در فرایند آموزش مورد استفاده قرار نگرفته است (داده‌های آزمون) مورد ارزیابی قرار می‌گیرد و این با فرایند مدل‌سازی مدل‌های میانگین متحرک خودرگرسیون انباشته که از یک نمونه به منظور شناسایی، تخمین و ارزیابی مدل استفاده می‌کند متفاوت است. دلیل این موضوع نیز آن است که شکل کلی مدل (خطی) برای مدل‌سازی میانگین متحرک خودرگرسیون انباشته از قبل مشخص است.^[۲۴] در صورتی که در مدل‌سازی شبکه‌های عصبی چنین نیست و در مدل‌سازی با شبکه‌های عصبی مصنوعی مدل غیرخطی همانند مرتبه‌ی مدل باید تخمین زده شود و ممکن است شبکه دچار برازش بیش از حد شود.

۳. مدل رگرسیون فازی

مدل‌های کلاسیک از مفهوم عبارت خطا استفاده می‌کنند؛ به عبارت دیگر تخمین‌های این‌گونه از مدل‌ها مقادیر دقیقی است و شامل جمله‌ی خطا نمی‌شود و این همان مفهوم پایه‌ی رگرسیون فازی است.^[۲] مفهوم اساسی نظریه‌ی فازی و رگرسیون فازی این است که جمله‌ی خطا از باقی‌مانده‌های بین مقادیر تخمین زده شده و مقادیر اصلی یا مشاهدات تولید نمی‌شود، بلکه در عدم قطعیت پارامترهای مدل و امکان

آورده شده است. حال با در نظر گرفتن سطح آستانه‌ی h برای میزان توابع عضویت تمامی مشاهدات مطابق رابطه‌ی ۹، مدل برنامه‌ریزی غیرخطی مطابق رابطه‌ی ۱۷ به دست می‌آید.

$$\mu_{\tilde{w}_{i,j}}(w_{i,j}) = \begin{cases} \frac{1}{b_{i,j} - a_{i,j}} (w_{i,j} - a_{i,j}) & \text{if } a_{i,j} \leq w_{i,j} \leq b_{i,j}, \\ \frac{1}{b_{i,j} - c_{i,j}} (w_{i,j} - c_{i,j}) & \text{if } b_{i,j} \leq w_{i,j} \leq c_{i,j}, \\ 0 & \text{otherwise,} \end{cases} \quad (13)$$

$$\mu_{\tilde{X}_{t,j}}(x_{t,j}) = \begin{cases} \frac{\left(x_{t,j} - g\left(\sum_{i=1}^p a_{i,j} \cdot y_{t-i}\right)\right)}{g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) - g\left(\sum_{i=1}^p a_{i,j} \cdot y_{t-i}\right)} & \text{if } g\left(\sum_{i=1}^p a_{i,j} \cdot y_{t-i}\right) \leq x_{t,j} \leq g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right), \\ \frac{\left(x_{t,j} - g\left(\sum_{i=1}^p c_{i,j} \cdot y_{t-i}\right)\right)}{g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) - g\left(\sum_{i=1}^p c_{i,j} \cdot y_{t-i}\right)} & \text{if } g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) \leq x_{t,j} \leq g\left(\sum_{i=1}^p c_{i,j} \cdot y_{t-i}\right), \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

$$\mu_{\tilde{w}_j}(w_j) = \begin{cases} \frac{1}{e_j - d_j} (w_j - d_j) & \text{if } d_j \leq w_j \leq e_j, \\ \frac{1}{e_j - f_j} (w_j - f_j) & \text{if } e_j \leq w_j \leq f_j, \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

$$\mu_{\tilde{Y}}(y_t) \cong \begin{cases} \left[\frac{-B_1}{rA_1} + \left[\left(\frac{B_1}{rA_1} \right)^2 - \frac{C_1 - f^{-1}(y_t)}{A_1} \right]^{1/2} \right]^{1/r} & \text{if } C_1 \leq f^{-1}(y_t) \leq C_2, \\ \left[\frac{B_2}{rA_2} + \left[\left(\frac{B_2}{rA_2} \right)^2 - \frac{C_2 - f^{-1}(y_t)}{A_2} \right]^{1/2} \right]^{1/r} & \text{if } C_2 \leq f^{-1}(y_t) \leq C_3, \\ 0 & \text{otherwise,} \end{cases} \quad (16)$$

به قسمی که:

$$A_1 = \sum_{j=1}^q (e_j - d_j) \cdot \left(g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) - g\left(\sum_{i=1}^p a_{i,j} \cdot y_{t-i}\right) \right),$$

$$B_1 = \sum_{j=1}^q (d_j \cdot \left(g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) - g\left(\sum_{i=1}^p a_{i,j} \cdot y_{t-i}\right) \right) + g\left(\sum_{i=1}^p a_{i,j} \cdot y_{t-i}\right) \cdot (e_j - d_j)),$$

$$A_2 = \sum_{j=1}^q (f_j - e_j) \cdot \left(g\left(\sum_{i=1}^p c_{i,j} \cdot y_{t-i}\right) - g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) \right),$$

$$B_2 = \sum_{j=1}^q (f_j \cdot \left(g\left(\sum_{i=1}^p c_{i,j} \cdot y_{t-i}\right) - g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) \right) + g\left(\sum_{i=1}^p b_{i,j} \cdot y_{t-i}\right) \cdot (f_j - e_j)),$$

به طوری که $\alpha' = (\alpha_1, \alpha_2, \dots, \alpha_n)$ و $c' = (c_1, c_2, \dots, c_n)$ بردار متغیرهای مجهول، و S کل ابهامی است که قبلاً تعریف شده است.

