

پیش بینی شاخص قیمت بورس سهام با استفاده از شبکه عصبی و تبدیل موجک

رضا راعی^۱، شاپور محمدی^۲، حنظله فندرسکی^{۳*}

۱- هیأت علمی گروه مدیریت مالی و بیمه دانشکده مدیریت دانشگاه تهران، تهران، ایران

raei@ut.ac.ir

۲- هیأت علمی گروه مدیریت مالی و بیمه دانشکده مدیریت دانشگاه تهران، تهران، ایران

shmohmad@ut.ac.ir

۳- کارشناسی ارشد مدیریت مالی دانشکده مدیریت دانشگاه تهران، تهران، ایران

hfendereski@ut.ac.ir

چکیده

شاخص بازار سرمایه به عنوان دماسنج اقتصادی هر کشور است. از این رو پیش‌بینی این متغیر جهت دید کلی از وضعیت اقتصادی و اخذ استراتژی‌های سرمایه‌گذاری، یکی از مسایل مهم به شمار می‌رود. از جمله روش‌های پیش‌بینی پرکاربرد در سری‌های زمانی مالی، شبکه عصبی است که با توجه به جامعیت این روش و عدم وجود برخی از پیش‌فرض‌ها در خصوص داده‌ها، گسترش زیادی نسبت به روش‌های آماری یافته است؛ اما وجود نویز در سری‌های زمانی به خصوص در سری‌های زمانی مالی و اقتصادی باعث کاهش دقت پیش‌بینی شبکه عصبی می‌شود. یکی از روش‌های نوفه‌زدایی در سری‌های زمانی، تبدیل موجک است. این پژوهش به مقایسه بین دقت پیش‌بینی شاخص کل بورس اوراق بهادار تهران در دو مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده با تبدیل موجک و شبکه عصبی با استفاده از داده‌های اولیه از ابتدای سال ۱۳۸۵ تا ۳۱ خرداد ۱۳۹۲ می‌پردازد. نتایج حاکی از بهبود معنادار در پیش‌بینی شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده است.

واژه‌های کلیدی: تبدیل موجک، شبکه عصبی، نوفه‌زدایی، آستانه

مقدمه

در کشورهایی که بازار سرمایه آنها توسعه یافته است و تعداد شرکت‌های پذیرفته شده در آن به نسبت تعداد شرکت‌های فعال در کشور زیاد است، بازار سرمایه آنها نمایشگر کاملی از وضعیت اقتصادی آن کشور است و شاخص بورس اوراق بهادار به عنوان یک نماگر اقتصادی به شمار می‌رود که با پیش‌بینی این نماگر می‌توان نسبت به وضعیت و مدیریت نقدینگی کشور و همچنین اخذ استراتژی‌های سرمایه‌گذاری مناسب اقدام نمود. به عنوان مثال در دوره‌هایی که شاخص برای یک دوره طولانی مدت روند صعودی به خود می‌گیرد، موجب می‌گردد که نقدینگی از سایر بازارهای موازی از قبیل ارز، سکه و مسکن به این بازار گسیل یابد. در این حالت معمولاً سرمایه‌گذاران استراتژی‌های خرید و نگهداری را اخذ می‌نمایند و اقدام به سرمایه‌گذاری با دیدهای بلندمدت می‌نمایند. در نقطه مقابل زمانی که شاخص برای یک دوره طولانی مدت روند نزولی به خود می‌گیرد، معمولاً نقدینگی به سمت بازارهای موازی بازار سرمایه انتقال می‌یابد و سرمایه‌گذاران در بازار سرمایه اقدام به اخذ استراتژی‌های سرمایه‌گذاری کوتاه‌مدت می‌نمایند. لذا همواره پیش‌بینی شاخص به عنوان یک نماگر اقتصادی مورد توجه سرمایه‌گذاران و مسئولان اقتصادی یک کشور است.

روش‌های پیش‌بینی در دو دسته کلی روش‌های ساختاری و سری زمانی طبقه‌بندی می‌گردد [۱۱]. در روش‌های ساختاری مقدار یک متغیر (متغیر وابسته) با توجه به سایر متغیرها (متغیرهای مستقل) تعیین می‌گردد در حالی که در مدل‌های سری زمانی، وقفه‌های زمانی متغیر وابسته به عنوان متغیر مستقل محسوب می‌گردد.

در روش‌های ساختاری علاوه بر پیش‌بینی متغیر وابسته، نیاز به پیش‌بینی مقدار متغیر یا متغیرهای مستقل نیز است در حالی که در روش‌های سری زمانی، وقفه‌های یک متغیر برای پیش‌بینی مقدار آتی همان متغیر به کار می‌رود. با توجه به سادگی مدل‌های سری زمانی، این روش مخصوصاً در حوزه‌های مالی و اقتصادی، گسترش زیادی یافته است. مدل‌های پیش‌بینی نیز به دو دسته کلی مدل‌های آماری و مدل‌های هوش مصنوعی تقسیم‌بندی می‌گردند. مدل‌های آماری شامل $ARIMA^1$ و $GARCH^2$ است. لازمه این مدل‌ها وجود رابطه خطی (به استثناء مدل $GARCH$)، توزیع نرمال و مانا بودن داده‌هاست که معمولاً این شرایط در بازارهای مالی مهیا نیست. مدل‌های هوش مصنوعی نیز شامل شبکه عصبی، سیستم فازی، الگوریتم ژنتیک و هستند که در بین آنها شبکه عصبی با توجه به ویژگی‌های آن، کاربرد گسترده‌ای یافته است. وجود برخی از مزیت‌ها سبب برتری مدل‌های هوش مصنوعی نسبت به مدل‌های آماری شده است. مزیتی که شبکه عصبی نسبت به مدل‌های اقتصادیسنجی دارد عدم نیاز به مانا بودن سری زمانی و وجود رابطه خطی است. به عبارت دیگر، شبکه عصبی به دلیل جامعیت توان تعیین هر گونه رابطه خطی و غیرخطی را دارا است [۲۸]. وجود عواملی چون رویدادهای سیاسی، سیاست‌های شرکت‌ها، شرایط کلی اقتصاد، نرخ‌های بانکی و ... بر قیمت سهام و شاخص تأثیر می‌گذارند. لازمه استفاده از چنین داده‌های خامی در شبکه عصبی، نوفه‌زدایی آنهاست. از جمله روش‌هایی که در نوفه‌زدایی کاربرد دارد، تبدیل موجک است. سوالی که در اینجا مطرح می‌گردد این

1. Autoregressive integrated moving average
2. Generalized autoregressive conditional heteroskedasticity

مطالعات جنسن در خصوص کاربرد تبدیل موجک در تخمین پارامترهای ناشناخته در فرآیندهای حافظه بلند مدت اشاره نمود.

در دسته سوم، توجه اصلی معطوف به نوفه‌زدایی در داده‌ها برای افزایش صحت پیش‌بینی است. در این حوزه می‌توان به مطالعات هاون و همکاران در خصوص پیش‌بینی قیمت اختیار خرید با استفاده از داده‌های نوفه‌زدایی شده اشاره نمود.

