

پیش‌بینی بازده بازار سهام تهران با استفاده از ترکیب تجزیه موجک و شبکه عصبی فازی تطبیقی

علی رئوفی^۱

تیمور محمدی^۲

تاریخ پذیرش: ۱۳۹۷/۴/۵

تاریخ ارسال: ۱۳۹۶/۶/۲۸

چکیده

همواره مدل‌سازی و پیش‌بینی متغیرهای مالی یکی از موضوع‌های مورد علاقه و مهم برای اقتصاددانان بوده است. در این مقاله، ساختاری برای پیش‌بینی سری‌های زمانی ارایه شده است که با استفاده از رویکرد محاسبات نرم این امکان را فراهم می‌آورد تا بتوان با دقت بیشتر مقادیر آینده یک سری زمانی را پیش‌بینی کرد. در این روش، با استفاده از تجزیه موجک، نویزهای تصادفی داده‌های ورودی شبکه عصبی فازی تطبیقی کاهش می‌یابد و از این رو، این عمل باعث کاهش خطا و بهبود در پیش‌بینی سری زمانی آشوبی مورد نظر می‌شود. در این مقاله، روش یادشده با استفاده از سری بازده بورس اوراق بهادار تهران در بازه زمانی ۱۳۹۰/۱/۸ تا ۱۳۹۵/۰۷/۱ مورد ارزیابی قرار گرفته که نتایج بیان‌کننده برتری روش پیشنهادی نسبت به سایر روش‌هاست. همچنین معناداری اختلاف در پیش‌بینی مدل‌های مختلف با استفاده از آزمون MGN مورد بررسی قرار گرفت که نتایج نشان‌دهنده اختلاف معنادار در پیش‌بینی مدل‌های مختلف بود.

واژگان کلیدی: تبدیل موجک، شبکه عصبی فازی تطبیقی، محاسبات نرم، نویززدایی، بورس اوراق بهادار.

طبقه‌بندی JEL: C45, C55, C58, G17.

۱- دانشجوی دکتری اقتصاد مالی، دانشگاه علامه طباطبائی (نویسنده مسئول)، پست الکترونیکی:

raoofi931@atu.ac.ir

۲- دانشیار دانشکده اقتصاد دانشگاه علامه طباطبائی، پست الکترونیکی: atmahmadi@gmail.com

۱- مقدمه

فرآیند سری‌های زمانی را می‌توان به سه طبقه خطی، تصادفی و آشوبناک دسته‌بندی کرد و براساس این، پیش‌بینی فرآیند خطی ممکن، فرآیند تصادفی غیرممکن و فرآیند آشوبی^۱ تا حدودی ممکن است. به عبارتی، تفاوت سری تصادفی و آشوبی در این است که سیستم مولد فرآیند آشوبی معین (غیرتصادفی) و غیرخطی است و داده‌ها تنها ظاهری نامنظم و تصادفی دارند، در حالی که تابع یک جریان معین یا یک فرمول ریاضی مشخص هستند، اما ناظری که از ساختار و عملکرد سیستم مولد سیگنال آشوب گونه بی-اطلاع است یا اطلاع عمیقی از آن ندارد، نمی‌تواند این سیگنال را از سیگنال واقعاً تصادفی تشخیص دهد؛ بنابراین، با توجه به ناشناخته بودن سیستم مولد داده‌ها، مقادیر آتی به نحوی محاسبه و پیش‌بینی می‌شود که دقت آن به‌طور دایم به صورت نمایی در حال کاهش است (رضایی، ۱۳۹۰). همچنین سیستم‌های آشوبی را می‌توان سیستم‌های دینامیکی غیرخطی دانست که اول، حساسیت بسیاری به وضعیت نخست دارند، دوم، جذب‌کننده‌های عجیب^۲ و پیچیده‌ای دارند و سوم، شکستگی‌های ناگهانی ساختاری در مسیر زمانی آنها به‌خوبی رؤیت می‌شود (پروخورف^۳، ۲۰۰۸).

با توجه به آنچه بیان شد، اگر بتوان فرآیند مولد داده‌های یک متغیر (خطی یا غیرخطی) را به‌دست آورد، پیش‌بینی آن متغیر راحت‌تر و با خطای کمتری امکان‌پذیر خواهد بود (مشیری و مروت، ۱۳۸۵). بررسی‌ها در بازار سرمایه نشان داده است که رفتار سهام از یک الگوی غیرخطی تبعیت می‌کند (رئوفی (۱۳۹۲) و زراءنژاد و رئوفی (۱۳۹۴)) و چهارچوب‌های خطی مدل‌های اقتصادسنجی مانند خودبازگشتی میانگین متحرک انباشته^۴ (ARIMA) و خودبازگشتی برداری^۵ (VAR) می‌توانند تنها بخشی از رفتار سهام در بازار را مدل‌سازی کنند (زراءنژاد و دیگران (۱۳۹۱) و خالوزاده

1- Chaotic Process

2- Strange Attractor

3- Prokhorov

4- Auto Regressive Integrated Moving Average

5- Vector Auto Regressive

(۱۳۸۲)). بنابراین، وجود یک سیستم پویای غیرخطی در رفتار متغیر، مدل‌های خطی را در عمل دچار ابهام خواهد کرد؛ از این رو، پیش‌بینی داده‌هایی که از سیستم‌های غیرخطی پیروی می‌کنند، نیازمند ابزارهای هوشمند و پیشرفته‌تری مانند روش‌های محاسبات نرم^۱ است. این روش‌ها به‌عنوان دسته‌ای از سیستم‌های هوشمند، می‌توانند رابطه غیرخطی بین ورودی و خروجی‌ها را تشخیص دهند (زراءنژاد و رئوفی، ۱۳۹۴). از جمله این روش‌ها می‌توان به روش شبکه‌های عصبی مصنوعی اشاره کرد که توانایی مدل‌سازی پیچیدگی‌های غیرخطی بازارهای مالی را داراست و برتری آن نسبت به روش‌های آماری و اقتصادسنجی در مطالعات متعددی بررسی شده است (رئوفی و دیگران (۲۰۱۶ و ۲۰۱۵)، آرمن و رئوفی (۱۳۹۳)، زراءنژاد و رئوفی (۱۳۹۴)، آتسالاکیس و والاوانیس^۲ (۲۰۰۹) و کاریل و دیگران^۳ (۲۰۰۵)).

در این مقاله، به‌عنوان یک نوآوری نسبت به مطالعات پیشین، از موجک به‌عنوان یک پیش‌پردازنده برای فیلتر کردن نویزهای تصادفی، قبل از مدل‌سازی و پیش‌بینی سری زمانی استفاده شده است. تجزیه موجک یک‌سری را به دو بخش روند و اجزا تقسیم می‌کند. این عمل امکان مدل‌سازی هر یک از دو بخش را به صورت جداگانه فراهم می‌آورد (لینش و جان^۴، ۲۰۱۰). از دیگر کاربردهای مهم تجزیه موجک حذف نویز^۵ از سیگنال مغشوش است. این موضوع باعث تسریع در هم‌گرا شدن شبکه‌های عصبی و سایر الگوریتم‌های تقریب توابع خواهد شد (جائو و دیگران^۶، ۲۰۰۴). تجزیه موجک در سال‌های اخیر، قدرت عملکرد و دقت بالای خود را به‌عنوان یک پیش‌پردازنده در بهبود مدل‌سازی و آنالیز سری‌های زمانی نشان داده است، به نحوی که با فیلتر کردن داده‌ها ابعاد و پیچیدگی مسئله کاهش می‌یابد.

در این مقاله، با طراحی یک الگوریتم ابتکاری برای بهبود در نتایج پیش‌بینی بازده

-
- 1- Soft Computing
 - 2- Atsalakis and Valavanis
 - 3- Karyl et al.
 - 4- Lineesh and John
 - 5- De- noising
 - 6- Gao et al.

شاخص کل بورس اوراق بهادار تهران، نتایج مدل شبکه عصبی فازی تطبیقی (ANFIS) با مدل شبکه عصبی فازی تطبیقی با یک لایه پیش‌پردازنده موجک برای سری زمانی بازده بورس اوراق بهادار در بازه زمانی ۱۳۹۰/۱/۸ تا ۱۳۹۵/۷/۱ مقایسه شده است.

۲- مبانی نظری

در یک تعریف کلی، گمانه‌زنی در مورد شرایط و حوادث آینده را پیش‌بینی^۱ و چگونگی انجام این عمل را پیش‌بینی کردن^۲ می‌نامند. در تجزیه و تحلیل‌های کمی، غالباً پیش‌بینی به صورت استفاده از اطلاعات حال و گذشته در قالب الگوهای تک‌متغیره یا چندمتغیره و به کار بردن الگوی موردنظر برای دوره‌های بعدی است. از این رو، می‌توان گفت، پیش‌بینی عبارت از برآورد احتمالی وقایع آینده است که براساس اطلاعات حال و گذشته انجام می‌شود.

روش‌های پیش‌بینی به‌طور کلی به دو گروه اصلی روش‌های کمی و روش‌های کیفی تقسیم می‌شوند. روش‌های کمی زمانی به کار گرفته می‌شوند که داده‌های مربوط به گذشته در دسترس باشد و انتظار می‌رود، الگوی داده‌ها در آینده نیز ادامه داشته باشد. از روش‌های کیفی زمانی استفاده می‌شود که برای تصمیم‌گیری، اطلاعات کامل و دقیق از آینده در دست نیست یا محیط به گونه‌ای آشفته است که نمی‌توان اطلاعات گذشته را ملاک پیش‌بینی آینده و تصمیم‌گیری قرار داد (رئوفی، ۱۳۹۲).