۴. فرموله کردن مدل ترکیبی

شبکه‌های عصبی مصنوعی مدل‌های پیش‌بینی دقیقی برای دامنه‌ی وسیعی از مسائل هستند، اما دارای محدودیت تعداد زیاد داده‌های گذشته به منظور کسب نتایج دقیق‌اند.^[۱] این در حالی است که امروزه به علت عدم قطعیت محیط و توسعه‌ی سریع فناوری‌های نوین، معمولاً باید موقعیت‌های آتی را با استفاده از داده‌های کم و در بازه‌های زمانی کوتاه مدت پیش‌بینی کرد. بنابراین در دنیای امروز پیش‌بینی روش‌هایی مورد نیاز است که داده‌های کم‌تری لازم داشته باشد و در چنین شرایطی کارایی لازم را باشند. مدل رگرسیون فازی یک مدل پیش‌بینی بازه‌ی مناسب در شرایط داده‌های قابل حصول کم است، اما همیشه عملکرد مناسبی ندارند. اساساً هدف اصلی این مقاله، بررسی صحت این فرضیه است که آیا فازی‌سازی شبکه‌های عصبی مصنوعی باعث افزایش دقت آنها و دست‌یابی به نتایج دقیق‌تر خواهد شد یا نه؟ اگرچه به لحاظ نظری، به کارگیری منطق فازی غالباً ابزاری مناسب برای رفع محدودیت تعداد داده‌های مورد نیاز شبکه‌های عصبی، و دست‌یابی به نتایج دقیق‌تر در این‌گونه از شرایط داده‌ی می‌شود، این موضوع در حالت کلی اثبات نشده است. لذا در این مطالعه به منظور نشان دادن درستی این موضوع در کاربرد خاص مطروحه، از نتایج تجربی حاصله استفاده شده است.

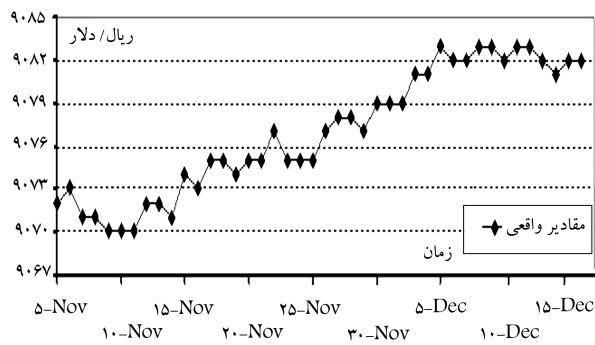
پارامترهای شبکه‌های عصبی (وزن‌ها و بایاس‌ها) مقادیری قطعی‌اند $(b_{i,j}, b_{0,j})$ و $(w_j, w_{i,j})$ ، در صورتی که در روش جدید به جای به کارگیری این مقادیر قطعی، پارامترهای شبکه به صورت اعداد فازی در نظر گرفته شده‌اند $(\tilde{w}_j, \tilde{w}_{i,j}, \tilde{b}_{0,j}, \tilde{b}_{i,j})$. علاوه بر این، متدولوژی ارائه شده در این نوشتار برای شرایطی که دامنه‌ی پیش‌بینی وسیع می‌شود،^[۲۶] به کار گرفته شده است. در حالت کلی، یک مدل ترکیبی با پارامترهای فازی به صورت زیر است:^[۲۷]

$$\tilde{y}_t = f(\tilde{b}_0 + \sum_{j=1}^q \tilde{w}_j \cdot g(\tilde{b}_{0,j} + \sum_{i=1}^p \tilde{w}_{i,j} \cdot y_{t-i})) \quad (11)$$

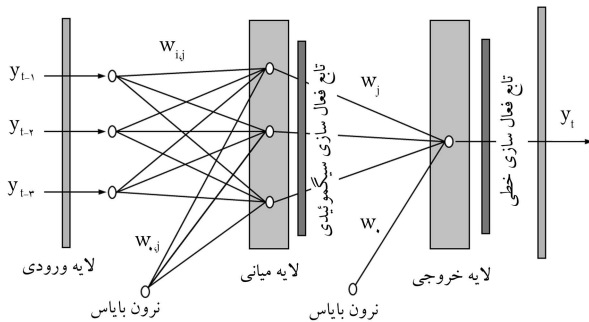
به قسمی که y_t مشاهدات $(\tilde{w}_j, \tilde{w}_{i,j}, \tilde{b}_{0,j}, \tilde{b}_{i,j})$ اعداد فازی، و f و g به ترتیب توابع فعال‌سازی لایه میانی و لایه‌ی خروجی‌اند. با یک جایگذاری، معادله‌ی ۱۱ تبدیل می‌شود به:

$$\tilde{y}_t = f(\tilde{b}_0 + \sum_{j=1}^q \tilde{w}_j \cdot \tilde{X}_{t,j}) = f\left(\sum_{j=0}^q \tilde{w}_j \cdot \tilde{X}_{t,j}\right) \quad (12)$$