در این پژوهش تمرکز بر روی کاربرد سوم تبدیل موجک و استفاده آن در نوفه‌زدایی داده‌هاست. کائو^۱ و همکاران (۲۰۰۵) به مقایسه بین روش‌های خطی فاما و فرنج و CAPM و روش‌های غیرخطی این دو روش (شبکه عصبی) پرداختند. آنها معتقد بودند که در کشورهای در حال توسعه عواملی متفاوت از فاکتورهای موجود در کشورهای توسعه یافته، بر بازده سهام تأثیر می‌گذارد. این مطالعه بر روی ۳۶۷ شرکت پذیرفته شده در بورس شانگهای بین سال‌های ۱۹۹۹ تا ۲۰۰۲ صورت پذیرفت. نتایج نشان می‌دهد که در بین مدل خطی فاما و فرنج و مدل CAPM، مدل CAPM توانایی بالاتری را در پیش‌بینی بازده سهام دارا است و این تفاوت از لحاظ آماری با استفاده از آزمون دایوولد ماریانو، معنادار است. این نتایج در خصوص روش‌های غیرخطی دو مدل نیز تأیید گردید. در مقایسه بین روش‌های خطی و غیرخطی نیز، شبکه عصبی توانایی بالاتری در خصوص پیش‌بینی بازده سهام نسبت به سایر روش‌ها را داراست [۱۲]. ژانگ و کیو^۲ (۲۰۰۵) به بررسی توانایی شبکه عصبی در پیش‌بینی سری‌های زمانی فصلی و دارای روند پرداخته‌اند. نتایج حاصل از

است که آیا نوفه‌زدایی سبب افزایش دقت شبکه عصبی در پیش‌بینی شاخص سهام می‌گردد؟

برای پاسخ به این پرسش ابتدا باید نسبت به نوفه‌زدایی شاخص کل بورس اوراق بهادار اقدام نمود و نتایج را با پیش‌بینی شبکه عصبی با استفاده از داده‌های اولیه مقایسه نمود.

پیشینه پژوهش

نظریه موجک به عنوان یک روش پژوهش در تحلیل علمی گسترش بسیاری یافته است و در بسیاری از شرایط از تحلیل موجک به عنوان ابزار پژوهش استفاده شده و در اکثر موارد نتایج رضایت‌بخش بوده و نسبت به پژوهش‌های قبلی شرایط مناسب‌تری تجربه شده است. تحلیل موجک اثر قابل توجهی بر اغلب رشته‌ها از قبیل ریاضی، آمار، تحلیل سیگنال‌ها، تحلیل تصاویر، فشرده‌سازی داده‌ها، ژئوفیزیک و آنالیز عددی داشته است.

علی‌رغم خواص بسیار جالب موجک، هنوز کاربردهای آن در مالی گسترش نیافته است. در هر حال کاربردهای موجک در سه دسته عمده می‌توان طبقه‌بندی نمود [۱۵]:

– در دسته اول، تحلیل موجک برای تجزیه و تحلیل چند بعدی به کار می‌رود. از جمله مطالعات صورت گرفته در این دسته می‌توان به پژوهش رمزی و لمپارت اشاره نمود که از تبدیل موجک برای تجزیه و تحلیل رابطه بین مصرف و درآمد و عرضه پول و درآمد استفاده نمودند.

– در دسته دوم، از روش موجک برای پیش‌بینی متغیرهای ناشناخته در مدل استفاده می‌گردد. از جمله پژوهش‌های صورت گرفته در این حوزه می‌توان به

1. Cao

2. Zhang & Qi

پیش‌بینی داده‌ها نویزدار، نشان از دقت پایین شبکه عصبی است و با افزایش مرحله نویز در داده‌ها، خطای پیش‌بینی افزایش یافته است. سپس با استفاده از داده‌های روندزادایی شده به پیش‌بینی با استفاده از شبکه عصبی پرداخته شد و نتایج با روش ARIMA، مقایسه شد که نشان از بالا بودن دقت پیش‌بینی شبکه عصبی نسبت به مدل ARIMA است [۲۹]. تان (۲۰۰۹) اقدام به پیش‌بینی نرخ برابری برخی از ارزها یک گام پیش‌رو و چند گام پیش‌رو می‌پردازد. وی در این پایان‌نامه روش جدیدی را برای پیش‌بینی با استفاده از تبدیل موجک ارائه می‌نماید. در این مدل سری اصلی با استفاده از تجزیه موجک مانا، به ۴ لایه تقسیم می‌شود. سپس ۵ شبکه عصبی شامل ضرایب جزئیات و ضرایب تخمین تشکیل می‌دهد و بعد از آموزش و تست اقدام به پیش‌بینی یک گام پیش‌رو با استفاده از هر یک از این شبکه‌ها می‌نماید. از سوی دیگر از سری اصلی، یک سری ویژگی‌های آماری از جمله، میانگین، واریانس و... استخراج می‌گردد. این ویژگی‌های آماری به همراه پیش‌بینی ۵ شبکه به شبکه عصبی جدیدی وارد و اقدام به پیش‌بینی یک گام پیش‌رو می‌نماید. وی این روش را با روش‌های شبکه عصبی، شبکه عصبی با داده‌های نوفه‌زادایی شده با تبدیل موجک و شبکه عصبی با داده‌های نوفه‌زادایی شده با بسته موجک^۱ مقایسه می‌نماید. نتایج نشان می‌دهد در هر دو حالت پیش‌بینی یک گام پیش‌رو و چند گام پیش‌رو، مدل جدید، میانگین مجذور خطای (MSE) کمتری نسبت به سه مدل دیگر دارد، لذا دقت پیش‌بینی بالاتری دارد [۲۶]. جمازی و آلویی^۲ (۲۰۱۱) اقدام به

پیش‌بینی قیمت نفت با استفاده از تبدیل موجک و شبکه عصبی نمودند. آنها بیان نمودند با توجه به شوک‌های موجود در قیمت نفت ناشی از عواملی از جمله جریان‌های سیاسی، جنگ و یا سایر مسایل، پیش‌بینی قیمت نفت با استفاده از شبکه عصبی را با مشکلاتی روبه‌رو ساخته که این موضوع موجب اختلال و کاهش صحت پیش‌بینی شبکه عصبی می‌شود. لذا ابتدا با استفاده از تبدیل موجک به نوفه‌زادایی داده‌ها پرداخته و سپس با استفاده از شبکه عصبی پس‌انتشار خطا و بهینه‌سازی تعداد لایه‌های پنهان و تعداد نرون‌های موجود، اقدام به پیش‌بینی قیمت نفت نمودند و نتایج را با نتایج حاصل از پیش‌بینی شبکه عصبی با استفاده از داده‌های اولیه مورد مقایسه قرار دادند. نتایج حاکی از توانایی بالای شبکه عصبی با داده‌های نوفه‌زادایی شده در پیش‌بینی قیمت نفت نسبت به شبکه عصبی با داده‌های اولیه است. در این پژوهش برای نوفه‌زادایی داده‌ها از تابع هتر و شبکه عصبی چند لایه‌ای استفاده شده است [۱۸]. هاون^۳ و همکاران (۲۰۱۲) به نوفه‌زادایی نوسان ضمنی با استفاده از تبدیل موجک پرداخته و سپس با بهره‌گیری از مدل بلک-شولز به تعیین قیمت اختیار اقدام نمودند. سپس قیمت‌ها را با داده‌های واقعی مقایسه کردند. نتایج حاکی از بهبود صحت پیش‌بینی با استفاده از داده‌های نوفه‌زادایی شده است [۱۷]. لی^۴ و همکاران (۲۰۱۴) اقدام به پیش‌بینی قیمت نفت نمودند. آنها بر خلاف جمازی و آلویی از مدل ARIMA در پیش‌بینی استفاده نمودند. از آنجایی که این مدل یک مدل خطی به شمار می‌آید، لذا از ماشین بردار پشتیبان^۵ برای پیش‌بینی بخش غیرخطی

3. Haven

4. Li

5. Support Vector Machine

1. Wavelet Packet

2. Jammazi & Aloui

خانواده موجک مناسب گردد. انتخاب تابع موجک مناسب باعث می‌گردد عملیات تجزیه و بازسازی با کمترین خطا صورت پذیرد. عواملی از جمله وجود دامنه نوسان محدودتر و همچنین حجم مینا سبب می‌گردد که داده‌های شاخص بورس تهران با سایر شاخص‌های بورس‌های کشورهای مختلف، رفتاری متفاوت داشته باشد، لذا عدم اکتفا به پژوهش‌های گذشته و تعیین خانواده موجک مناسب محسوس است. همان‌طور که در ادامه اشاره شده است پس از آزمون، خانواده هتر به عنوان خانواده موجک مناسب در این مقاله انتخاب گردیده است.