شکل ۱- انواع مدل‌های پیش‌بینی

مدل‌های کمی به دو دسته تقسیم می‌شوند؛ مدل‌های تک‌متغیره^۱ و مدل‌های علی^۲ یا چندمتغیره^۳. در مدل‌های تک‌متغیره تنها براساس الگوهای تاریخی متغیر موردنظر، ارزش آینده آن پیش‌بینی می‌شود. در مدل‌های علی، یک متغیر وابسته و یک یا چند متغیر مستقل وجود دارد. در این مدل‌ها با استفاده از روابط بین متغیر موردنظر و سایر متغیرها، ارزش آتی متغیر موردنظر پیش‌بینی می‌شود. از جمله مهم‌ترین مدل‌های تک‌متغیره می‌توان به: مدل‌های خودبازگشت میانگین متحرک^۴، هموارسازی نمایی^۵، هلت- وینترز^۶ و مدل

- 1- Univariate Model
- 2- Causal Model
- 3- Multivariate Model
- 4- Auto Regressive Moving Average Models
- 5- Exponential Smoothing Models
- 6- Holt-Winters Model

باکس - جنکینز^۱ اشاره کرد. از جمله مدل‌های چندمتغیره نیز می‌توان به مدل‌های اقتصادسنجی و روش‌های محاسبات نرم که امروزه بسیار مورد توجه قرار گرفته‌اند، اشاره کرد.

کاربرد محاسبات نرم در اقتصاد از اواخر دهه ۸۰ میلادی با مطالعه وایت^۲ (۱۹۸۸)، در بازار مالی و پیش‌بینی قیمت سهام شرکت IBM آغاز شد. البته، هدف اصلی آن مطالعه آزمون فرضیه کارایی بازار بود. هرچند نتایج مطالعه نشان داد که الگوریتم‌های حداقل‌سازی استفاده شده در اقتصادسنجی بهتر از الگوریتم‌های شبکه عصبی است، اما به دلیل ساده بودن شبکه مورد استفاده، نتایج این مطالعه توسط محققان مختلف به مجادله گرفته شد. پس از وایت، مطالعات متعددی در زمینه کاربرد محاسبات نرم با موفقیت صورت پذیرفت. این موفقیت در حوزه اقتصاد مالی، توجه متخصصان اقتصاد کلان و اقتصادسنجی را به خود جلب کرد. با این حال، این روش، به‌رغم کاربردهای بسیار مفید و مثبت علمی، از برخی لحاظ نیز مورد نقد قرار گرفته است. از دید تعدادی از آماردانان و اقتصاددانان نکته قوت شبکه‌های عصبی از لحاظ آزادی از قید و بندهای مفروض در مدل‌های آماری، خود یک نکته ضعف به شمار می‌رود. با این حال، امروزه، در بیشتر مطالعاتی که به پیش‌بینی متغیرهای مالی با رفتارهای غیرخطی^۳ می‌پردازند، روش‌های محاسبات نرم جایگاه ویژه‌ای کسب کرده‌اند. به‌طور خاص با توجه به اینکه عوامل بسیاری از جمله: رویدادهای سیاسی، شرایط اقتصادی، انتظارات معامله‌گران و سایر عوامل محیطی بر شاخص سهام اثرگذارند، طبیعت آن به‌طور کلی بسیار پرنوسان، پویا، غیرخطی، پیچیده و آشوبناک است (یودانگ و لینان، ۲۰۰۹). بنابراین، وجود چنین سیستم پویای غیرخطی در رفتار متغیر، مدل‌های خطی را در عمل دچار ابهام خواهد کرد و در نتیجه، این مدل‌ها از کارایی لازم برخوردار نخواهند بود (عباسی و ابوییک^۴ (۲۰۰۸)، مشیری و مروت (۱۳۸۵)

1- Box- Jenkins Model

2- White

۳- بررسی غیرخطی بودن ماهیت داده‌ها در عملکرد مدل‌های پیش‌بینی از اهمیت ویژه‌ای برخوردار است. برای این منظور مراجعه کنید به: آرمن و رئوفی، ۱۳۹۳.

4- Abbasi & Abouec

و رثوفی (۱۳۹۲)). همچنین با توجه به تحقیقات قبلی برای پیش‌بینی در افق‌های بلندمدت قیمت سهام با استفاده از مدل‌های خطی، ناکارآیی مدل‌های خطی، در فرآیند پیش‌بینی بلندمدت محرز شده است، در حالی که روش‌های محاسبات نرم عملکرد بهتری در افق‌های بلندمدت داشته‌اند (خالوزاده، ۱۳۸۲).

براساس توضیحات پیشین برتری روش‌های محاسبات نرم بر سایر روش‌های آماری و اقتصادسنجی در پیش‌بینی متغیرهایی که از رفتار غیرخطی تبعیت می‌کنند (عباسی و ابوییک (۲۰۰۸)، کاریل و دیگران (۲۰۰۵)، رثوفی (۱۳۹۲) و خالوزاده (۱۳۸۲)) محرز شده است. بنابراین، در این مقاله، به‌جای مقایسه نتایج مدل‌های خطی و غیرخطی بر چگونگی بهبود عملکرد نتایج مدل‌های غیرخطی تک‌متغیره تمرکز شده است.

۲-۱- تجزیه موجک^۱

گسترش کاربرد موجک در دهه‌های اخیر موجب استفاده از آن در علوم مختلف مانند مهندسی برق، آمار، رایانه، فیزیک، اقتصاد و علوم مالی شده است. از این تبدیل به‌عنوان ابزاری علمی برای روشن ساختن ساختارهای پیچیده‌ای که در تلاطم ظاهر می‌شوند و سیگنال‌های گذرایی که به‌سرعت تغییر می‌کنند، مانند جریان‌ات جوی و متغیرهای مالی، استفاده می‌شود. این تبدیل به‌عنوان یک ابزار عددی می‌تواند تا حد زیادی از پیچیدگی محاسبات با مقیاس بزرگ بکاهد، بدین ترتیب که با تغییر هموار ضرایب، ماتریس‌های متراکم را به شکل تُنکی که به‌سرعت قابل محاسبه باشند، درمی‌آورد (صادقی و ذوالفقاری، ۱۳۸۹).

موجک را می‌توان با لنزهای یک دوربین مقایسه کرد؛ لنزهای دوربین این امکان را ایجاد می‌کند که هم تصاویر وسیع را ثبت کرد و هم بر جزئیات بسیار ریز که در حالت عادی از چشم انسان پنهان است، متمرکز شد (مشیری و دیگران، ۱۳۸۹). به عبارتی، موجک‌ها می‌توانند داده‌ها را به اجزایی با فرکانس متفاوت برای تجزیه و تحلیل جداگانه تقسیم کنند؛ به‌طور مثال، در تحلیل سری‌های زمانی مالی، موجک این امکان را فراهم

می‌آورد که سری موردنظر را به دو بخش روند (سیگنال با فرکانس پایین) و اجزا (سیگنال با فرکانس بالا) تفکیک و بتوان هر کدام را به صورت مجزا در مدل‌سازی‌ها استفاده کرد. تبدیل موجک با استفاده از توابع پایه‌ای^۱، سری زمانی را به فضای فرکانس می‌برد و سپس، سری یادشده را در زمان و مقیاس‌های مختلف نشان می‌دهد. بین نظریه موجک و تبدیل فوریه^۲ ارتباط تنگاتنگی وجود دارد، اما تفاوت‌های بین آنها باعث برتری موجک شده است. در تبدیل فوریه یگانه تابع پایه‌ای آن را تابع سینوسی تشکیل می‌دهد. توابع سینوسی از جمله توابع دوره‌ای هستند و ذاتاً غیرموضعی عمل می‌کنند، بنابراین، هرگونه تغییری در یک نقطه خاص از دامنه زمان، تأثیرش روی کل سری احساس می‌شود. به عبارتی، در این حالت فرض می‌شود که ظرفیت فرکانس تابع پایه در طول محور زمان ثابت است، اما توابع پایه‌ای مورد استفاده در موجک دارای انرژی محدود هستند و به صورت موضعی عمل می‌کنند.