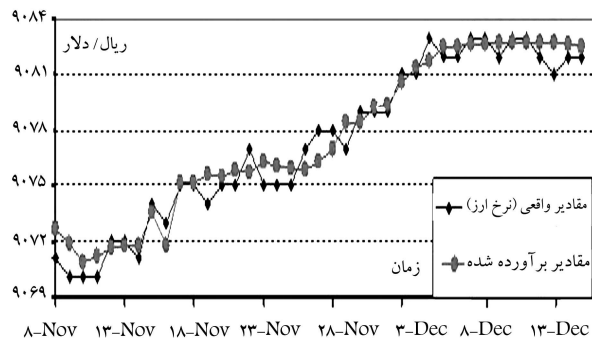
به قسمی که $\tilde{X}_{t,j} = g(\tilde{b}_{0,j} + \sum_{i=1}^p \tilde{w}_{i,j} \cdot y_{t-i})$ است. وزن‌های اتصالی مابین لایه‌ی ورودی و لایه‌ی میانی در این معادله به صورت اعداد فازی مثلثی $(\tilde{w}_{i,j} = (a_{i,j}, b_{i,j}, c_{i,j}))$ (رابطه‌ی ۱۳) در نظر گرفته شده است. به طوری که $\mu_{\tilde{w}}(w_{i,j})$ تابع عضویت هر مجموعه‌ی فازی است که با پارامترهای $w_{i,j}$ مشخص می‌شود. حال با استفاده از اصل گسترش،^[۲۸] تابع عضویت $\tilde{X}_{t,j} = g(\sum_{i=1}^p \tilde{w}_{i,j} \cdot y_{t-i})$ در حالت کلی در پیوست ۱ آورده شده است. با در نظر گرفتن داده‌های فازی مثلثی $\tilde{X}_{t,j}$ با تابع عضویتی (مطابق رابطه‌ی ۱۴) و پارامترهای فازی مثلثی w_j (مطابق رابطه‌ی ۱۵) تابع عضویت $\tilde{y}_t = f(\tilde{b}_0 + \sum_{j=1}^q \tilde{w}_j \cdot \tilde{X}_{t,j}) = f(\sum_{j=0}^q \tilde{w}_j \cdot \tilde{X}_{t,j})$ مطابق رابطه‌ی ۱۶ خواهد بود. اثبات رابطه‌ی مزبور در حالت کلی در پیوست ۲



شکل ۱. داده‌ی روزانه‌ی نرخ ارز (دلار آمریکا در مقابل ریال ایران) از ۱۴ آبان ماه تا ۲۵ آذرماه سال ۱۳۸۴.



شکل ۲. ساختار شبکه‌ی عصبی چندلایه‌ی پیشخور.



شکل ۳. مقادیر واقعی و برآورده‌شده توسط شبکه‌ی عصبی.

فاز ۱. با توجه به اصول طراحی شبکه‌های عصبی مصنوعی [۴۱] و به‌کارگیری نرم‌افزار MATLAB شبکه‌ی ایدال، یک شبکه‌ی $N^{(2-3-1)}$ است که در شکل ۲ نشان داده شده است. وزن‌ها و بایاس‌های مذکور نیز در جدول ۱ ارائه شده است.

همچنین مقادیر واقعی و برازش‌شده توسط شبکه در شکل ۳ نشان داده شده است. معیارهای عملکردی شبکه‌ی مذکور نیز در جدول ۲ ارائه شده است.

فاز ۲. (تعیین ابهام کمینه): با قراردادن اوزان به دست آمده $(w_0, w_1, w_2, w_3) = (0, 5847, 0, 5847, 985, 3, 0, 278)$ پارامترهای فازی مدل با استفاده از معادله‌ی ۱۷ به دست آمده $(h = 0)$ و نتایج حاصله در شکل ۴ آورده شده است.

$$\tilde{y}_t = (0, 5847, 0, 7000) x_{1,t} + (958, 3, 0, 71069) x_{2,t} + (0, 2781, 0, 7000) x_{3,t} \quad (18)$$

$$C_1 = \sum_{j=0}^q \left(d_j \cdot g \left(\sum_{i=0}^p a_{i,j} \cdot y_{t-i} \right) \right)$$

$$C_r = \sum_{j=0}^q \left(f_j \cdot g \left(\sum_{i=0}^p c_{i,j} \cdot y_{t-i} \right) \right),$$

$$C_r = \sum_{j=0}^q \left(e_j \cdot g \left(\sum_{i=0}^p b_{i,j} \cdot y_{t-i} \right) \right),$$

$$\text{Min} \sum_{t=1}^k \sum_{j=0}^q \left(f_j \cdot g \left(\sum_{i=0}^p c_{i,j} \cdot y_{t-i} \right) \right) - \left(d_j \cdot g \left(\sum_{i=0}^p a_{i,j} \cdot y_{t-i} \right) \right)$$

$$\text{Subject to} \begin{cases} \left[\frac{-B_1}{rA_1} + \left[\left(\frac{B_1}{rA_1} \right)^2 - \frac{C_1 - f^{-1}(y_t)}{A_1} \right]^{1/2} \right] \leq h \\ \text{if } C_1 \leq f^{-1}(y_t) \leq C_r, \text{ for } t = 1, 2, \dots, k \\ \left[\frac{B_r}{rA_r} + \left[\left(\frac{B_r}{rA_r} \right)^2 - \frac{C_r - f^{-1}(y_t)}{A_r} \right]^{1/2} \right] \leq h \\ \text{if } C_r \leq f^{-1}(y_t) \leq C_r, \text{ for } t = 1, 2, \dots, k \end{cases} \quad (17)$$

مراحل روش ترکیبی به‌طور خلاصه در سه فاز انجام می‌شود:

فاز ۱. آموزش یک مدل شبکه‌ی عصبی با استفاده از اطلاعات موجود در مشاهدات (که به‌صورت غیرفازی اند). نتیجه‌ی این فاز، جواب بهینه‌ی پارامترهای شبکه‌ی عصبی مصنوعی $w^* = (w_0^*, w_{i,j}^*)$ (for $j = 0, 1, \dots, q, i = 0, 1, \dots, p$) است که به‌عنوان یکی از مجموعه داده‌های ورودی در فاز دوم مورد استفاده قرار می‌گیرد.

فاز ۲. تعیین کم‌ترین ابهام با استفاده از معادله‌ی ۱۷ و اوزان حاصله $w^* = (w_0^*, w_{i,j}^*)$ (for $j = 0, 1, \dots, q, i = 0, 1, \dots, p$)

فاز ۳. با توجه به نظرات ایشیوچی داده‌های حد بالا و پائین مدل، وقتی که دامنه‌ی مدل ترکیبی وسیع شود، حذف خواهد شد. به‌منظور ساختن مدلی شامل همه‌ی شرایط ممکن، اگر مجموعه داده‌ها شامل تفاوت‌های مشخص یا موارد خارج از محدوده باشند، c_j ها بسیار گسترده خواهند شد. طبق نظرات ایشیوچی داده‌های اطراف مرزهای بالا و پائین مدل حذف می‌شود، سپس مدل مجدداً فرمول‌بندی می‌شود.