مقاله پیش‌رو به بررسی عملکرد استفاده از داده‌های نوفه‌زدایی شده در پیش‌بینی شاخص بورس اوراق بهادار با استفاده از شبکه عصبی می‌پردازد. در این پژوهش از داده‌های روزانه شاخص استفاده شده است. برای نوفه‌زدایی داده‌ها از تبدیل موجک گسسته حداکثر هم پوشانی استفاده شده است که نسبت به تبدیل موجک گسسته که در بیشتر پژوهش‌های گذشته مورد استفاده قرار گرفته است، دارای مزایای است که در ادامه معرفی می‌شود. همچنین برای تعیین نوع خانواده موجک مناسب برای نوفه‌زدایی، برخی از خانواده‌های پر کاربرد با طول موج‌های مختلف آنها مورد مقایسه قرار می‌گیرد تا در نهایت بهترین خانواده موجک با طول مناسب انتخاب گردد.

روش پژوهش

در این پژوهش هدف بررسی دقت شبکه عصبی در پیش‌بینی شاخص بورس اوراق بهادار با استفاده از داده‌های نوفه‌زدایی است. لذا ابتدا داده‌ها با استفاده از تبدیل موجک، نوفه‌زدایی شده و سپس داده‌های

استفاده کردند. در این مقاله برای جلوگیری از تاثیر نوسان‌های قیمت نفت، از نوفه‌زدایی موجک با استفاده از آسانته اقدام نمودند که نتایج حاکی از بهبود قدرت پیش‌بینی مدل بوده است [۲۱]. عباسی‌نژاد و نادری (۱۳۹۱) در بورس اوراق بهادار تهران آشوب را تحلیل و پیش‌بینی‌پذیری را بررسی کرده و نیز عملکرد انواع مدل‌های شبکه عصبی را با کمک داده‌های تجزیه شده با روش موجک ارزیابی کردند. در این مقاله از تابع موجک گسسته استفاده شده است. برای انتخاب خانواده موجک گسسته براساس پژوهش‌های گذشته از تابع دابشیز (db3) استفاده شده است. براساس نتایج این مطالعه، سری بازدهی بورس در دوره بررسی شده، پیش‌بینی پذیر بوده و آثار غیرخطی معین و آشوبی داشته است. یافته دیگر این پژوهش نیز به برتری عملکرد مدل‌های شبکه عصبی چندلایه پیش‌خور و شبکه عصبی فازی مبتنی بر داده‌های تجزیه شده به کمک تجزیه موجک در مقابل به کارگیری سطح داده‌ها دلالت دارد [۵]. این مقاله از تبدیل موجک گسسته استفاده نموده است در حالی که در سال‌های اخیر با توسعه کاربرد تبدیل موجک، نوع اصلاح شده‌ای از تبدیل موجک گسسته، با نام تبدیل موجک گسسته حداکثر هم‌پوشانی معرفی شده است. این تبدیل دارای مزایایی نسبت به تبدیل موجک گسسته است که از جمله آن می‌توان به عدم حساسیت به نقطه شروع زمانی اشاره نمود. در ادامه به برخی از مزایای این تبدیل اشاره شده است. همچنین در این مقاله براساس پژوهش‌های گذشته، از خانواده دابشیز ۳ استفاده شده است. تبدیل موجک دارای خانواده‌های زیادی هستند که هر کدام از آنها متناسب با داده‌های خاصی هستند لذا لازم است با توجه به نوع داده‌ها اقدام به انتخاب

می‌دهد. ضرایب موجک نیز از روابط زیر به دست می‌آیند:

مدل (۲)

$$s_{j,k} \approx \int_{j,k}(t) f(t) dt$$

مدل (۳)

$$d_{j,k} \approx \int \Psi_{j,k}(t) f(t) dt$$

که $s_{j,k}$ ها را هموار^۱ سطح لام و $d_{j,k}$ ها را جزئیات^۲ سطح لام می‌نامند.

از سوی دیگر تابع تبدیل موجک گسسته نسبت به نقطه شروع سری زمانی حساس است، لذا از نسخه اصلاح شده تابع تبدیل موجک گسسته، تبدیل موجک گسسته حداکثر همپوشانی (MOWDT) برای نوفه‌زدایی داده‌ها استفاده می‌گردد. از جمله مزایای دیگر MOWDT نسبت به DWT، می‌توان به موارد ذیل اشاره نمود [۴]:

۱- با MODWT سری زمانی با هر طول را می‌توان بررسی کرد، اما در DWT محدود به سری‌های زمانی با طول‌های 2^j هستیم؛ ۲- وقایع و حرکات موجود در سری زمانی با تحلیل چند وقتی MODWT همراستایی بیشتری دارند؛ ۳- MODWT نسبت به جابجایی در زمان به صورت دوار سری زمانی حساس نیست؛ ۴- تخمین‌گر واریانس بر مبنای MODWT از تخمین‌گر DWT کارا تر است.

موجک‌ها دارای جنسیت هستند. موجک پدر^۳ معمولاً با \square و موجک مادر^۴ با نماد Ψ نمایش داده می‌شود که به صورت زیر تعریف می‌شوند:

مدل (۴)

$$j,k(t) = s^{-\frac{j}{2}} \left(\frac{t - s^j k}{s^j} \right)$$

نوفه‌زدایی شده به عنوان داده‌های ورودی برای پیش‌بینی به شبکه عصبی منتقل می‌شود. به عبارت دیگر، تبدیل موجک به عنوان یک روش پیش‌پردازش به کار می‌رود. لذا این بخش به دو قسمت تبدیل موجک و شبکه عصبی تقسیم می‌شود.

موجک‌ها توابع ریاضی هستند که داده‌ها را به اجزای فراوانی (فرکانس) تفکیک کرده و هر جز را با نمایش متناسب با مقیاس آن جز مطالعه می‌کنند [۸]. توابع موجک برای رفع نواقص تبدیل فوریه ارایه شده است. یکی از مزایای تبدیل موجک نسبت به روش‌های فوریه سنتی، توان بالای تحلیل آنها در شرایطی است که سیگنال‌ها با گسستگی و جهش‌های سریع همراه هستند. استفاده از روش موجک در مالی مزایای عمده‌ای دارد که از جمله آنها می‌توان به مطالعه مستقیم سری‌های زمانی نامانا و بررسی مدل‌ها و رفتارهای مالی در زمان مقیاس‌های متفاوت اشاره نمود.

تبدیل موجک پیوسته معمولاً حاوی حجم بالایی از اطلاعات است که نیازی به همه آنها در بازسازی تابع و مشاهده و تحلیل پیوستگی‌ها نیست و حتی می‌توان با نمونه‌برداری، ضرایب موجک را به نحوی به دست آورد که تابع اصلی را از آنها بازسازی نمود.