فرم ریاضی تبدیل فوریه را می‌توان به صورت زیر نشان داد:

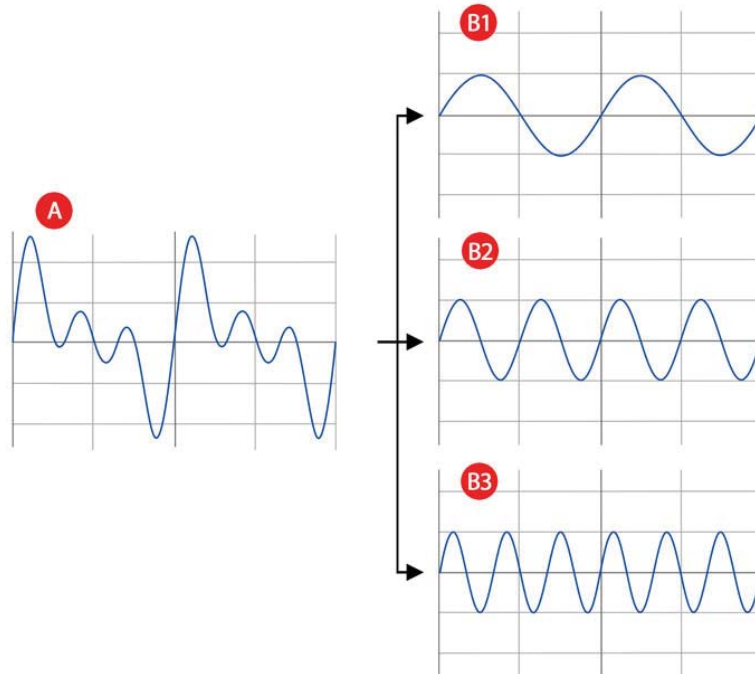
$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \quad (1)$$

که در حقیقت، برابر با مجموع حاصل ضرب سیگنال اصلی $f(t)$ در یک تابع نمایی مختلط است (این تابع نمایی مختلط دارای قسمت حقیقی و موهومی از جنس توابع سینوسی است).

معادله یادشده را می‌توان به صورت زیر نوشت:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) (\cos\omega t - j\sin\omega t) dt \quad (2)$$

نتیجه تبدیل فوریه ضرایب فوریه هستند که در صورت ضرب کردن در توابع سینوسی با فرکانس مربوط می‌توان اجزای سینوسی تشکیل دهنده موج اصلی را به وجود آورد (شکل شماره ۲).



شکل ۲- تجزیه یک سیگنال توسط تبدیل فوریه

ماخذ: رتوفی، ۱۳۹۲.

همان‌گونه که تبدیل فوریه یک شکل موج را به مجموعه‌ای از سیگنال‌های سینوسی تبدیل می‌کند، تبدیل موجک پیوسته نیز عملکرد تقریباً مشابهی دارد. تبدیل موجک پیوسته را می‌توان به صورت مجموع حاصل ضرب سیگنال اصلی در موجک‌های تغییر مقیاس یافته و انتقال داده شده در زمان نشان داد.

$$\gamma(s, \tau) = \int_{-\infty}^{+\infty} f(t) \psi_{s, \tau}(t) dt \quad (3)$$

در رابطه بالا $\psi_{s, \tau}(t)$ موجک مادر تغییر مقیاس یافته به اندازه s و انتقال یافته در زمان به اندازه τ است. نتیجه تبدیل موجک پیوسته ضرایب موجک γ است که تابعی از مقیاس و ضریب جابه‌جایی است؛ بنابراین، با استفاده از تبدیل موجک می‌توان یک سیگنال را با استفاده از مجموعه‌ای از توابع که از تغییر مقیاس و انتقال دادن یک تابع اصلی به نام

موجک مادر به دست می‌آیند به زیرسیگنال‌ها تجزیه کرد.

برخلاف تبدیل فوریه در تبدیل موجک توابع پایه‌ای (موجک مادر) متفاوتی وجود دارد که بسته به نوع کاربرد و ماهیت سری می‌توان از هر کدام از آنها استفاده کرد. تمام توابع پایه‌ای مورد استفاده در تبدیل موجک انرژی محدودی دارند و مقدار متوسط آنها برابر صفر است (صادقی و ذوالفقاری، ۱۳۸۹)؛ بنابراین، به صورت موضعی عمل می‌کنند. این خاصیت موجک سبب شده است که این تبدیل بتواند سری‌های زودگذر و نامانا را بررسی کند (عباسی‌نژاد و نادری، ۱۳۹۱). یکی دیگر از مزیت‌های تبدیل موجک نسبت به تبدیل فوریه، الگوریتم سریع‌تر آن برای ارزیابی انبساط سری است (بجستانی و همکاران، ۲۰۰۶).

در کل، تبدیل موجک دارای ویژگی‌های زیر است:

۱- ایجاد و ارایه یک مسیر برای آنالیز تحلیل شکل موج‌های مختلف در دو حوزه زمان و فرکانس. برخلاف تبدیل فوریه که تنها به تحلیل در حوزه فرکانس می‌پردازد (تان^۱، ۲۰۰۹).

۲- در تبدیل موجک، دامنه دید قابل تغییر است و می‌توان هم‌زمان، توابع پایه‌ای کوتاه‌مدت و بلندمدت را داشت. این در حالی است که در تبدیل فوریه، محتوای فرکانس تابع پایه در طول محور زمانی مانا فرض می‌شود (عباسی‌نژاد و محمدی، ۱۳۸۴).

۳- در تبدیل موجک توابع پایه‌ای بی‌شماری وجود دارد که بسته به نوع کار می‌توان از آن استفاده کرد، در حالی که در تبدیل فوریه تنها تابع پایه‌ای تابع سینوسی است (عباسی‌نژاد و نادری، ۱۳۹۱).

۴- توانایی تقریب سیگنال‌ها با کیفیت بسیار مناسب (صادقی و ذوالفقاری، ۱۳۸۹).
تبدیل موجک را براساس نوع داده‌ها می‌توان به دو دسته تبدیل موجک پیوسته^۲ (CWT) و تبدیل موجک گسسته^۳ (DWT)، تقسیم کرد (کریم و همکاران، ۲۰۱۱).

1- Tan

2- Continuous Wavelet Transform

3- Discrete Wavelet Transform

در صورت استفاده از تبدیل موجک پیوسته حجم اطلاعات و محاسبات بسیار زیاد می‌شود، زیرا S به صورت پیوسته تغییر می‌کند، در ضمن، در برخی موارد ممکن است به یک تابع غیرتحلیلی برخورد کنیم که امکان انتگرال‌گیری از آن وجود نداشته باشد یا اینکه انتگرال جواب خاصی نداشته باشد، از این‌رو، بهتر است از تبدیل موجک گسسته استفاده شود. این تبدیل (تبدیل موجک گسسته) ویژگی‌های بسیار مناسبی را ارائه می‌دهد که از جمله این ویژگی‌ها می‌توان به موارد زیر اشاره کرد (صادقی و ذوالفقاری، ۱۳۸۹):

- ۱- به اندازه کافی اطلاعات مورد نیاز برای تحلیل و بررسی موج اصلی ارائه می‌دهد.
 - ۲- مقدار محاسبات مورد نیاز را به مقدار بسیار مناسبی (نسبت به تبدیل موجک پیوسته) کاهش می‌دهد.
 - ۳- می‌توان توسط آن سیگنال اصلی را در فرکانس‌های مختلف و با وضوح‌های کاملاً متفاوت بررسی و تحلیل کرد.
 - ۴- می‌توان سیگنال اصلی را به سیگنال‌هایی با تقریب نه‌چندان خوب (نسبت به تبدیل موجک پیوسته)، اما با اطلاعات بسیار خوب و مفید تجزیه کرد.
- در تبدیل موجک گسسته متغیرهای S و T که در حالت پیوسته، مقادیر پیوسته‌ای داشتند، در حالت گسسته مقادیر گسسته‌ای به خود می‌گیرند و پیاده‌سازی تبدیل به صورت ساده‌تری انجام می‌گیرد.
- معمول‌ترین حالت برای پذیرفتن مقدار متوسط پارامترهای S و T به صورت توان‌هایی با پایه ۲ بوده که در ادامه، نشان داده شده است:

$$s = 2^j \cdot \tau = n2^j \quad (4)$$

حال، با توجه به این توضیحات می‌توان تبدیل موجک گسسته را به صورت زیر نشان داد:

$$\begin{aligned} d_{j,n} &= Wf(2n^j \cdot n^j) = \langle f(t, \psi_{j,n}(t)) \rangle \\ &= 2^{-\frac{j}{2}} \int_{-\infty}^{+\infty} f(t) \cdot \psi(2^{-j}t - n) dt \end{aligned} \quad (5)$$

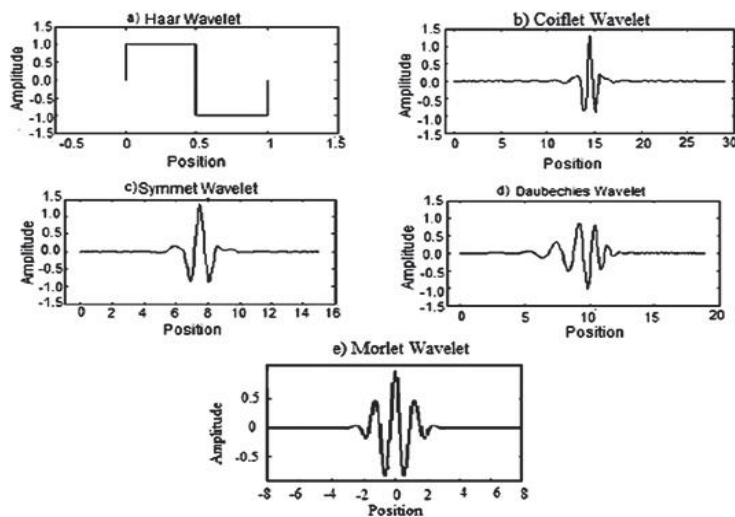
$$\psi_{j,n} = 2^{-\frac{j}{2}} \psi(2^j t - n)$$

در حالت گسسته نیز می‌توان به راحتی تابع اصلی را از تبدیل موجک گسسته بازسازی کرد:

$$f(t) = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} d_{j,n} \psi_{j,n}(t) \quad (6)$$

$$= \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} W f(n2^j, 2^j) (2^{-\frac{j}{2}} \psi(2^{-j}t - n))$$

از جمله مهم‌ترین توابع پایه‌ای گسسته عبارت‌اند از: هار^۱، دایشن^۲، سیملتس^۳، میر^۴ و کویفلت^۵.