۵. به‌کارگیری مدل ترکیبی به‌منظور پیش‌بینی

در این قسمت، از مدل حاصله به‌منظور پیش‌بینی دامنه‌ی تغییرات نرخ ارز (دلار آمریکا در مقابل ریال ایران) استفاده شده است. اطلاعات استفاده شده در این تحقیق شامل ۴۲ داده‌ی روزانه‌ی نرخ ارز (دلار آمریکا در مقابل ریال ایران) از شنبه چهاردهم آبان‌ماه تا جمعه بیست و پنجم آذرماه سال ۱۳۸۴، مطابق با پنجم نوامبر تا شانزدهم دسامبر سال ۲۰۰۵ است [۳۹] که در شکل ۱ نمایش داده شده است.

۱.۵. پیش‌بینی

با به‌کارگیری مدل ترکیبی ابتدا ۳۵ مشاهده (پنج هفته‌ی اول) به‌منظور آموزش شبکه و سپس ۷ مشاهده (هفته‌ی آخر) برای سنجش عملکرد مدل مورد استفاده قرار گرفته است.

جدول ۱. مقادیر مربوط به وزن‌ها و بایاس‌های شبکه.

وزن‌های لایه‌ی ورودی			وزن‌های لایه‌ی مخفی	بایاسها	
$W_{z_1,1}^1$	$W_{z_1,2}^1$	$W_{z_1,3}^1$	Weights	b_j^1	b^2
۳,۷۸۹	۴۲,۱۰۴۱	۱۵۵,۲۶۷	۰,۵۸۴۶۷	-۶,۵۸۳۶	-۹۸۵,۱۹۱
۲,۳۶۵۱	-۱۱,۴۹۳۴	۱۷۱,۷۱۳۷	۹۸۵,۳۱۹۴	۱۱,۴۴۸۵	
۴,۵۲۲۶	-۲۶,۰۸۷۶	۱۵۸,۰۹۰۲	۰,۲۷۸۰۸	-۲۷,۱۲۹۱	

جدول ۲. معیارهای عملکرد شبکه‌ی عصبی در پیش‌بینی نرخ ارز.

آموزش		تست					
MSE	MAE	MSE	SSE	RMSE	ME	MAPE	MAE
۰,۹۳۲	۰,۸۲۵	۰,۶۸۶	۴,۸۰۲	۰,۸۲۸	-۰,۵۳۹	۰,۰۰۰۸	۰,۶۹۲

جدول ۳. حدود بالا و پایین و مقادیر اصلی داده‌های آزمون.

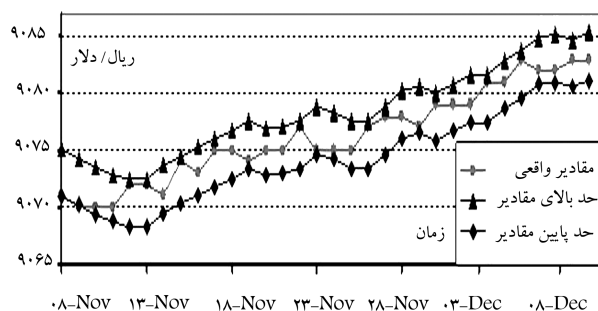
تاریخ	مقادیر واقعی	حد پایین مقادیر		حد بالای مقادیر	
		بعد از حذف	قبل از حذف	بعد از حذف	قبل از حذف
شنبه (۱۹-آذر)	۹۰۸۲	۹۰۸۱	۹۰۸۱	۹۰۸۴	۹۰۸۴
یکشنبه (۲۰-آذر)	۹۰۸۳	۹۰۸۲	۹۰۸۱	۹۰۸۴	۹۰۸۴
دوشنبه (۲۱-آذر)	۹۰۸۳	۹۰۸۲	۹۰۸۱	۹۰۸۴	۹۰۸۴
سه‌شنبه (۲۲-آذر)	۹۰۸۲	۹۰۸۱	۸۰۹۱	۹۰۸۴	۹۰۸۴
چهارشنبه (۲۳-آذر)	۹۰۸۱	۹۰۸۱	۹۰۸۱	۹۰۸۴	۹۰۸۴
پنج‌شنبه (۲۴-آذر)	۹۰۸۲	۹۰۸۱	۹۰۸۱	۹۰۸۴	۹۰۸۴
جمعه (۲۵-آذر)	۹۰۸۲	۹۰۸۱	۹۰۸۱	۹۰۸۴	۹۰۸۴

پیش‌بینی شده‌اند. نتایج مقادیر پیش‌بینی شده در جدول ۳ آورده شده است. این نکته قابل ذکر است که داده‌های مربوط به حد بالا و پایین مقادیر در جدول گرد شده‌اند. چنان‌که مشاهده می‌شود، نتایج حاصله و فواصل فازی نسبت به مدل قبلی مطلوب‌تر است.

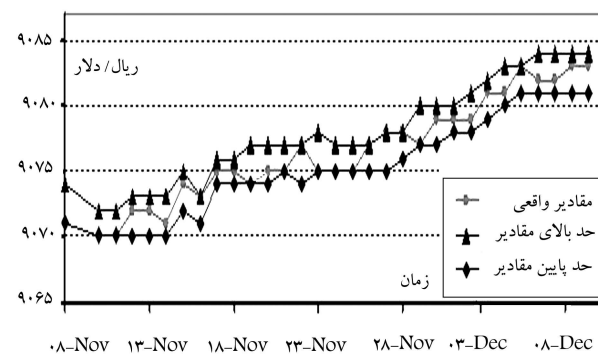
۲.۵. مقایسه با سایر روش‌های مشابه

در این قسمت به منظور نشان دادن برتری و کارایی مدل پیشنهادی در تقابل با سایر روش‌ها در پیش‌بینی سری‌های زمانی در شرایط داده‌های قابل حصول کم، مدل پیشنهادی در پیش‌بینی نرخ ارز (دلار آمریکا در مقابل ریال ایران) با سایر مدل‌ها مقایسه شده است.