تقریب هر تابع گسسته یا سری زمانی با استفاده از توابع موجک به صورت زیر به دست می‌آید:

مدل (۱)

$$f(t) = \sum_{k=1}^{\infty} s_{j,k,j,k}(t) + \sum_{k=1}^{\infty} d_{j,k} \Psi_{j,k}(t) + \sum_{k=1}^{\infty} d_{j-1,k} \Psi_{j-1,k}(t) + \dots + \sum_{k=1}^{\infty} d_{1,k} \Psi_{1,k}(t)$$

که در فرمول فوق J تعداد سطوح آنالیز یا مقیاس‌ها و k مقدار حرکت در زمان در هر سطح را نشان

1. Smooth
2. Detail
3. Father Wavelet
4. Mother Wavelet

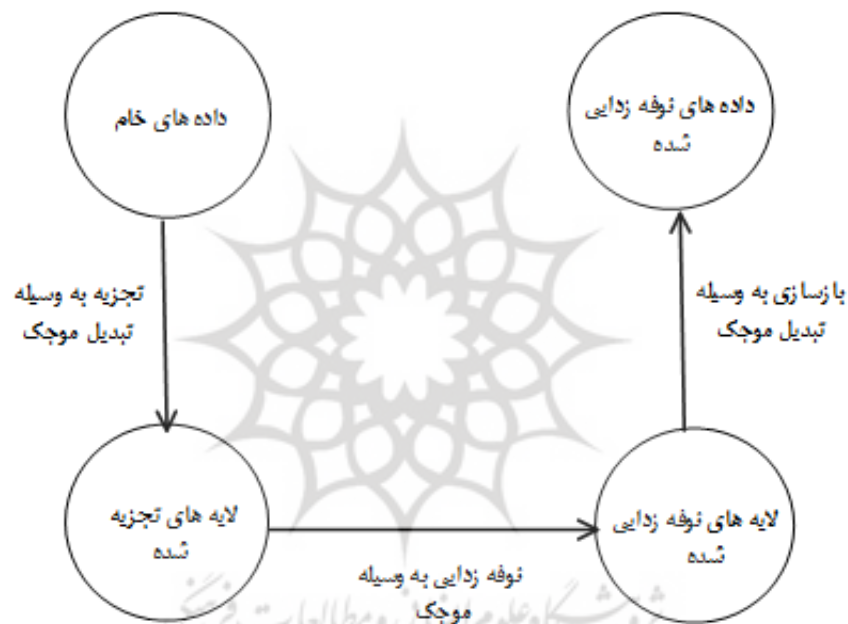
مدل (۵)

$$\Psi_{j,k}(t) = s^{-\frac{j}{2}} \Psi\left(\frac{t - s^j k}{s^j}\right)$$

سیملت، کویفلت، دابشیز، مورلت و... داشته باشند که هر یک از این اشکال دارای ویژگی‌هایی هستند که استفاده از آنها برای داده‌هایی با آن ویژگی مناسب است [۴].

در مدل‌های پیش‌بینی که در آن تبدیل موجک به عنوان یک روش نوفه‌زدایی به کار می‌رود، فرآیند نوفه‌زدایی در سه مرحله صورت می‌پذیرد.

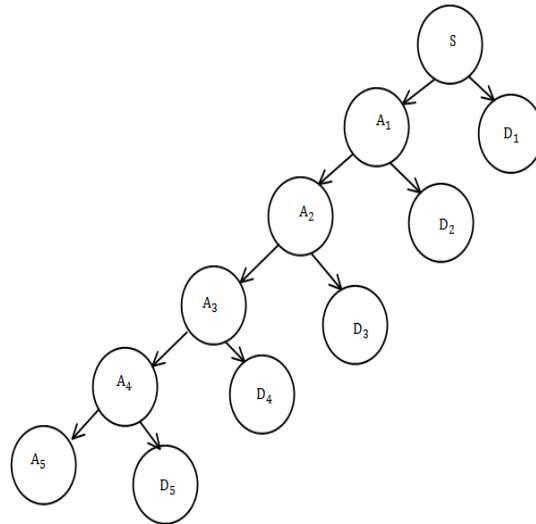
موجک پدر بخش هموار، روند (فرکانس پایین) سیگنال و موجک مادر بخش‌های جزئی (فرکانس بالا) را نشان می‌دهد. به طور کلی به (t) ها توابع مقیاسگر و به $\Psi(t)$ ها توابع موجک گفته می‌گویند. توابع Ψ و \square می‌توانند اشکال مختلفی از جمله هتر، کلاسه مکزیکی،



شکل (۱) نوفه زدایی با استفاده از تبدیل موجک

ابتدا داده‌های اولیه به وسیله تبدیل موجک به لایه‌هایی تجزیه می‌شود. تبدیل موجک را می‌توان به صورت تجزیه تکه به تکه اطلاعات در یک سری زمانی در نظر گرفت. در این فرآیند ابتدا داده‌های اولیه به وسیله تبدیل موجک به لایه‌های تجزیه می‌شوند که به این عمل تجزیه یا تحلیل^۱ تابع می‌گویند. هر بخش از سیگنال تجزیه شده را می‌توان به صورت یک ضریب

موجک و یک ضریب مقیاسگر در نظر گرفت. ضرایب موجک و ضرایب مقیاسگر در هر مرحله از اعمال فیلتر موجک و فیلتر مقیاسگر بر سری زمانی اصلی به دست می‌آیند که تکرار این فرآیند در غالب الگوریتم هرمی^۲ زیر صورت می‌پذیرد [۱].



شکل (۲) الگوریتم هرمی تجزیه موجک

S: داده‌های اصلی؛ A_i: لایه iام داده‌های کم‌فراوانی؛ D_i: لایه iام داده‌های پرفراوانی

نرم^۳ و ... اشاره نمود. در آستانه نرم همه ضرایب موجک به سمت صفر حرکت می‌کند. ضرایب بزرگتر از مقدار آستانه به سمت صفر و ضرایب کمتر از آستانه صفر می‌شود؛ در حالی که در آستانه سخت ضرایب یا صفر در نظر گرفته می‌شد و یا همان مقدار اولیه بدون تغییر باقی می‌ماند، لذا می‌توان آستانه نرم را پیوسته در نظر گرفت.

مدل (۷)

$$\delta_{\eta}^S(o_t) = \text{sign}(o_t)(|o_t| - \eta) \text{sign}(o_t) = \begin{cases} +1 & \text{if } o_t > 0 \\ 0 & \text{if } o_t = 0 \\ -1 & \text{if } o_t < 0 \end{cases}$$

o_t : مقدار ضریب موجک و η : مقدار آستانه

در این پژوهش با توجه به پیوسته بودن آستانه نرم و با عنایت به اینکه این آستانه با کلیه داده‌ها سروکار دارد که در نهایت به ارایه تابع هموارتری از سیگنال‌های اصلی منجر می‌شود، از این آستانه برای نوفه‌زدایی موجک استفاده شده است.

در واقع در تجزیه موجک داده‌ها به دو سری داده‌های پرفراوانی و کم‌فراوانی تقسیم می‌شوند. داده‌های کم‌فراوانی که از طریق اعمال موجک پدر بر سری اصلی حاصل می‌گردند، نشان‌دهنده ویژگی‌های اصلی سری است. داده‌های پرفراوانی نیز از طریق اعمال موجک مادر بر سری اصلی حاصل می‌گردند که اغلب نویز نامیده می‌شوند. در واقع هدف اصلی تجزیه موجک جداسازی ویژگی‌های اصلی سری از نویز است.

نوفه‌زدایی به وسیله موجک به دو دسته خطی و غیرخطی طبقه‌بندی می‌شود، اما به دلیل قابلیت انطباق روش موجک غیرخطی به تغییرات سریع در ویژگی‌های نامانایی داده‌ها، این روش نوفه‌زدایی موجک کاربرد بیشتری نسبت به روش دیگر دارد. یکی از پرکاربردترین روش نوفه‌زدایی موجک غیرخطی به وسیله آستانه^۱ است. روش‌های آستانه‌ای متفاوتی وجود دارد که از جمله آنها می‌توان به آستانه سخت^۲، آستانه

تخمین زنده انحراف معیار شدیداً تحت تأثیر ضرایب است، لذا معمولاً از تخمین زنده میانه قدر مطلق انحراف^۳ (MAD) مقیاس ضرایب موجک استفاده می‌گردد [۱۴].

مدل (۹)

$$\tilde{\sigma}_{MAD} = \frac{\text{median}(|w_{1,0}|, |w_{1,1}|, \dots, |w_{1, \frac{N}{2}-1}|)}{0.6745}$$

برای L متغیر مستقل گوسی X_1, X_2, \dots, X_L با انحراف معیار σ داریم:

$$E(\text{median}(|X_i|, 1 \leq i \leq L)) \approx 0.6745\sigma$$

پس از نوفه‌زدایی داده‌ها در مرحله سوم عملیات معکوس تجزیه، بر روی ضرایب موجک انجام می‌پذیرد تا سری زمانی نوفه‌زدایی شده تولید گردد [۲۱]:

مدل (۱۰)

$$x(t) = \frac{1}{c_{\Psi}} \int_0^{\infty} \int_{-\infty}^{\infty} w(u, s) \Psi_{u,s}(t) du \frac{ds}{s^2}$$

که به این عمل بازسازی یا سنتز^۴ گفته می‌شود. سپس داده‌های نوفه‌زدایی شده به عنوان ورودی به شبکه عصبی انتقال می‌یابد.