شکل ۳- انواع توابع پایه‌ای موجک

مأخذ: رثوفی، ۱۳۹۲.

۲-۱-۱- نويززدایی با استفاده از تبدیل موجک

به‌طور اصولی نويز يك نوسان ناخواسته در سری مورد بررسی است که وجود آن عامل مزاحمی برای تشخیص سیگنال حقیقی و دلخواه است (تان، ۲۰۰۹). نويز اطلاعات موجود

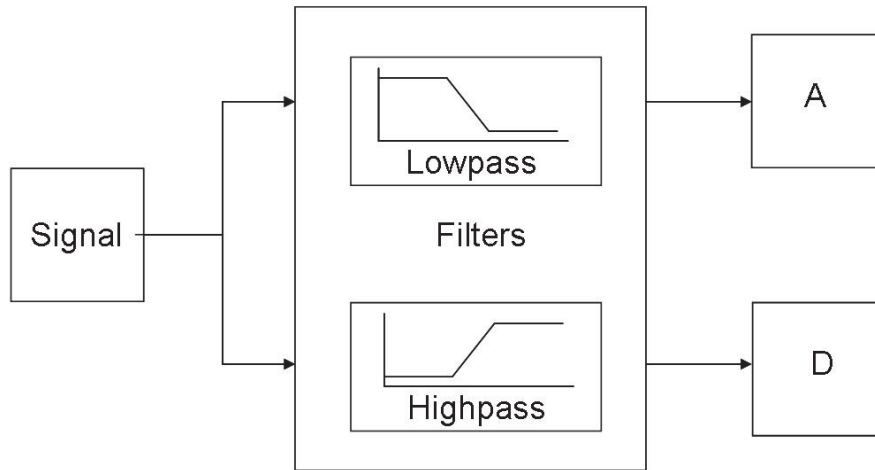
- 1- Haar
- 2- Daubechies
- 3- Symlet
- 4- Meyer
- 5- Coiflet

در یک سیگنال است که می‌توان آن را به وسیله متوسط‌گیری و ضرایب جزئی در تبدیل موجک فیلتر کرد.

از کاربردهای مهم تبدیل موجک، حذف نویز از سیگنال‌های مغشوش شده دریافتی است. حذف نویز و تجزیه سری‌های زمانی اغلب با استفاده از روش‌های اقتصادسنجی و سری‌های فوریه انجام می‌پذیرد. علاوه بر این روش‌ها، فیلترهایی مانند هودریک-پرسکات^۱، باکستر-کینگ^۲ و کریستیانو-فیتزجرالد^۳ برای تجزیه سری‌های زمانی در اقتصاد به کار گرفته می‌شوند. در تجزیه برپایه آنالیز فوریه نویز به‌عنوان یک نوسان با فرکانس بالا در طول سری زمانی (غیرموضعی بودن آنالیز فوریه) فرض می‌شد، اما در روش حذف نویز برپایه تجزیه موجک فرض می‌شود که تجزیه و تحلیل سری‌های زمانی در تفکیک‌های متفاوت و مقیاس‌های مختلف (ویژگی موضعی عمل کردن توابع پایه‌ای) در طول سری ممکن است تفکیک نویز از سری واقعی را بهبود بخشد (صادقی و ذوالفقاری، ۱۳۸۹).

در بیشتر سیگنال‌ها، مؤلفه‌های فرکانس پایین اهمیت بسیاری دارند. این مؤلفه‌ها، مشخصات کلی سیگنال را مشخص می‌کنند. از سوی دیگر، مؤلفه‌های فرکانس بالا، جزئیات ریز سیگنال را بیان می‌کنند؛ برای مثال، صدای انسان را در نظر بگیرید، اگر مؤلفه‌های فرکانس بالا را از صدا حذف کنید در آن صورت، صدای شخص تغییر می‌کند، اما باز هم مشخص است که شخص چه می‌گوید. در صورتی که اگر مقداری از مؤلفه‌های فرکانس پایین را حذف کنیم در آن صورت از صدای شخص چیزی فهمیده نمی‌شود. عملیات نویززدایی به صورت کلی در شکل شماره ۴، نمایش داده شده است.

1- Hodrick Prescott
2- Baxter Kings
3- Christiano Fitzgerald



شکل ۴- تجزیه سیگنال با استفاده از بانک فیلتری

مأخذ: رئوفی، ۱۳۹۲.

در شکل شماره ۴، سیگنال اصلی از دو فیلتر مکمل عبور داده و به دو سیگنال مجزا تجزیه می‌شود.

ساده‌ترین روش برای حذف نویز به کمک تجزیه موجک آستانه‌گذاری است. در این روش، ابتدا سیگنال مغشوش دریافتی در چندین مرتبه تجزیه می‌شود و سپس، روی ضرایب تجزیه در مراتب مختلف آستانه‌گذاری صورت می‌گیرد. این روش که آب‌گرفتنی^۱ ضرایب تجزیه موجک نامیده می‌شود، ابتدا توسط دونهو^۲ و جان استون^۳ ارایه شد (باروس و جوپیناس^۴، ۱۹۹۸). آستانه‌گذاری روی ضرایب به دو روش سخت و نرم صورت می‌گیرد. تفاوت این دو روش در آن است که در روش نرم ضرایب بالاتر از آستانه دست‌کاری می‌شود، در حالی که در روش سخت از خود ضرایب استفاده می‌شود. آستانه‌گذاری همواره با حذف بخشی از اطلاعات روبه‌رو است. از این رو، برای تعیین آستانه مناسب روش‌های متفاوتی مانند آستانه‌گذاری فراگیر، تخمین نااریب یا

- 1- Shrinkage
- 2- Donoho
- 3- John Stone
- 4- Burrus and Gopinath

کوچک‌ترین حداکثر به کار می‌رود (باروس و جویناس، ۱۹۹۸). در روش آستانه‌گذاری فراگیر تمام ضرایب با یک آستانه مجزا برای هر مرتبه تجزیه، تعیین می‌شوند. تفاوت این روش با دو روش دیگر اعوجاج کمتر در سیگنال نویززدایی شده است.

۲-۲- شبکه عصبی فازی تطبیقی ANFIS

در بین روش‌های مدل‌سازی نوین، سیستم‌های فازی جایگاه ویژه‌ای را کسب کرده‌اند. این موضوع را می‌توان معلول توانایی پیاده‌سازی دانش بشری با استفاده از مفاهیم و عبارت‌های زبانی و قواعد فازی، روابط غیرخطی و قابلیت تطبیق‌پذیری این نوع سیستم‌ها دانست. به‌طور خلاصه، یک سیستم فازی یک سیستم مبتنی بر قواعد منطقی «اگر-آنگاه» است. نقطه شروع ساخت یک سیستم فازی تعریف و تنظیم مجموعه‌ای از قواعد اگر-آنگاه فازی توسط فرد خبره در حوزه موردنظر است. به‌دست آوردن این قواعد مهم‌ترین و دشوارترین مرحله کار است و به دانش لازم و توان کافی برای پیاده‌سازی صحیح آن قواعد نیاز دارد.

اگر موفق به پیاده‌سازی یک سیستم فازی با قابلیت یادگیری شویم، در آن صورت، قادر خواهیم بود که خطای مقادیر خروجی را با استفاده از روش حداقل مربعات خطا به کمترین مقدار برسانیم. با ترکیب روش یادگیری یادشده با روش پس‌انتشار خطا می‌توان به یک روش آموزش ترکیبی دست یافت. عملکرد روش آموزش ترکیبی به این صورت است که در هر دور آموزش، هنگام حرکت رو به جلو خروجی‌های گره‌ها^۱ به صورت عادی تا لایه آخر محاسبه می‌شوند و سپس، پارامترهای نتیجه توسط روش کمترین مجموع مربعات خطا محاسبه می‌شوند. در ادامه، پس از محاسبه خطا در بازگشت رو به عقب نسبت خطا روی پارامترهای شرط، پخش و با استفاده از روش شیب نزولی خطا مقدار آنها تصحیح می‌شود.

به‌طور کلی، شبکه‌های عصبی مصنوعی برای اهدافی که ما از آن استفاده می‌کنیم، توانایی بالایی در توسعه یک مدل را در زمانی منطقی ندارد. از سوی دیگر، مدل‌سازی

فازی برای کاربرد ادغام تصمیمات از متغیرهای متفاوت، نیازمند رویکردی برای یادگیری از تجارب (داده‌های جمع‌آوری شده) است. شبکه عصبی مصنوعی و مدل فازی در بسیاری از زمینه‌های کاربردی استفاده شده‌اند و در هر کدام محاسن و معایبی هست؛ بنابراین، ترکیب موفقیت‌آمیز این دو دیدگاه، مدل‌سازی شبکه‌های عصبی مصنوعی و فازی، موضوع مطالعات آینده قرار گرفته است.