این مدل‌ها متشکل‌اند از شبکه‌های عصبی چندلایه‌ی پیش‌خور، مدل‌های خودرگرسیون میانگین متحرک انباشته و مدل‌های ترکیبی میانگین متحرک خودرگرسیون انباشته فازی. از مقایسه‌ی طول بازه پیش‌بینی شده توسط روش ترکیبی پیشنهادی با میانگین قدرمطلق خطای شبکه‌ی چندلایه‌ی پیش‌خور، با فرض توزیع یکنواخت مقادیر واقعی در اطراف تابع برآورد شده و فضای پیش‌بینی، نتایج حاصله بیانگر برتری روش ترکیبی بر شبکه‌های عصبی مصنوعی است. گفتنی است طول بازه



شکل ۴. مقادیر واقعی و حد بالا و پایین آنها (قبل از حذف).



شکل ۵. مقادیر واقعی و حد بالا و پایین آنها (بعد از حذف).

چنان‌که در شکل ۴ مشاهده می‌شود مقادیر واقعی در فواصل فازی قرار گرفته‌اند، اما طول فواصل فازی وسیع‌شده و مدل ترکیبی فواصل مناسبی را ارائه نمی‌دهد. فاز ۳. از نتایج فوق مشخص می‌شود که مشاهده‌ی مربوط به هجدهم آذرماه (نهم نوامبر) در مرز پایینی قرار گرفته است، بنابراین محدودیت خطی که توسط این مشاهده تولید شده، حذف و سپس فاز دوم مجدداً تکرار می‌شود ($h = 0$). نتایج به دست آمده در شکل ۵ آورده شده است.

$$\hat{y}_t = (0,5847, 0,0000) x_{1,t} + (958,3, 0,0940) x_{2,t} + (0,2781, 0,0000) x_{3,t} \quad (19)$$

در انتها نیز با استفاده از مدل ترکیبی بازبینی شده مقادیر آتی متغیر وابسته

۶. نتیجه‌گیری

تغییرات تکنولوژیکی و جهانی‌شدن تجارت و بازارهای مالی باعث شده تا توانایی پیش‌بینی دقیق‌تر و سریع‌تر الگوهای موجود در حفظ توان رقابتی اهمیت بیشتری پیدا کند. اما تغییرات سریع این‌گونه محیط‌های ناشناخته، به‌ویژه بازارهای مالی، پیش‌بینی‌کنندگان را از نظر تأمین داده‌های لازم به‌منظور حصول نتایج مطلوب دچار مشکل کرده است. در این نوشتار براساس اصول و مفاهیم پایه‌ی شبکه‌های عصبی مصنوعی و رگرسیون فازی، مدلی از ترکیب دو روش مزبور به‌منظور پیش‌بینی سری‌های زمانی پیشنهاد شده است. مدل پیشنهادی با فازی در نظر گرفتن پارامترهای شبکه‌ی چندلایه‌ی پیشخور، داده‌های مورد نیاز برای کسب نتایج مطلوب را کاهش داده است. نتایج به دست آمده بیانگر آن است که مدل ترکیبی نه تنها توانایی انجام پیش‌بینی مناسب را دارد بلکه برای تصمیم‌گیرندگان بهترین و بدترین حالت ممکن را نیز فراهم می‌سازد. مدل پیشنهادی همچنین در شرایط برابر عملکرد بهتری نسبت به شبکه‌های چندلایه‌ی پیشخور و روش ترکیبی میانگین متحرک خودرگرسیون انباشته‌ی فازی داشته است.

پیش‌بینی‌شده توسط روش ترکیبی ۲/۷۸ ریال است در حالی که میانگین قدرمطلق خطای شبکه برابر با ۰/۷۶۴ ریال است.

روش ترکیبی نسبت به مدل‌های میانگین متحرک خودرگرسیون انباشته و همچنین روش ترکیبی فازی آن نیز نتایج مطلوب‌تری در پیش‌بینی نرخ ارز ارائه کرده است. مدل‌های ترکیبی میانگین متحرک خودرگرسیون انباشته‌ی فازی، حالت فازی‌شده‌ی این‌گونه از مدل‌ها هستند.^[۳۹] طول ۹۵٪ اطمینان به دست آمده از مدل میانگین متحرک خودرگرسیون انباشته، ۱۶/۲ ریال، و طول بازه حاصل از روش ترکیبی میانگین متحرک خودرگرسیون انباشته‌ی فازی ۴/۵ ریال^[۳۹] است، در صورتی که طول بازه پیش‌بینی‌شده توسط روش ترکیبی ۲/۷۸ ریال است و این خود به‌ترتیب بیانگر بیش از ۸۳٪ و ۳۱٪ بهبود در دقت نتایج نسبت به این دو مدل است. مزایای روش ترکیبی را می‌توان چنین خلاصه کرد:

۱. رفع مسئله‌ی کمبود داده‌ها؛

۲. ارائه‌ی بهترین و بدترین حالات ممکن؛

۳. دستیابی به نتایج دقیق‌تر.