شبکه عصبی مصنوعی برگرفته از شبکه عصبی انسان است. خروجی و ورودی شبکه عصبی مصنوعی همانند متغیر مستقل و وابسته در مدل رگرسیون هستند. ورودی همانند رگرسیون کننده و خروجی همانند رگرسیون شونده است. پیش‌بینی پارامترها نیز در شبکه عصبی مشابه مدل رگرسیون است. پیش‌بینی پارامترها در شبکه عصبی اغلب آموزش نامیده می‌شود [۲۰].

پارامتر مهم در دستور آستانه ذکر شده، مقدار آستانه (η) است. آستانه بهینه زمانی اتفاق می‌افتد که پارامتر آستانه برابر با نویز باشد ($\eta = \sigma_{\square}$). انتخاب پارامتر آستانه کوچکتر از نویز ($\eta < \sigma_{\square}$) باعث باقی ماندن نویز در داده‌ها در زمان پیش‌بینی و انتخاب پارامتر آستانه بزرگتر از نویز ($\eta > \sigma_{\square}$) باعث از بین رفتن برخی از اطلاعات که مربوط به سیگنال هستند، می‌گردد. همچنین تعیین مقدار پارامتر آستانه برابر با مقدار نویز، به راحتی مقدور نیست و نیاز به دانستن σ_{\square} است که معین نیست، لذا روش‌هایی برای تعیین مقدار بهینه پارامتر آستانه تعیین گردیده است که بهترین آنها آستانه جامع است که به صورت ذیل محاسبه می‌گردد:

مدل (۸)

$$\eta^U = \tilde{\sigma}_{\epsilon} \sqrt{2 \log N}$$

منطق استفاده از پارامتر آستانه جامع^۱ عبارت است از: فرض می‌کنیم Z_1, Z_2, \dots, Z_N دارای توزیع یکسان و مستقل^۲ (IID) $N(0, \sigma_{\square}^2)$ باشد داریم:

$$P\left[\max(|z|) \leq \sigma_{\epsilon} \sqrt{2 \log N}\right] \rightarrow 1$$

$N \rightarrow \infty$

همانطور که اشاره شد این روش انتخاب آستانه در نمونه‌هایی با اندازه بزرگ، یک روش مناسب است، ولی در خصوص نمونه‌هایی با اندازه کوچک ممکن است استفاده از سایر روش‌ها مناسب باشد.

در این روش نیاز به تعیین σ_{\square} است در حالی که در اکثر موارد، انحراف معیار نویز مشخص نیست. اگر ما فرض کنیم که نویز در سیگنال وجود داشته باشد،

3. Median Absolute Deviation
4. Synthesizing or Reconstructing

1. Universal Thresholding
2. Independent and Identically Distributed (IID)

خطای پیش‌بینی اقدام گردد. در حالی که در شبکه عصبی بازگشتی، خروجی نرون‌ها در لایه‌های مختلف می‌تواند به لایه‌های قبلی نیز انتقال یابد.

مدل شبکه عصبی پیش‌خور چند لایه‌ای (MLP)^۵ یکی از پرکاربردترین مدل‌های پیش‌بینی در شبکه عصبی است. در MLP، اولین لایه، لایه ورودی است که اطلاعات خارجی و داده‌ها را دریافت می‌نماید. لایه خروجی نیز لایه‌ای است که جواب مسأله و نتیجه مدل را ارائه می‌نماید. لایه خروجی و لایه ورودی به وسیله یک یا چند لایه پنهان از یکدیگر متمایز می‌گردند.

برخلاف ویژگی‌های ممتاز شبکه عصبی، ساخت یک شبکه عصبی مناسب برای پیش‌بینی کار ساده‌ای نیست و در صورت استفاده از یک مدل نامناسب عملکرد شبکه کاهش می‌یابد. عواملی که در شبکه عصبی ممکن است بر روی عملکرد شبکه تأثیرگذار باشند، باید مدنظر قرار گیرند.

به طور معمول، شبکه عصبی ترکیبی از نرون‌های موجود در لایه‌ها است. در شبکه‌های چند لایه‌ای (MLP) معمولاً یک یا چند لایه پنهان وجود دارد. در طراحی یک شبکه چند لایه‌ای ابتدا باید موارد ذیل تعیین گردد: تعداد نرون‌های ورودی؛ تعداد لایه‌های پنهان و تعداد نرون‌های موجود در این لایه‌ها و تعداد نرون‌های خارجی. تعیین هر یک از موارد فوق وابسته به نوع کاربرد شبکه عصبی و هدفی است که شبکه برای آن طراحی می‌گردد و نمی‌توان یک راه‌حل کلی ارائه نمود.

تعداد نرون‌های ورودی با تعداد متغیرهایی که در بردار ورودی برای پیش‌بینی وارد می‌شود، تعیین می‌گردد. در مدل‌های سری زمانی تعداد نرون‌های

در شبکه عصبی مصنوعی هر نرون سیگنالی را از نرون‌های دیگر دریافت و براساس تابع فعال‌سازی^۱ پردازش نموده و با توجه به وزن آن به سایر نرون‌ها یا نرون خارجی انتقال می‌دهد. وزن ارزش محاسباتی برای قدرت نسبی ارتباط برای انتقال داده از یک لایه به لایه دیگر است. بعد از جمع نمودن وزن داده‌های ورودی با بایس^۲ و به کار بردن تابع فعال‌سازی، خروجی شبکه حاصل می‌گردد که این فرآیند را می‌توان به طور خلاصه در فرمول ذیل بیان نمود [۲۴]:

مدل (۱۱)

$$y_i = f\left(\sum_{j=1}^m (x_j w_{ij} + \theta_i)\right)$$

w_{ij} : وزن ارتباط بین نرون i و نرون j ; x_j : ارزش نرون j ; f : تابع فعال‌سازی؛ θ_i : بایس نرون i .

توانایی تخمین شبکه عصبی مصنوعی ناشی از ساختار موازی آن است که این موضوع به عنوان مزیت شبکه عصبی نسبت به سایر مدل‌ها به شمار می‌آید. شالوده مدل شبکه عصبی، نرون‌ها و لایه‌های پنهان هستند که برای پردازش اطلاعات به کار می‌روند.

مدل‌های متفاوتی از شبکه عصبی تاکنون ارائه شده‌اند که از جمله آنها می‌توان به شبکه عصبی پیش‌خور^۳ و شبکه عصبی بازگشتی^۴ اشاره نمود. در شبکه عصبی پیش‌خور، خروجی نرون در هر لایه (به جز لایه خروجی) به عنوان ورودی به لایه بعدی انتقال می‌یابد. در این شبکه جهت تعیین خروجی، ورودی‌ها یکبار از شبکه عبور می‌کنند سپس نتایج با مقدار واقعی مقایسه می‌گردند تا با تعدیل در وزن‌ها نسبت به کاهش

1. Activation Function
2. Bias
3. Feed Forward ANN
4. Recurrent ANN

5. Multi-layer Perceptron

را که تفاضل بین مقدار پیش‌بینی و مقدار واقعی است را ارزیابی می‌کند. از جمله معیار ارزیابی عملکرد صحت پیش‌بینی شبکه عصبی عبارت است از:

معیار مجذور میانگین خطا [۲۱]:

مدل (۱۲)

$$MSE = \left(T^{-1} \sum_{n=1}^T (C_n^{MP} - C_n)^2 \right)$$

که در آن T اندازه نمونه C_n^{MP} داده‌های واقعی C_n

داده‌های پیش‌بینی شده است.