یکی از متداول‌ترین روش‌های ترکیبی سیستم استنتاج تطبیقی عصبی فازی (ANFIS) است که توسط ژانگ^۱ در سال ۱۹۹۳ معرفی شد و حاوی مزیت‌های هر دو سیستم فازی و شبکه عصبی است (ژانگ و همکاران (۱۹۹۷) و سرینیواسان و فیشر^۲ (۱۹۹۵)). این مدل یک سامانه فازی^۳ از نوع تاکاگی-سوگنو^۴ را در یک ساختار عصبی اجرا می‌کند و برای فرآیند آموزش از روش پس‌انتشار خطا یا ترکیبی از روش‌های آموزش پس‌انتشار خطا و حداقل مربعات خطا استفاده می‌کند.

در ANFIS قوانین ثابت هستند و عامل‌های توابع عضویت بهینه‌سازی می‌شود که برای تعیین عامل‌های توابع عضویت (همان شکل توابع عضویت) از روند نماهای آموزشی شبکه عصبی استفاده می‌شود. نوع توابع عضویت (مانند مثلثی و گوسی) و تعداد توابع عضویت برای ورودی‌ها و خروجی‌ها به روش سعی و خطا تعیین می‌شود. در ANFIS لازم است که در لایه اول نوع تابع عضویت و تعداد آنها مشخص شود.

ANFIS مدلی مانند f را شناسایی نموده و به کار می‌گیرد که تقریباً بتواند به جای تابع f (تولیدکننده داده‌ها) مورد استفاده قرار گیرد. هر سیستم فازی شامل N قاعده فازی است

1- Jang

2- Srinivasan and Fisher

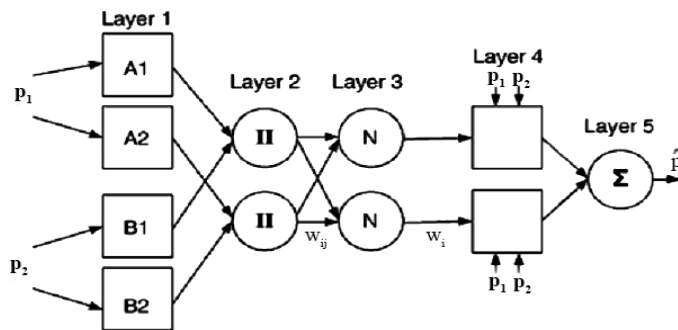
۳- در ادبیات منطق فازی، انواع مختلفی از سامانه‌های فازی پیشنهاد شده و معروف‌ترین آنها سیستم استدلال تاکاگی سوگنو (۱۹۸۵) و سیستم استدلال ممدانی (۱۹۷۷)، است. این دو روش در بسیاری از جنبه‌ها یکسان هستند مانند فازی‌سازی ورودی‌ها و عملگرهای فازی، اما تفاوت اصلی بین این دو روش در شکل خروجی آنهاست. در سیستم ممدانی خروجی به صورت یک مجموعه فازی است که باید غیرفازی (Defuzzify) شوند، اما در سیستم سوگنو خروجی به صورت خطی یا ثابت است (راهنمای نرم‌افزار متلب، ۲۰۱۰).

4- Takagi-Sugeno

که برای پیش‌بینی استفاده و به صورت زیر ارایه می‌شود:

$$f(x) = \frac{\sum_{i=1}^N E_t \left(\prod_{i=1}^n \mu_{A_i(ji)}(E_{t-p_i}) \right)}{\sum_{i=1}^N \left(\prod_{i=1}^n \mu_{A_i(ji)}(E_{t-p_i}) \right)} \quad (7)$$

که در آن، E_{t-p} ، μ_{A_i} و i به ترتیب نشان‌دهنده وقفه‌های متغیر تا مرتبه p ، مرتبه عضویت، مجموعه‌های فازی و قاعده فازی است. همچنین $\mu_{A_i(ji)}(x_i)$ بیان‌کننده مرتبه عضویت از ورودی به مقدار قاعده فازی نام $A_i(ji)$ است (شکل شماره ۵).



شکل ۵- ساختار ANFIS با دو ورودی

مأخذ: رئوفی، ۱۳۹۲.

لایه اول، گره‌های ورودی است. هر گره از این لایه، مقادیر عضویت متعلق به هر یک از مجموعه‌های فازی مناسب را با استفاده از تابع عضویت تولید می‌کند.

$$\begin{aligned} O_{1.i} &= \mu_{A_i}(X) \quad i = 1,2. \\ O_{2.i} &= \mu_{B_{i-2}}(X) \quad i = 3,4. \end{aligned} \quad (8)$$

در این لایه، x و y ورودی‌های غیرفازی به گره i هستند و A_i و B_i متغیرهای زبانی منطبق با این گره هستند.

لایه دوم، لایه‌ای است که در آن هر گره به صورت یک گره ثابت به نام P بوده که خروجی گره‌ها حاصل ضرب تمام سیگنال‌های ورودی است.

$$O_{2.i} = W_i = \mu_{A_i}(x_i) * \mu_{B_i}(y_i) \quad i = 1,2. \quad (9)$$

هر گره خروجی بیان‌کننده شدت برانگیختگی^۱ یک قاعده است. لایه سوم، لایه‌ای است که در آن هر گره به صورت یک گره ثابت به نام N است که N امین گره، نسبت N امین قاعده شدت برانگیختگی را برای تمام قاعده‌های شدت برانگیختگی به صورت زیر محاسبه می‌کند:

$$O_{3.i} = \bar{W}_i = \frac{W_i}{W_1 + W_2} \quad i = 1, 2. \quad (10)$$

لایه چهارم، لایه‌ای است که تابع گره چهارمین لایه توزیع N امین قانون را به کل خروجی محاسبه می‌کند و به صورت زیر تعریف می‌شود:

$$O_{4.i} = \bar{W}_i f_i = \bar{W}_i (p_i x + q_i y + r_i) \quad i = 1, 2. \quad (11)$$

که در آن، \bar{W}_i یک شدت برانگیختگی نرمال شده از لایه سوم و p_i ، q_i و r_i مجموعه پارامترهای این گره‌اند که به پارامترهای استنتاجی موسوم هستند.

لایه پنجم، گره خروجی است. تنها گره این لایه، یک گره ثابت به نام \sum است که تمام خروجی‌ها را صورت حاصل جمع همه سیگنال‌های ورودی به صورت زیر محاسبه می‌کند:

$$O_{5.i} = \sum \bar{W}_i f_i = \frac{\sum w_i f_i}{\sum w_i} \quad i = 1, 2. \quad (12)$$

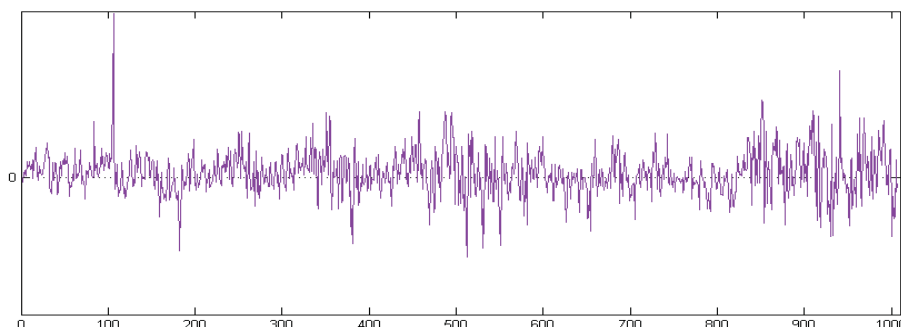
۳- معرفی داده

در این مطالعه، از داده‌های سری زمانی بازده شاخص بورس اوراق بهادار تهران (TEPIX) در بازه زمانی ۱۳۹۰/۱/۸ تا ۱۳۹۵/۷/۱ با حجم نمونه ۱۰۰۷ مشاهده استفاده شده است. بازده روزانه شاخص بورس اوراق بهادار تهران به صورت درصد تغییر در مقدار شاخص قیمت (TEPIX) تعریف و به صورت زیر محاسبه می‌شود:

$$R_t = \ln \left(\frac{P_t}{P_{t-1}} \right) * 100 \quad (13)$$

که در آن، R_t بازده روزانه شاخص قیمت سهام، P_t مقدار کنونی شاخص و P_{t-1} مقدار آن در دوره قبل است.

روند سری زمانی بازده شاخص کل سهام در شکل شماره ۶، مشاهده می‌شود.



شکل ۶- نمودار بازده شاخص کل بورس اوراق بهادار

در این مطالعه، از ۸۵ درصد مشاهدات (۸۵۷ مشاهده) برای مدل‌سازی (داده‌های آموزش) و از ۱۵۰ مشاهده دیگر برای پیش‌بینی و مقایسه دقت پیش‌بینی‌ها (داده‌های آزمون) استفاده شده است. هرچه حجم نمونه برای داده‌های آموزش غنی‌تر (تعداد مشاهدات کافی) باشد، امکان رسیدن به نتایج مطلوب‌تر برای داده‌های آزمون محتمل‌تر است. بدین منظور در بیشتر مطالعات (منتظر حجت و دیگران (۲۰۱۶)، صادقی و ذوالفقاری (۱۳۹۲)، کیانی و همکاران (۱۳۸۵)، رئوفی (۱۳۹۲) و زراءنژاد و دیگران (۲۰۱۲))، بخش عمده‌ای از مشاهدات را به داده‌های آموزش اختصاص داده‌اند.