منابع (References)

1. Khashei, M. "Forecasting and analysis of Esfahan steel company productions price in Tehran metals exchange using artificial neural networks", Master of Science Thesis, Isfahan University of Technology (2005).
2. Zhang, G., Patuwo, B.E. and Hu, M.Y. "Forecasting with artificial neural networks: The state of the art", *Neurocomputing*, **56**, pp. 205-232 (2004).
3. Nam, K. and Schaefer, T. "Forecasting international airline passenger traffic using neural networks", *Logistics and Transportation*, **31**(3), pp. 239-251 (1995).
4. Kang, S. "An investigation of the use of feedforward neural networks for forecasting", Ph.D. Thesis, Kent State University (1991).
5. Tanaka, H. "Fuzzy data analysis by possibility linear models", *Fuzzy Sets and Systems*, **24**(3), pp. 363-375 (1987).
6. Tanaka, H. and Ishibuchi, H., *Possibility Regression Analysis Based On Linear Programming*, in: J. Kacprzyk, M. Fedrizzi (Eds.), *Fuzzy Regression Analysis*, Omnitech Press, Warsaw and Physica-Verlag, Heidelberg, pp. 47-60 1992.
7. Song, Q. and Chissom, B.S. "Forecasting enrollments with fuzzytime series -part I", *Fuzzy Sets and Systems*, **54**(1), pp. 1-9 (1993).
8. Chen, S.M. "Forecasting enrollments based on fuzzy time series", *Fuzzy Sets and Systems*, **81**(3), pp. 311-31 (1996).
9. Reid, M.J. "Combining three estimates of gross domestic product", *Economica*, **35**, pp. 431-444 (1968).
10. Bates, J.M.C. and Granger, W.J. "The combination of forecasts", *Oper. Res. Q.*, **20**, pp. 451-468 (1969).
11. Clemen, R., "Combining forecasts: A review and annotated bibliography with discussion", *Int. J. Forecasting*, **5**, pp. 559-608 (1989).
12. Khashei, M. and Bijari, M. "Combination of linear and nonlinear autoregressive integrated moving average and artificial neural networks for improving financial fore-

- casts performance”, *Economic Research Journal*, Tarbiat Modares University, **8**(2), pp. 82-100 (2008).
13. Cadenas, E. and Rivera, W. “Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model”, *Renewable Energy*, **35**, pp. 2732-2738 (2010).
 14. Cheng, M., Tsai, H. and Sudjono, E. “Evolutionary fuzzy hybrid neural network for project cash flow control”, *Engineering Applications of Artificial Intelligence*, **23**, pp. 604-613 (2010).
 15. Zhang, G.P. “Time series forecasting using a hybrid ARIMA and neural network model”, *Neurocomputing*, **50**, pp. 159-175 (2003).
 16. Khashei, M. and Bijari, M. “An artificial neural network (p,d,q) model for time series forecasting”, *Expert Systems with Applications*, **37**, pp. 479-489 (2010).
 17. Pai, P. and Lin, C. “A hybrid ARIMA and support vector machines model in stock price forecasting”, *Omega*, **33**, pp. 497-505 (2005).
 18. Tseng, F., Yu, H. and Tzeng, G. “Combining neural network model with seasonal time series ARIMA model”, *Technological Forecasting & Social Change*, **69**, pp. 71-87 (2002).
 19. Ginzburg, I. and Horn, D. “Combined neural networks for time series analysis”, *Adv. Neural Inf. Process. Systems*, **6**, pp. 224-231 (1994).
 20. Chen, K. and Wang, C. “A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan”, *Expert Systems with Applications*, **32**, pp. 254-264 (2007).
 21. Armano, G., Marchesi, M. and Murru, A. “A hybrid genetic-neural architecture for stock indexes forecasting”, *Information Sciences*, **170**, pp. 3-33 (2005).
 22. Kim, H. and Shin, K. “A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets”, *Applied Soft Computing*, **7**, pp. 569-576 (2007).
 23. Yu, L., Wang, S. and Lai, K. “A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates”, *Computers and Operations Research*, **32**, pp. 2523-2541 (2005).
 24. Khashei, M., Bijari, M. and Raissi, G.H.A. “Improving the hybrid ANN/ ARIMA models with probabilistic neural networks (PNNs) for time series forecasting”, *Industrial Journal*, **44**(2), pp. 181-193 (2010).
 25. Amin-Naseri, M.R. and Soroush, A.R. “Combined use of unsupervised and supervised learning for daily peak load forecasting”, *Energy Conversion and Management*, **49**, pp. 1302-1308 (2008).
 26. Aladag, C., Egrioglu, E. and Kadilar, C. “Forecasting nonlinear time series with a hybrid methodology”, *Applied Mathematics Letters*, **22**, pp. 1467-1470 (2009).
 27. Shafie-khah, M., Parsa Moghaddam, M. and Sheikh-El-Eslami, M.K. “Price forecasting of day-ahead electricity markets using a hybrid forecast method”, *Energy Conversion and Management*, **52**, pp. 2165-2169 (2011).
 28. Chang, P., Liu, C. and Wang, Y. “A hybrid model by clustering and evolving fuzzy rules for sales decision supports in printed circuit board industry”, *Decision Support Systems*, **42**, pp. 1254-1269 (2006).
 29. Lin, Y. and Cobourn, W.G. “Fuzzy system models combined with nonlinear regression for daily ground-level ozone predictions”, *Atmospheric Environment*, **41**, pp. 3502-3513 (2007).
 30. Khashei, M., Bijari, M. and Raissi, G.A. “Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks”, *Neurocomputing*, **72**, pp. 956-967 (2009).
 31. Wong, B.K., Vincent S.L. and Jolie, L. “A bibliography of neural network business applications research: 1994-1998”, *Computers & Operations Research*, **27**, pp. 1023-1044 (2000).
 32. Zhang, G.P., Patuwo, E.B. and Hu, M.Y., “A simulation study of artificial neural networks for nonlinear time-series forecasting”, *Comput. Oper. Res.*, **28**, pp. 381-396 (2001).
 33. Demuth, H. and Beale, B., *Neural Network Toolbox User Guide*, The Mathworks Inc, Natick (2004).
 34. Fildes, R. and Makridakis, S. “The impact of empirical accuracy studies on time series analysis and forecasting”, *Int. Statist. Rev.*, **63**, pp. 289-308 (1995).
 35. Dubois, D. and Prade, H., *Theory and Applications, Fuzzy Sets and Systems*, Academic Press, New York (1980).
 36. Ishibuchi, H. and Tanaka, H. “Interval regression analysis based on mixed 0-1 integer programming problem”, *J. Japan Soc. Ind. Eng.*, **40**(5), pp. 312-319 (1988).
 37. Khashei, M., Hejazi, S.R. and Bijari, M. “A new hybrid artificial neural networks and fuzzy regression model for time series forecasting”, *Fuzzy Set and Systems*, **159**, pp. 769-786 (2008).
 38. Zadeh, L.A., *The Concept of a Linguistic Variable and Its Application to Approximate Reasoning*, Part 1, 2 and 3, Information Sciences 8, pp. 43-80 (1976).
 39. Khashei, M. and Bijari, M. “Foreign exchange rate forecasting using a hybrid fuzzy and auto regressive integrated moving average model”, *Esteghlal Journal*, Isfahan University of Technology, pp. 67-75 (2008).
 40. Khashei, M. and Bijari, M. “Using artificial neural networks (ANNs) to forecast and analysis of esfahan steel company (ESC) productions price in Tehran metals exchange (part I)”, *Steel Journal*, **133**, pp. 28-31 (2006).
 41. Khashei, M. and Bijari, M. “Using artificial neural networks (ANNs) to forecast and analysis of Esfahan steel company (ESC) productions price in Tehran metals exchange (part II)”, *Steel Journal*, **134**, pp. 28-31 (2006).
 42. Hong, D.H. and Yi, H.C. “A note on fuzzy regression model with fuzzy input and output data for manpower forecasting”, *Fuzzy Sets and Systems*, **138**, pp. 301-305 (2003).
 43. Lee, H.T. and Chen, H. “Fuzzy regression model with fuzzy input and output data for manpower forecasting”, *Fuzzy Sets and Systems*, **119**, pp. 205-213 (2001).