برای مقایسه بین دقت پیش‌بینی شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده در مقابل شبکه عصبی با استفاده از داده‌های اولیه، از شاخص کل بورس اوراق بهادار تهران استفاده شده است.

همانطور که اشاره شد هدف از انجام پژوهش، پیش‌بینی شاخص بورس به عنوان یک دماسنج اقتصادی است در بازار سرمایه کشور ما دو بورس در حوزه اوراق بهادار فعالیت می‌نماید. شاخص فرابورس ایران به دلیل نوپا بودن و همچنین محدود بودن شرکت‌های پذیرفته شده در آن توانایی بالایی در خصوص ارایه وضعیت اقتصادی ندارد، لذا جامعه آماری این پژوهش شاخص‌های بورس اوراق بهادار تهران است که از بین آنها با توجه به اینکه شاخص کل میانگین موزون تمام شرکت‌های پذیرفته شده در این بورس است، به عنوان نمونه مورد استفاده قرار گرفته شده است. اطلاعات مربوط به شاخص کل از ابتدای سال ۱۳۸۵ تا انتهای خرداد ۱۳۹۲ به صورت روزانه از تارنمای شرکت بورس اوراق بهادار تهران به آدرس www.irbourse.com استخراج شده است. تعداد کل مشاهدات در این دوره ۱۶۷۷ مشاهده بوده که تعداد ۳۰ مشاهده برای مقایسه دقت پیش‌بینی دو مدل و سایر

ورودی وابسته به تعداد وقفه‌های مورد استفاده در مدل است.

تعداد لایه‌های پنهان و تعداد نرون‌های موجود در این لایه‌ها نقش مهمی در دستیابی موفقیت پیش‌بینی با استفاده از شبکه عصبی ایفا می‌نمایند. تعداد نرون‌های موجود در لایه پنهان است که به شبکه عصبی اجازه می‌دهند ویژگی‌ها و الگوهای موجود در داده‌ها را تشخیص و رابطه غیرخطی بین متغیرهای ورودی و خروجی را تعیین نمایند. به عبارت دیگر، شبکه عصبی بدون لایه پنهان و با تابع خطی در لایه خروجی همانند یک مدل آماری خطی عمل می‌نماید. پژوهش‌های صورت پذیرفته شده بیان می‌نمایند که استفاده بیش از ۲ لایه پنهان در شبکه عصبی بهبود معناداری در عملکرد شبکه عصبی ارایه نمی‌نماید. همچنین در بیشتر پژوهش‌ها استفاده از یک لایه پنهان در شبکه عصبی نتایج قابل قبولی ارایه نموده است. در خصوص تعداد نرون‌های موجود در لایه‌های پنهان می‌توان بیان نمود که تعداد نرون‌های کم باعث کاهش مسایل بیش‌برازشی می‌گردد. از سوی دیگر پایین بودن زیاد تعداد نرون‌ها در لایه پنهان موجب کاهش توانایی شبکه در یادگیری و مدل‌سازی داده‌ها می‌شود. تعیین تعداد نرون‌های خارجی به سادگی با توجه به مسأله تحت بررسی تعیین می‌گردد. در مباحث پیش‌بینی، تعداد نرون‌های خارجی برابر با افق پیش‌بینی است [۲۸].

معیارهای مختلفی برای ارزیابی عملکرد شبکه عصبی استفاده می‌شود که از جمله آنها می‌توان به زمان آموزش، زمان مدل‌سازی و... اشاره کرد. یکی از مهمترین معیارهای ارزیابی عملکرد معیار صحت پیش‌بینی شبکه عصبی است. این معیار خطای پیش‌بینی

مشاهده‌ها برای آموزش و آزمون شبکه عصبی استفاده شده است.

یافته‌های پژوهش

پژوهش فرآیندی نظام‌مند است که در آن پژوهشگر به دنبال پاسخی برای مسأله پژوهش است و لازمه پاسخگویی به مسأله پژوهش تجزیه و تحلیل داده‌ها، تعبیر و تفسیر آنها و بررسی نتایج به دست آمده است.

بنابر آنچه گفته شد برای آزمون فرضیه پژوهش مبنی بر دقت بالاتر شبکه عصبی با داده‌های نوفه‌زدایی شده نسبت به شبکه عصبی با استفاده از داده‌های اولیه، ابتدا اقدام به نوفه‌زدایی داده‌ها می‌گردد. با توجه به مزایای موجک گسسته حداکثر همپوشانی، از این موجک برای نوفه‌زدایی داده‌ها استفاده می‌گردد، لیکن نیاز به انتخاب خانواده موجک مناسب است. انتخاب

تابع موجک مناسب باعث می‌گردد عملیات تجزیه و بازسازی با کمترین خطا صورت پذیرد. در این پژوهش نیز به منظور پیدا کردن نوع و طول موجک مناسب برای تجزیه سری زمانی، برای طول‌های مختلف توابع موجک دابشیز، کویفلت و سیملت، سری زمانی شاخص را تجزیه کرده و سپس با انجام عمل معکوس با استفاده از تبدیل موجک گسسته حداکثر همپوشانی، سری را بازسازی می‌کنیم. سپس ریشه میانگین مجذور خطا را میان سری اولیه و سری بازسازی شده محاسبه و خانواده موجکی که کمترین مقدار معیار RMSE را به خود اختصاص دهد به عنوان خانواده موجک مناسب انتخاب می‌گردد. مقادیر مربوط به RMSE به شرح جدول ۱ است: (با توجه به کوچک بودن RMSE میان خانواده‌های مختلف موجک، اعداد جدول ذیل برای نمایش بهتر در 10^6 ضرب شده‌اند).

جدول (۱) مقایسه انواع خانواده موجک برای شاخص بورس اوراق بهادار تهران

خانواده موجک	هتر	دابشیز ۲	دابشیز ۳	دابشیز ۴	دابشیز ۵	سیملت ۲	سیملت ۳
RMSE	۰/۰۰۰۰۰۶۷۵	۰/۰۰۰۰۰۷۶۱۷	۰/۲۱۰۲	۰/۰۳۹۰	۰/۰۶۲۹	۰/۰۲۷۱	۰/۳۵۲۱
خانواده موجک	سیملت ۴	سیملت ۵	کویفلت ۱	کویفلت ۲	کویفلت ۳	کویفلت ۴	کویفلت ۵
RMSE	۰/۰۱۴۷	۰/۳۲۵۵	۹۰/۶۳۸۰	۰/۰۱۸۶	۰/۲۱۰۲	۰/۰۱۵۰	۰/۰۰۵۱

بین میانگین دو جامعه می‌نماییم. برای بررسی معنادار بودن تفاوت بین دو جامعه RMSE را در دو حالت با استفاده از فرضیه زیر و آزمون مقایسه زوجی با استفاده از نرم‌افزار EXCEL به شرح ذیل بررسی می‌نماییم:

H_0 : بین دو جامعه میانگین مجذور خطای موجک هتر و موجک دابشیز ۲ در استفاده از تبدیل موجک د رخصوص شاخص کل بورس اوراق بهادار تهران تفاوت معناداری وجود ندارد.

با توجه به جدول فوق، RMSE برای خانواده‌های مختلف تبدیل موجک بسیار پایین است به عبارت دیگر تفاوت قابل توجهی میان استفاده از خانواده‌های مختلف تبدیل موجک وجود ندارد، اما در جدول فوق موجک هتر و موجک دابشیز ۲، کمترین مقدار ریشه میانگین مجذور خطا را به خود اختصاص دادند. با توجه به اینکه مقدار RMSE بسیار کوچک است، لذا اقدام به آزمون مقایسه زوجی برای معنادار بودن تفاوت

خصوص شاخص کل بورس اوراق بهادار تهران تفاوت معناداری وجود دارد.