خلاصه‌ای از آماره‌های توصیفی برای بازده شاخص بورس تهران در جدول شماره ۱، مشاهده می‌شود. براساس جدول، میانگین سری موردنظر برابر ۰/۰۷۲ است. همچنین شواهدی مبنی بر چولگی^۱ مثبت در توزیع بازده وجود دارد که نشان‌دهنده احتمال زیاد برای افزایش‌های بزرگ نسبت به کاهش‌ها در بازده شاخص است. این پدیده نشان می‌دهد که بازده شاخص به احتمال نامتقارن است. همچنین سری موردنظر دارای کشیدگی^۲ پایین است. آماره جارک-^۳ نشان می‌دهد که با احتمال بسیار بالا فرضیه صفر مبنی بر نرمال

1- Skewness
2- Kurtosis
3- Jarque-Bera

بودن توزیع بازده رد می‌شود، بنابراین، بازده شاخص کل از توزیع نرمال تبعیت نمی‌کند.

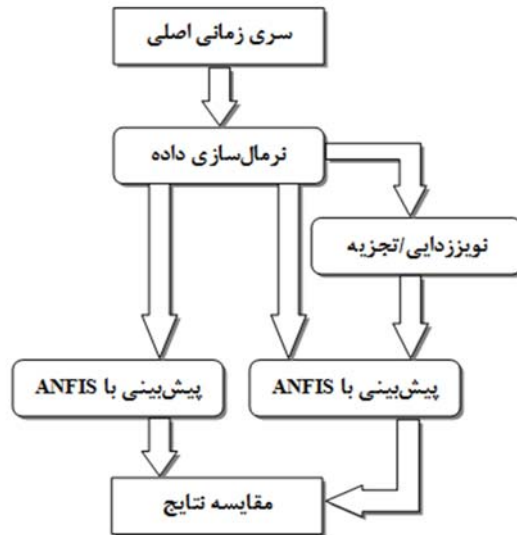
جدول ۱- آماره‌های توصیفی بازده روزانه

تعداد مشاهدات	مجموع	بازگشت - باز	کمیابی	تولگی	انحراف معیار	سکول	حد اکثر	حد کم	میان	واریانس	میان
۱۰۰۷	۷/۹	۵۹/۶	۶/۸	۰/۴۲	۰/۳۱	-۱/۱	۲/۲۸	۰/۰۵۷	۰/۰۷۲	بازده روزانه	

مأخذ: یافته‌های تحقیق.

۴- روش کار و معرفی الگوریتم پیشنهادی

همان‌طور که در شکل شماره ۷، مشاهده می‌شود، در الگوریتم پیشنهادی از دو روش شبکه عصبی فازی با لایه پیش‌پردازش تبدیل موجک و بدون لایه پیش‌پردازنده استفاده شده است.



شکل ۷- شمای الگوریتم پیشنهادی

در هر دو روش، ابتدا داده‌ها به صورت زیر نرمال‌سازی شده‌اند:

$$\forall x_t \in R_t . nR_t = \frac{x_t - \min(R_t)}{\max(R_t) - \min(R_t)} \quad (14)$$

در روش پیشنهادی از یک لایه پیش‌پردازنده تبدیل موجک برای نويززدایی و کاهش ابعاد مسئله استفاده شده است. این لایه پیش‌پردازنده موجب تضعیف نويزهای تصادفی سیگنال می‌شود و این موضوع باعث تسریع در هم‌گرا شدن الگوریتم‌های تقریب توابع و همچنین کاهش خطای این الگوریتم‌ها در پیش‌بینی می‌شود. همچنین شبکه آموزش دیده با استفاده از الگوریتم پیشنهادی این امکان را داراست که برای انتخاب ورودی هم از وقفه‌های خود متغیر استفاده کند هم از وقفه‌های نويز زدایی شده متغیر. با توجه به اینکه یکی از اهداف این مقاله، بررسی کارآیی تبدیل موجک در بالا بردن کارآیی شبکه ANFIS است، از این رو، ساختار شبکه عصبی فازی اعم از شکل و تعداد توابع عضویت و تعداد ورودی‌ها در دو مدل شبکه عصبی فازی و شبکه عصبی فازی همراه با تبدیل موجک یکسان فرض شده است. همچنین پس از پیش‌بینی توسط دو روش یادشده با استفاده از شاخص‌های متداول ارزیابی عملکرد مانند MAE، RMSE، U-thile و MAPE، کارآیی دو روش مقایسه شده است.

معیارهای ارزیابی یادشده از جمله معیارهای کاربردی و مفید در زمینه بررسی قدرت پیش‌بینی روش‌های مختلف هستند، اما هیچ‌کدام از آنها قادر نیستند برتری یک روش را به صورت آماری بررسی کنند؛ بنابراین، برای اینکه از نظر آماری فرضیه برابری صحت پیش‌بینی‌ها در مدل‌های رقیب مورد آزمون قرار گیرند از آزمون مورگان-گرنجر-نیولند (MGN) استفاده شد.

براساس این فرمول، ابتدا خطای پیش‌بینی دو روش مختلف که با $e_{1,t}$ و $e_{2,t}$ نشان داده می‌شود، محاسبه می‌شود. سپس، مجموع (S_t) و تفاوت (D_t) خطای پیش‌بینی براساس روابط زیر محاسبه می‌شود:

$$\begin{aligned} S_t &= e_{1,t} + e_{2,t} \\ D_t &= e_{1,t} - e_{2,t} \\ e_{1,t} &= x_m - x_{1p} \\ e_{2,t} &= x_m - x_{2p} \end{aligned} \quad (15)$$

که در آن، x_{1p} و x_{2p} به ترتیب مقادیر پیش‌بینی شده خروجی توسط مدل اول و دوم (مدل‌های رقیب) و x_m مقادیر اندازه‌گیری شده خروجی هستند. آزمون برابری میانگین مربعات خطای دو روش مختلف، در مقابل کمتر بودن خطای مدل اول را می‌توان با استفاده از آماره MGN مورد بررسی قرار داد.

$$MGN = \frac{\hat{\rho}_{sd}}{\sqrt{\frac{1 - \hat{\rho}_{sd}^2}{N - 1}}} \quad (۱۶)$$

که در آن، $\hat{\rho}_{sd}$ ضریب همبستگی بین s_t ، d_t و N تعداد مشاهدات است. مقدار MGN محاسبه شده با t جدول با درجه آزادی $N-1$ مقایسه می‌شود (دیبولد ماریانو، ۲۰۰۲).

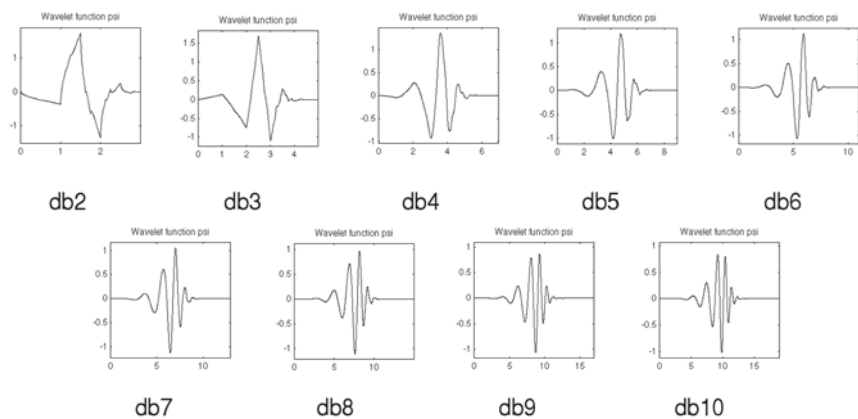
۵- برآورد مدل و نتایج

مطابق الگوریتم پیشنهادی پس از نرمال‌سازی داده‌های بازده بورس اوراق بهادار اقدام به تجزیه آن می‌کنیم. این کار توسط جعبه ابزار موجک نرم‌افزار متلب^۱ صورت می‌گیرد. برای انتخاب سطح بهینه تجزیه داده‌ها، ابتدا داده‌ها به پنج سطح تجزیه و سپس، به کمک جعبه ابزار متلب، سطح بهینه مشخص شد. براساس این، سطح بهینه تجزیه داده‌ها برابر یک است.

شکل شماره ۸، نمودار تجزیه داده‌ها را تا سطح یک با تابع «دایبزش^۲» نشان می‌دهد. علت انتخاب این تابع همخوانی ظاهری آن با نوسانات موجود در بازده شاخص بورس بوده است (محمدی و همکاران، ۱۳۹۶).

1- Matlab

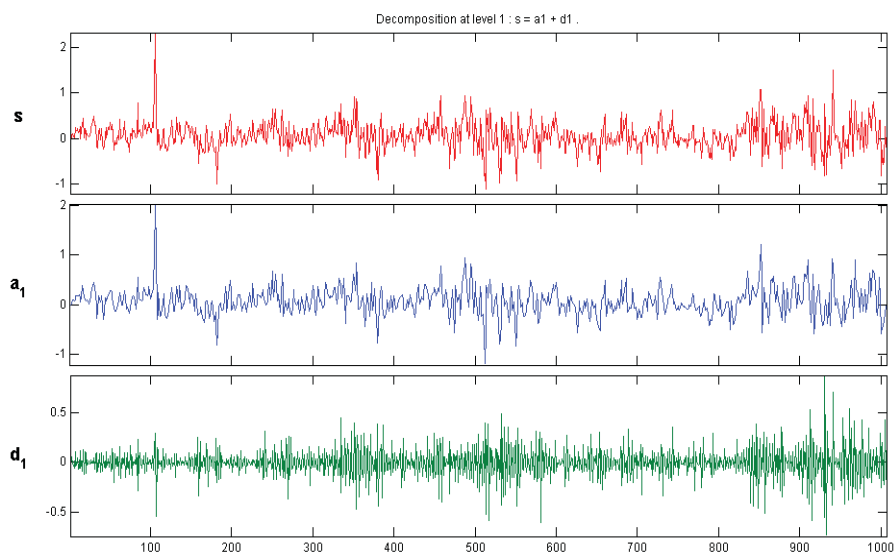
2- db 3



شکل ۸- خانواده موجک دابیشز (db)

مأخذ: رثوفی، ۱۳۹۲.