$a_i, d_i \geq$ آنکه به شرط کلی، در حالت کلی، به شرط آنکه $\tilde{Y} = \tilde{\beta}_1 \cdot \tilde{x}_1 + \tilde{\beta}_2 \cdot \tilde{x}_2 + \dots + \tilde{\beta}_n \cdot \tilde{x}_n$ مطابق زیر خواهد بود: $^{[۲۲]}$ for $i = 1, 2, \dots, n$

$$\mu_{\tilde{Y}}(y_t) \cong \begin{cases} \left(\frac{-B_1}{rA_1} + \left[\left(\frac{B_1}{rA_1} \right)^2 - \frac{C_1 - f^{-1}(y_t)}{A_1} \right]^{1/2} \right)^{1/r} & \text{if } C_1 \leq f^{-1}(y_t) \leq C_r, \\ \left(\frac{B_r}{rA_r} + \left[\left(\frac{B_r}{rA_r} \right)^2 - \frac{C_r - f^{-1}(y_t)}{A_r} \right]^{1/2} \right)^{1/r} & \text{if } C_r \leq f^{-1}(y_t) \leq C_{r+1}, \\ 0 & \text{otherwise,} \end{cases} \quad (۲۱)$$

به قسمی که:

$$\begin{aligned} A_1 &= \sum_{i=1}^n (b_i - a_i)(e_i - d_i), \\ B_1 &= \sum_{i=1}^n (a_i(e_i - d_i) + d_i(b_i - a_i)), \\ A_r &= \sum_{i=1}^n (c_i - b_i)(f_i - e_i), \\ B_r &= \sum_{i=1}^n (c_i(f_i - e_i) + f_i(c_i - b_i)), \\ C_1 &= \sum_{i=1}^n (a_i d_i), \quad C_r = \sum_{i=1}^n (c_i f_i), \\ C_{r+1} &= \sum_{i=1}^n (b_i e_i), \end{aligned}$$

اثبات: با توجه به توابع عضویت پارامترهای فازی $\tilde{\beta}_i$ و داده‌های فازی مثلثی \tilde{X}_i ، مجموعه‌های برش α - این دو عدد فازی به ترتیب عبارت‌اند از: $\tilde{X}_i^\alpha = [d_i + \alpha(e_i - d_i), a_i + \alpha(b_i - a_i)]$ و $\tilde{\beta}_i^\alpha = [a_i + \alpha(b_i - a_i), c_i + \alpha(b_i - c_i)]$ با در نظر گرفتن اصل گسترش حاصل ضرب دو بازه‌ی برش α ، یک بازه‌ی برش α - خواهد بود، لذا خواهیم داشت:

$$\tilde{Z}_i^\alpha = \tilde{\beta}_i^\alpha \cdot \tilde{X}_i^\alpha = [a_i + \alpha(b_i - a_i)] \times [d_i + \alpha(e_i - d_i)], \\ [c_i + \alpha(b_i - c_i)] \times [f_i + \alpha(e_i - f_i)]$$

حال با توجه به رابطه‌ی فوق تابع عضویت \tilde{Z}_i^α به آسانی قابل محاسبه است. $^{[۲۳]}$

۱. تابع عضویت $\tilde{Y} = f(\tilde{\beta}_1 \cdot x_1 + \tilde{\beta}_2 \cdot x_2 + \dots + \tilde{\beta}_n \cdot x_n)$

با فرض پارامترهای فازی $\tilde{\beta}_i = (a_i, b_i, c_i)$ با تابع عضویتی مشابه رابطه‌ی ۱۵ تابع عضویت $\tilde{Y} = f(\tilde{\beta}_1 \cdot x_1 + \tilde{\beta}_2 \cdot x_2 + \dots + \tilde{\beta}_n \cdot x_n)$ در حالت کلی مطابق زیر خواهد بود.