H_1 : بین دو جامعه میانگین مجذور خطای موجک هتر و موجک دابشیز ۲ در استفاده از تبدیل موجک در

جدول (۲) نتایج آزمون مقایسه زوجی خانواده موجک

آزمون t: مقایسه زوجی میانگین		
دابشیز ۲	هتر	
۵/۸E-21	۴/۵۵E-21	میانگین
۵/۶۹E-41	۲/۱۱E-41	واریانس
۱۶۷۷	۱۶۷۷	مشاهده‌ها
	۰/۷۰۴۰۰۱	همبستگی پیرسون
	۰	تفاوت میانگین آزمون شده
	۱۶۷۶	درجه آزادی
	-۹/۴۶۰۹۳	آماره t
	۰/۰۰۰۰	$P(T \leq t)$ یک طرفه
	۱/۶۴۵۷۶۳	t بحرانی یک طرفه
	۰/۰۰۰۰	$P(T \leq t)$ دو طرفه
	۱/۹۶۱۳۸	t بحرانی دو طرفه

شبکه بهینه برای هر پیش‌بینی از کدنویسی در نرم افزار متلب (MATLAB) استفاده شده است. در این فرآیند با توجه به پژوهش‌های پیشین تعداد نرون‌های موجود در لایه اول (وقفه‌های شاخص) بین ۱ تا ۱۰ وقفه، تعداد لایه‌های پنهان بین ۱ و ۲ لایه و تعداد نرون‌های موجود در لایه اول بین ۱۰ تا ۲۰ نرون و نرون‌های موجود در لایه پنهان دوم نیز بین ۰ تا ۱۰ نرون متغیر است. سپس شبکه بهینه براساس میانگین مجذور خطا (MSE) تعیین و با استفاده از آن مقدار شاخص برای ۱۶۴۸ امین داده پیش‌بینی می‌گردد (یک گام پیش‌رو) سپس مقدار واقعی داده ۱۶۴۸ جایگزین مقدار پیش‌بینی می‌شود و مجدداً به عنوان ورودی به شبکه منتقل می‌گردد تا پس از انتخاب شبکه بهینه، نسبت به پیش‌بینی مقدار شاخص برای ۱۶۴۹ امین داده اقدام گردد. این فرآیند تا پیش‌بینی ۱۶۷۷ امین داده (۳۰ پیش‌بینی) ادامه می‌یابد.

با توجه به جدول فوق و آماره t، فرضیه H_0 مبنی بر عدم وجود تفاوت معنادار رد و فرضیه H_1 تأیید می‌گردد. بنابراین بین دو جامعه میانگین مجذور خطا (MSE) موجک هتر و موجک دابشیز ۲، تفاوت معناداری وجود دارد و با توجه به اینکه میانگین MSE برای موجک هتر در نمونه مورد بررسی از موجک دابشیز ۲ کمتر است، موجک هتر به طور معناداری از موجک دابشیز در تبدیل موجک شاخص کل بورس اوراق بهادار تهران در دوره مورد بررسی، بهتر است. لذا برای تجزیه و بازسازی سری زمانی در فرآیند نوفه‌زدایی شاخص، از موجک هتر استفاده می‌گردد.

پس از نوفه‌زدایی داده‌های در مرحله اول ۱۶۴۷ داده به عنوان ورودی به شبکه عصبی منتقل می‌گردد که از این تعداد ۸۰ درصد مربوط به بخش آموزش و ۲۰ درصد مربوط به بخش تست هستند. برای تعیین

در جدول زیر شبکه بهینه برای هر پیش‌بینی و مقدار
پیش‌بینی شاخص با استفاده از داده‌های نوفه‌زدایی شده
و با استفاده از داده‌های اولیه ارایه شده است:

جدول (۳) مقدار پیش‌بینی شاخص سهام با استفاده از داده‌های اولیه

ردیف	مقدار واقعی شاخص	مقدار پیش‌بینی شاخص	تعداد وقفه‌های بهینه	تعداد نرون‌های موجود در لایه پنهان اول	تعداد نرون‌های موجود در لایه پنهان دوم
۱	۴۴/۷۰۷	۴۳/۰۳۲	۵	۱۹	۸
۲	۴۴/۶۸۴	۴۳/۵۴۸	۶	۱۴	۶
۳	۴۵/۰۹۱	۳۹/۸۱۹	۶	۱۶	۴
۴	۴۵/۵۱۳	۴۰/۵۴۹	۸	۱۴	۵
۵	۴۴/۹۶۳	۴۵/۳۴۰	۳	۱۴	۶
۶	۴۴/۱۲۹	۴۲/۲۸۶	۴	۱۴	۴
۷	۴۳/۸۶۴	۴۴/۸۱۷	۹	۱۶	۶
۸	۴۳/۳۰۱	۴۱/۴۶۸	۷	۱۷	۱
۹	۴۲/۷۹۶	۴۰/۴۹۰	۸	۱۵	۶
۱۰	۴۳/۰۴۷	۴۱/۷۵۵	۷	۱۸	۳
۱۱	۴۲/۸۵۶	۴۱/۹۸۰	۳	۱۵	۱
۱۲	۴۲/۷۷۷	۳۹/۵۵۱	۸	۱۶	۲
۱۳	۴۳/۴۸۵	۴۱/۵۵۵	۵	۱۸	۴
۱۴	۴۳/۹۱۱	۴۱/۹۴۳	۵	۱۶	۳
۱۵	۴۴/۱۰۷	۴۱/۰۵۸	۳	۱۸	۴
۱۶	۴۴/۴۷۱	۴۱/۷۷۷	۳	۱۹	۷
۱۷	۴۵/۱۶۲	۴۱/۵۵۹	۶	۱۸	۷
۱۸	۴۴/۹۴۵	۴۳/۰۴۷	۱۹	۱۹	۰
۱۹	۴۵/۲۴۸	۴۱/۹۳۶	۷	۱۸	۷
۲۰	۴۵/۵۹۰	۴۳/۵۶۵	۱۹	۱۹	۰
۲۱	۴۵/۴۸۸	۴۱/۳۰۸	۳	۱۵	۲
۲۲	۴۵/۱۹۹	۴۶/۲۸۵	۲	۱۴	۰
۲۳	۴۵/۲۰۷	۴۱/۷۳۱	۶	۱۹	۶
۲۴	۴۵/۵۵۳	۴۳/۶۰۰	۷	۱۶	۷
۲۵	۴۵/۶۷۲	۴۱/۵۹۹	۲	۱۳	۰
۲۶	۴۶/۶۲۳	۴۰/۹۶۶	۸	۱۸	۶
۲۷	۴۷/۸۱۸	۴۳/۴۷۳	۷	۱۶	۷
۲۸	۴۸/۳۳۳	۵۰/۴۲۳	۳	۱۷	۰
۲۹	۴۸/۷۸۳	۴۲/۹۸۸	۵	۱۵	۵
۳۰	۴۸/۹۱۷	۴۴/۶۹۵	۳	۱۳	۰