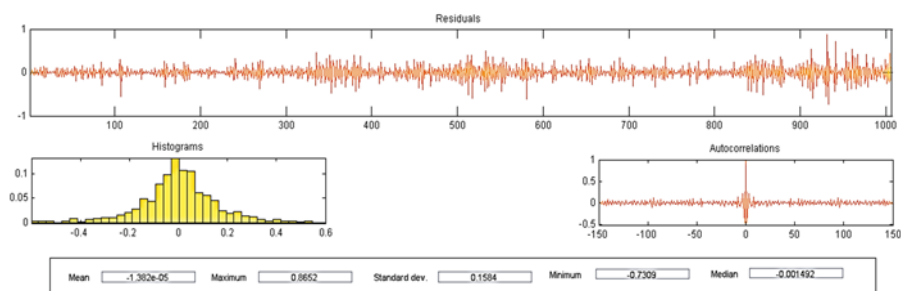
براساس شکل، s همان سیگنال اصلی بازده بورس اوراق بهادار، a_1 سری تقریب زده شده و d_1 جزئیات فیلتر شده هستند.



شکل ۹- سری زمانی بازده بورس و مؤلفه‌های تجزیه شده توسط موجک دابیشز ۳

مأخذ: یافته‌های پژوهش.

با بررسی مقادیر فیلتر شده از داده اصلی (سری d_1) می‌توان به ماهیت داده‌های فیلتر شده پی برد. در شکل شماره ۱۰، مقادیر فیلتر شده، هیستوگرام، خودهمبستگی و همچنین آماره‌های توصیفی مقادیر فیلتر شده مشاهده می‌شود. با توجه به هیستوگرام این داده‌ها می‌توان گفت که توزیع این متغیر تقریباً نرمال است. همچنین با توجه به نمودار خودهمبستگی داده‌ها می‌توان نتیجه گرفت که خودهمبستگی بین داده‌ها بسیار ناچیز و حدود صفر است. آماره‌های توصیفی نیز نشان می‌دهند که میانگین حدوداً صفر است. با توجه به توصیفات بالا مشخص می‌شود که سری یادشده تمام خصوصیات نوفه سفید^۱ را داراست؛ بنابراین، می‌توان نتیجه گرفت که مقادیر فیلتر شده هیچ‌گونه بار اطلاعاتی ندارد و کاملاً تصادفی هستند.

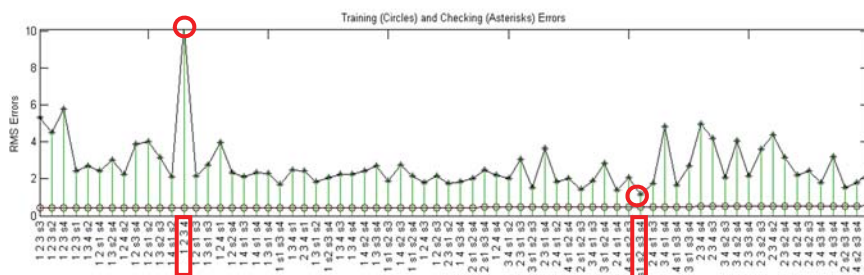


شکل ۱۰- مقادیر فیلتر شده از سری بازده بورس اوراق بهادار

مأخذ: یافته‌های پژوهش.

حال، با توجه به الگوریتم پیشنهادی شبکه عصبی فازی را با داده‌های اصلی و داده‌های فیلتر شده آموزش خواهیم داد. در این مطالعه، از ۸۵ درصد مشاهدات (۸۵۷ مشاهده) برای آموزش شبکه و از ۱۵۰ مشاهده دیگر برای پیش‌بینی و مقایسه دقت پیش‌بینی‌ها استفاده شده است. برای مدل‌سازی و آموزش شبکه از طریق تغییر مداوم توابع عضویت مختلف، تعداد توابع عضویت و از طریق آزمون و خطا، شبکه‌ای با پنج تابع عضویت «گوسی» برای ورودی و توابع خطی برای متغیر خروجی انتخاب شده است. شکل شماره ۱۱، مقدار خطای آموزش شبکه با ۴ ورودی از ترکیبات وقفه‌های داده اصلی (وقفه‌های ۱، ۲، ۳ و ۴)

و چهار وقفه از داده فیلتر شده (وقفه‌های s1، s2، s3 و s4) را نشان می‌دهد. همان‌طور که در این شکل مشاهده می‌شود، کمترین خطا مربوط به زمانی است که از چهار داده فیلتر شده استفاده شده و بیشترین خطا را شبکه آموزش دیده به وسیله داده‌های خام داراست.



شکل ۱۱- مقدار خطای RMSE شبکه عصبی فازی با داده‌های ورودی متفاوت

مأخذ: یافته‌های پژوهش.

برای مقایسه و ارزیابی قدرت پیش‌بینی مدل‌ها از معیارهای ارزیابی کارآیی مدل‌های پیش‌بینی مانند RMSE، MSE، MAE، MAPE و TIC استفاده شده است. همان‌طور که در جدول شماره ۲، مشاهده می‌شود، عملکرد شبکه عصبی فازی با یک لایه پیش‌پردازنده تبدیل مویک به مراتب بهتر از شبکه عصبی فازی تعلیم دیده با داده‌های خام است.

جدول ۲- مقایسه عملکرد دو روش پیشنهادی ANFIS

شرح	MSE	RMSE	MAE	MAPE	U-Thile
با داده‌های اصلی	۰/۰۵۶۱	۰/۲۴۶۷	۰/۱۹۰۲	۲۰۶/۷۷۴	۰/۴۴۷۰
با داده‌های تجزیه شده	۰/۰۴۷۸	۰/۲۱۲۹	۰/۱۶۹۰	۲۰۱/۰۰۹	۰/۳۳۳۳

مأخذ: یافته‌های پژوهش.

همچنین با استفاده از آماره MGN به بررسی معنادار بودن اختلاف در پیش‌بینی‌ها

پرداخته شد.

جدول ۳- نتیجه آزمون MGN

شرح	مقایسه دو مدل پیشنهادی
مقدار آماره MGN	۲/۱۸۰

مأخذ: یافته‌های پژوهش.

با مقایسه آماره به‌دست آمده با جدول توزیع نرمال مشاهده می‌شود که مدل شبکه عصبی فازی با لایه پیش‌پردازنده تبدیل موجک به صورت معناداری نسبت به روش دیگر، پیش‌بینی دقیق‌تری داشته است.

۶- جمع‌بندی و نتیجه‌گیری

در بین بازارهای مالی، بازار بورس اوراق بهادار از اهمیت ویژه‌ای برخوردار است، از این رو، پیش‌بینی آن برای کاهش ریسک در سرمایه‌گذاری و تصمیم‌گیری در بازارهای سرمایه بسیار مهم است. به‌طور معمول، برای پیش‌بینی از روش‌های آماری و اقتصادسنجی سری‌های زمانی استفاده می‌شود، اما این مدل‌ها با ضعف‌هایی همراه هستند که به محقق اجازه نمی‌دهند ساختار پیچیده و غیرخطی متغیرها را شناسایی و سری موردنظر را با خطایی قابل قبول پیش‌بینی کنند. از این رو، پیش‌بینی داده‌هایی که از سیستم‌های غیرخطی پیروی می‌کنند، نیازمند ابزارهای هوشمند و پیشرفته‌تری مانند روش‌های محاسبات نرم است. مقالات متعددی به برتری این روش‌ها بر روش‌های آماری و اقتصادسنجی پرداخته‌اند (رئوفی و دیگران (۲۰۱۵)، عباسی و ابوییک (۲۰۰۸)، کاریل و دیگران (۲۰۰۵) و رئوفی (۱۳۹۲))، از این رو، این مقاله به‌جای مقایسه نتایج مدل‌های خطی و غیرخطی بر بهبود عملکرد نتایج مدل‌های غیرخطی تک‌متغیره متمرکز است. روش‌هایی مانند شبکه‌های عصبی که امروزه از رایج‌ترین روش‌های پیش‌بینی به شمار می‌رود، گاهی با توجه به الگوریتم یادگیری، در کمینه محلی^۱ گیر می‌افتد که این مشکل بر میزان خطا و قابلیت اطمینان در یادگیری الگوی موردنظر اثر می‌گذارد. برای رفع این مشکل از ترکیب ابزارهای مختلف محاسبات نرم، مانند الگوریتم‌های تکاملی (برای مثال، الگوریتم ژنتیک)