$$\mu_{\tilde{Y}}(y) = \begin{cases} \frac{1}{f(\sum_{i=1}^n b_i \cdot x_i) - f(\sum_{i=1}^n a_i \cdot x_i)} \left(y - f(\sum_{i=1}^n a_i \cdot x_i) \right) & \text{if } -f(\sum_{i=1}^n a_i \cdot x_i) \leq y \leq f(\sum_{i=1}^n b_i \cdot x_i), \\ \frac{1}{f(\sum_{i=1}^n b_i \cdot x_i) - f(\sum_{i=1}^n c_i \cdot x_i)} \left(y - f(\sum_{i=1}^n a_i \cdot x_i) \right) & \text{if } -f(\sum_{i=1}^n b_i \cdot x_i) \leq y \leq f(\sum_{i=1}^n c_i \cdot x_i), \\ 0 & \text{otherwise} \end{cases} \quad (۲۰)$$

اثبات: با استفاده از اصل گسترش مجموعه‌های فازی و تعریف اعداد فازی مثلثی، حاصل ضرب یک عدد فازی مثلثی در یک عدد حقیقی مثبت، مجموع چند عدد فازی مثلثی و $f(\cdot)$ یک عدد فازی مثلثی، یک عدد فازی مثلثی خواهد بود. به عنوان مثال اگر $\tilde{A} = (a, b, c)$ یک عدد فازی مثلثی و x یک عدد حقیقی مثبت باشد، آنگاه $\tilde{A} \cdot x = (ax, bx, cx)$ خواهد بود. همچنین اگر $\tilde{A} = (a, b, c)$ و $\tilde{B} = (d, e, f)$ دو عدد فازی مثلثی باشند آنگاه $\tilde{A} + \tilde{B} = (a+d, b+e, c+f)$ خواهد بود و به همین ترتیب برای تابع حقیقی و یک به یک f خواهیم داشت: $f(\tilde{A}) = (f(a), f(b), f(c))$. بنابراین، با توجه به موارد یادشده و ترکیب آنها خواهیم داشت: $\tilde{Y} = (f(\sum_{i=1}^n a_i), f(\sum_{i=1}^n b_i), f(\sum_{i=1}^n c_i))$ که تابع عضویت \tilde{Y} یک عدد مثلثی مطابق فوق خواهد بود.

۲. تابع عضویت $\tilde{Y} = \tilde{\beta}_1 \cdot \tilde{x}_1 + \tilde{\beta}_2 \cdot \tilde{x}_2 + \dots + \tilde{\beta}_n \cdot \tilde{x}_n$

با فرض پارامترهای فازی $\tilde{\beta}_i = (a_i, b_i, c_i)$ و داده‌های فازی مثلثی $\tilde{X}_i = (d_i, e_i, f_i)$ با تابع عضویتی مشابه رابطه‌ی ۱۵ تابع عضویت

IMPROVEMENT OF ARTIFICIAL NEURAL NETWORK PERFORMANCE USING FUZZY LOGIC FOR EXCHANGE RATE FORECASTING

M. Khashei(corresponding author)

khashei@in.iut.ac.ir

M. Bijari

bijari@cc.iut.ac.ir

F. Mokhatab Rafiei

farimah@cc.iut.ac.ir

**Dept. of Industrial Engineering
Isfahan University of Technology**

Sharif Industrial Engineering and Management Journal
Volume 29, Issue 2, Page 63-71, Original Article

© Sharif University of Technology

- Received 3 April 2011; received in revised form 8 August 2011; accepted 17 April 2012.

Abstract

Time series forecasting is an active research area that has drawn considerable attention for applications in a variety of areas. With the time series approach to forecasting, historical observations of the same variable are analyzed to develop a model describing the underlying relationship. Then, the established model is used in order to extrapolate the time series into the future. Improving forecasting, especially accurate time series forecasting, is an important yet often difficult task facing decision makers in many areas. Computational intelligence approaches, such as artificial neural networks (ANNs) and fuzzy logic, have gradually established themselves as popular tools for forecasting complicated financial markets. Fuzzy is one of the most important soft computing tools, which can provide a powerful framework in order to cope with vague or ambiguous problems, and can express linguistic values and human subjective judgments of natural language.

Artificial neural networks are flexible computing frameworks and universal approximators that can be applied to a wide range of forecasting problems with a high de-

gree of accuracy. The major advantage of neural networks is their flexible nonlinear modeling capability. With ANNs, there is no need to specify a particular model form. Rather, the model is adaptively formed based on the features presented in the data. This data-driven approach is suitable for many empirical data sets, where no theoretical guidance is available to suggest an appropriate data generating process. Despite the advantages cited for them, ANNs have weaknesses, one of the most important of which is their requirement of large amounts of data in order to yield accurate results. Both theoretical and empirical findings have indicated that integration of different models can be an effective way of improving upon their predictive performance and also overcoming the limitations of single models, especially when the models in combination are quite different.

In this paper, a new hybrid model of artificial neural networks is proposed based on the basic concepts of fuzzy logic, in order to overcome the data restriction of neural networks and yield more accurate results than traditional ANNs in situations of short time spans. In the proposed model, instead of using crisp parameters in each layer, fuzzy parameters in the form of triangular fuzzy numbers are applied for related parameters of these layers. In this way, the proposed model can search the feasible spaces easily and more efficiently for finding the optimum values of parameters. The empirical results of exchange rate forecasting indicate that the hybrid model is more satisfactory than its components, i.e, artificial neural networks and fuzzy regression models.

Key Words: Artificial neural networks (anns), fuzzy regression, time series forecasting, combined forecasts, exchange rate.