1- Local Minima

یا سیستم استنتاج فازی بهره گرفته‌اند. در این مقاله، برای رفع این مشکل از روش شبکه عصبی فازی تطبیقی (ANFIS) که ترکیبی از شبکه عصبی و استنتاج فازی است، استفاده شده، علاوه بر آن، یک لایه پیش‌پردازنده برای بهبود در نتایج این روش به کار گرفته شده است. تجزیه موجک در سال‌های اخیر، قدرت عملکرد و دقت بالای خود را در بهبود مدل‌سازی و آنالیز سری‌های زمانی نشان داده است، به نحوی که با فیلتر کردن داده‌ها ابعاد مسئله را کاهش می‌دهد. این مسئله باعث تسریع در هم‌گرا شدن شبکه‌های عصبی و سایر الگوریتم‌های تقریب توابع خواهد شد. مهم‌ترین عامل برای رسیدن به خطایی قانع‌کننده، هموارسازی سری زمانی است که تبدیل موجک این وظیفه را با حذف نویزهای فرکانس بالا انجام می‌دهد. در این مقاله، از تبدیل موجک برای فیلتر کردن نویزهای تصادفی به همراه شبکه عصبی فازی تطبیقی استفاده و نتایج آن با شبکه عصبی فازی بدون لایه پیش‌پردازنده موجک مقایسه شده است. با استفاده از معیارهای ارزیابی MAE ، $RMSE$ ، $MAPE$ و TIC به مقایسه نتایج این دو مدل پرداخته شد که در تمام این شاخص‌ها مدل شبکه عصبی فازی تطبیقی با لایه پیش‌پردازنده تبدیل موجک از عملکرد بهتری نسبت به مدل رقیب برخوردار بود. همچنین با استفاده از آزمون مورگان-گرنجر-نیوبلد (MGN) به بررسی معنادار بودن اختلاف در قدرت پیش‌بینی مدل‌ها پرداخته شد که نتایج نشان‌دهنده برتری معنادار پیش‌بینی شبکه عصبی فازی تطبیقی با لایه پیش‌پردازنده تبدیل موجک بود. در نهایت، به‌عنوان پیشنهاد برای پژوهش‌های آتی، توصیه می‌شود، نتایج الگوریتم پیشنهادی با استفاده از داده‌های شبیه‌سازی شده با روش مونت کارلو^۱ آزمایش شود تا برتری کلی این الگوریتم در پیش‌بینی متغیرهای مالی بر سایر روش‌ها مشخص شود.

منابع

- آرمن، سیدعزیز و علی رئوفی (۱۳۹۳)، ارزیابی پیش‌بینی‌پذیری قیمت طلا و مقایسه پیش‌بینی روش‌های خطی و غیرخطی، *نظریه‌های کاربردی اقتصاد*، شماره ۳، صص ۱-۲۴.
- رضایی، وحید (۱۳۹۰)، بررسی رفتار آشوبناک در سری زمانی قیمت طلا، *پایان‌نامه کارشناسی ارشد*، دانشگاه شهید چمران اهواز، دانشکده اقتصاد و علوم اجتماعی.
- رئوفی، علی (۱۳۹۲)، شناسایی سیستم مولد داده‌های شاخص بورس اوراق بهادار تهران، *مدل‌سازی و پیش‌بینی آن با استفاده از محاسبات نرم*، پایان‌نامه کارشناسی ارشد، دانشکده اقتصاد و علوم اجتماعی دانشگاه شهید چمران اهواز.
- زراءنژاد، منصور و علی رئوفی (۱۳۹۴)، ارزیابی و مقایسه عملکرد پیش‌بینی روش‌های خطی و غیرخطی برای بازده روزانه بورس اوراق بهادار تهران، *دو فصلنامه اقتصاد پولی - مالی*، شماره ۹، صص ۱-۲۹.
- زراءنژاد، منصور، پویان کیانی، صلاح ابراهیمی و علی رئوفی (۱۳۹۱)، پیش‌بینی قیمت نفت خام اوپک با استفاده از مدل میانگین متحرک خودرگرسیون انباشته فازی، *پژوهشنامه اقتصاد انرژی ایران*، شماره ۵، صص ۱۰۷-۱۲۷.
- صادقی، حسین و مهدی ذوالفقاری (۱۳۸۹)، پیش‌بینی کوتاه‌مدت تقاضای برق کشور با استفاده از شبکه‌های عصبی و تبدیل موجک، *فصلنامه اقتصاد مقداری (بررسی‌های اقتصادی سابق)*، دوره ۷، شماره ۲، صص ۲۷-۵۶.
- عباسی‌نژاد، حسین و شاپور محمدی (۱۳۸۴)، تحلیل سیکل‌های تجاری با استفاده از نظریه موجک‌ها، *مجله تحقیقات اقتصادی*، شماره ۷۵، صص ۱-۲۰.
- کیانی بجستانی، آرمان، یاسر محمدیان روشن، ناصر پریرز و محمدرضا اکبرزاده توتونچی (۱۳۸۵)، پیش‌بینی سری‌های زمانی آشوب‌گرا مبتنی بر ترکیب تبدیل موجک و

شبکه‌های تطبیقی عصبی - فازی، چهاردهمین کنفرانس بین‌المللی مهندسی برق.

محمدی، تیمور، عاطفه تکلیف و ساحل زمانی (۱۳۹۶)، پیش‌بینی قیمت گاز طبیعی با استفاده از ترکیب تبدیل موجک و شبکه عصبی مصنوعی، فصلنامه پژوهش‌های اقتصادی ایران، شماره ۷۱، صص ۲۶-۱.

مشیری، سعید، کامران پاکیزه، منوچهر دبیریان و ابوالفضل جعفری (۱۳۸۹)، بررسی رابطه میان بازدهی سهام و تورم با استفاده از تجزیه موجک در بورس اوراق بهادار تهران، فصلنامه پژوهش‌های اقتصادی ایران، شماره ۴۲، صص ۷۴-۵۵.

نادری، اسماعیل (۱۳۹۱)، تحلیل آشوب و بررسی عملکرد مدل‌های خطی و غیرخطی سری زمانی در پیش‌بینی شاخص بازدهی بورس تهران، پایان‌نامه کارشناسی ارشد، دانشگاه تهران، دانشکده اقتصاد.

- Abbasi, E., & Abouec, A. (2008). Stock price forecast by using neuro-fuzzy inference system. In *Proceedings of World Academy of Science, Engineering and Technology*, 36, 320-323.
- Burrus, C. S., Gopinath, R. A., Guo, H., Odegard, J. E., & Selesnick, I. W. (1998). *Introduction to wavelets and wavelet transforms: a primer* (Vol. 1). New Jersey: Prentice hall.
- Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1), 134-144.
- Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, 23(3), 665-685.
- Karim, S. A. A., Karim, B. A., Ismail, M. T., Hasan, M. K., & Sulaiman, J. (2011). Applications of wavelet method in stock exchange problem. *Journal of Applied Sciences*, 11(8), 1331-1335.
- Cao, Q., Leggio, K. B., & Schniederjans, M. J. (2005). A comparison between Fama and French's model and artificial neural networks in predicting the Chinese stock market. *Computers & Operations Research*, 32(10), 2499-2512.
- Lineesh, M. C., & John, C. J. (2010). Analysis of non-stationary time series using wavelet decomposition. *Nature and Science*, 8(1), 53-59.
- Raofi, A. (2014). Identifying data generator process of Tehran stock

- exchange, modeling and forecasting using soft computing. Master's Thesis, Faculty of Economics and social Science, Shahid Chamran University. (in Persian).
- Raofi, A., Zarranezhad, M., Bayani, O. (2015). Assessment and Comparison of linear and non- linear Methods for Forecasting Returns on Stock Market Index, *International Journal of Scientific Management and Development*, 3 (10), PP. 1-9.
- Raofi, A., Montazer-Hojjat, A. H., & Kiani, P. (2016). Comparison of several combined methods for forecasting Tehran stock exchange index. *International Journal of Business Forecasting and Marketing Intelligence*, 2(4), 315-333.
- Srinivasan, K., & Fisher, D. (1995). Machine learning approaches to estimating software development effort. *IEEE Transactions on Software Engineering*, 21(2), 126-137.
- Tan, C., & Pedersen, C. N. S. (2009). Financial time series forecasting using improved wavelet neural network. *Master of Computer Science Faculty of Science. Univesity of Copenhagen. Copenhagen*.
- Gao, X., Xiao, F., Zhang, J., & Cao, C. (2004). Short-Term prediction of chaotic time series by wavelet networks. In *Intelligent Control and Automation, WCICA 2004. Fifth World Congress on* (Vol. 3, pp. 1931-1935). IEEE.
- Zarranezhad, M, Kiyani, P., Ebrahimi, S., Raofi, A. (2013). Forecasting OPEC crude oil price Using Fuzzy Autoregressive Integrated Moving Average (FARIMA) Model, *Iranian Energy Economics Research*, 5, 107-127. (in Persian).
- Zarranezhad, M., Raofi, A. (2015). Evaluation and comparison of forecast performance of linear and non-linear methods for daily returns of tehran stock exchange, *Financial Monetary Economics*, 22(9), 1-28. (in Persian).
- Zarranezhad, M., Raofi, A. and Kiani, P. (2012). Evaluation and comparison of performance of ANFIS and ARIMA in forecasting the daily gold prices, *The First international conference on econometrics and methods applications*. (in Persian).