جدول (۴) مقدار پیش‌بینی شاخص با استفاده از داده‌های نوفه‌زدایی شده

ردیف	مقدار واقعی شاخص	مقدار پیش‌بینی شاخص	تعداد وقفه‌های بهینه	تعداد نرون‌های موجود در لایه پنهان اول	تعداد نرون‌های موجود در لایه پنهان دوم
۱	۴۴/۷۰۷	۴۲/۰۹۵	۵	۱۴	۵
۲	۴۴/۶۸۴	۴۲/۲۴۲	۵	۱۱	۷
۳	۴۵/۰۹۱	۴۲/۴۸۰	۶	۱۷	۶
۴	۴۵/۵۱۳	۴۳/۱۳۷	۱	۱۸	۰
۵	۴۴/۹۶۳	۴۳/۰۴۱	۹	۱۸	۰
۶	۴۴/۱۲۹	۴۱/۶۳۹	۸	۱۷	۴
۷	۴۳/۸۶۴	۴۱/۶۳۹	۸	۱۳	۵
۸	۴۳/۳۰۱	۴۱/۴۷۹	۲	۱۷	۰
۹	۴۲/۷۹۶	۴۴/۳۷۳	۴	۱۲	۰
۱۰	۴۳/۰۴۷	۴۲/۵۷۶	۲	۱۲	۰
۱۱	۴۲/۸۵۶	۴۱/۰۸۱	۶	۲۰	۶
۱۲	۴۲/۷۷۷	۴۳/۱۵۴	۵	۱۸	۷
۱۳	۴۳/۴۸۵	۴۲/۹۲۷	۷	۱۵	۸
۱۴	۴۳/۹۱۱	۴۲/۰۷۷	۵	۱۹	۰
۱۵	۴۴/۱۰۷	۴۲/۱۶۸	۵	۱۵	۶
۱۶	۴۴/۴۷۱	۴۴/۳۰۹	۵	۱۴	۷
۱۷	۴۵/۱۶۲	۴۴/۳۵۶	۴	۱۵	۸
۱۸	۴۴/۹۴۵	۴۳/۲۸۰	۵	۱۷	۱
۱۹	۴۵/۲۴۸	۴۳/۴۵۹	۲	۲۰	۸
۲۰	۴۵/۵۹۰	۴۳/۸۱۶	۶	۱۱	۱
۲۱	۴۵/۴۸۸	۴۳/۲۳۷	۷	۱۷	۸
۲۲	۴۵/۱۹۹	۴۴/۵۳۸	۷	۱۸	۰
۲۳	۴۵/۲۰۷	۴۳/۰۹۲	۶	۱۴	۸
۲۴	۴۵/۵۵۳	۴۴/۲۷۶	۵	۱۳	۰
۲۵	۴۵/۶۷۲	۴۳/۷۲۳	۲	۱۴	۸
۲۶	۴۶/۶۲۳	۴۵/۳۸۱	۴	۱۰	۴
۲۷	۴۷/۸۱۸	۴۵/۳۹۳	۶	۱۷	۷
۲۸	۴۸/۳۳۳	۴۴/۱۸۳	۲	۱۵	۶
۲۹	۴۸/۷۸۳	۴۶/۹۶۷	۴	۱۶	۰
۳۰	۴۸/۹۱۷	۴۶/۶۰۵	۵	۱۲	۰

مقایسه بین دقت پیش‌بینی هر یک از دو مدل، میانگین مجذور خطا (MSE) به شرح ذیل محاسبه شده است:

بعد از پیش‌بینی شاخص با استفاده از هر یک از مدل‌های شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده و شبکه عصبی با استفاده از داده‌های اولیه، جهت

جدول (۵) معیار میانگین مجذور خطا در پیش‌بینی دو مدل

میانگین مجذور خطا در شبکه عصبی با استفاده از داده‌های اولیه	میانگین مجذور خطا در شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده	ردیف	میانگین مجذور خطا در شبکه عصبی با استفاده از داده‌های اولیه	میانگین مجذور خطا در شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده	ردیف
۷/۲۵۶/۰۲۰	۲۶/۱۴۷	۱۶	۲/۸۰۶/۹۶۵	۶/۸۲۴/۶۳۴	۱
۱۲/۹۸۳/۰۵۰	۶۴۹/۹۵۸	۱۷	۱/۲۹۰/۲۶۹	۵/۹۶۲/۸۷۶	۲
۳/۶۰۲/۷۸۴	۲/۷۷۲/۵۵۸	۱۸	۲۷/۷۹۸/۲۰۲	۶/۸۱۹/۴۱۰	۳
۱۰/۹۶۸/۰۱۹	۳/۱۹۹/۸۰۵	۱۹	۲۴/۶۳۸/۳۱۸	۵/۶۴۳/۹۵۰	۴
۴/۱۰۰/۶۲۵	۳/۱۴۷/۰۷۶	۲۰	۱۴۱/۹۰۳	۳/۶۹۵/۲۳۷	۵
۱۷/۴۷۴/۰۷۲	۵/۰۶۷/۹۰۱	۲۱	۳/۳۹۶/۶۴۹	۶/۲۰۰/۱۰۰	۶
۱/۱۷۹/۱۷۹	۴۳۷/۰۵۳	۲۲	۹۰۷/۶۳۷	۴/۹۵۱/۹۶۰	۷
۱۲/۰۷۹/۱۰۰	۴/۴۷۱/۱۱۰	۲۳	۳/۳۵۹/۸۸۹	۳/۳۱۹/۶۸۴	۸
۳/۸۱۵/۷۷۲	۱/۶۳۱/۷۵۱	۲۴	۵/۳۱۵/۳۳۰	۲/۴۸۸/۵۰۶	۹
۱۶/۵۸۸/۵۱۴	۳/۷۹۸/۲۱۱	۲۵	۱/۶۷۰/۰۳۹	۲۲۲/۱۲۴	۱۰
۳۲/۰۰۱/۶۴۹	۱/۵۴۲/۵۶۴	۲۶	۷۶۶/۸۵۰	۳/۱۴۹/۵۶۰	۱۱
۱۸/۸۷۸/۱۵۶	۵/۸۸۰/۱۴۰	۲۷	۱۰/۴۰۸/۳۶۶	۱۴۱/۹۷۸	۱۲
۴/۳۶۶/۸۴۶	۱۷/۲۲۴/۹۹۰	۲۸	۳/۷۲۵/۶۷۲	۳۱۱/۵۸۷	۱۳
۳۳/۵۸۲/۰۲۵	۳/۲۹۷/۸۵۶	۲۹	۳/۸۷۱/۸۴۳	۳/۳۶۲/۴۵۶	۱۴
۱۷/۸۲۲/۷۵۱	۵/۳۴۳/۹۵۷	۳۰	۹/۲۹۸/۸۴۰	۳/۷۶۱/۲۷۲	۱۵

برای بررسی معناداری تفاوت عملکرد مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده (WDANN) و مدل شبکه عصبی با استفاده از داده‌های اولیه (ANN) و آزمون فرضیه پژوهش، از آزمون دایبولد ماریانو و آماره تعدیل شده آن، آماره هاروی-لیورن-نیبولد به منظور بررسی تفاوت میانگین مجذور خطا برای شاخص کل بورس اوراق بهادار تهران استفاده شده که نتایج به شرح ذیل است:

H_0 : بین دو جامعه میانگین مجذور خطای مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده و

همان‌گونه که جدول فوق نشان می‌دهد در اکثر موارد مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده دقت بالاتری نسبت به مدل شبکه عصبی بدون استفاده از داده‌های نوفه‌زدایی شده در پیش‌بینی شاخص کل دارد و در صورتی که تفاوت بین معیار عملکرد از لحاظ آماری معنادار باشد، بیانگر عملکرد بهتر مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده نسبت به مدل شبکه عصبی با استفاده از داده‌های اولیه در پیش‌بینی شاخص بورس اوراق بهادار تهران برای یک گام پیش‌رو است.

مدل شبکه عصبی با داده‌های اولیه تفاوت معناداری ندارد.

مدل شبکه عصبی با داده‌های اولیه تفاوت معناداری ندارد.

H_1 : بین دو جامعه میانگین مجذور خطای مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده و

جدول (۶) نتایج آزمون معناداری دایبولد ماریانو

	میانگین	آماره‌ی دایبولد - ماریانو	آماره هاروی - لیورن - نیوبولد
$MSE_{WDANN} - MSE_{ANN}$	-۶/۰۲۵/۰۰۰	-۳/۳۳۷۸	-۳/۲۸۱۷

استفاده از داده‌های اولیه است و شبکه عصبی برای تنها یک مرحله پیش‌رو طراحی شده است از ۱۶۴۷ مشاهده برای آموزش و تست استفاده شده است. سپس بعد از تعیین شبکه بهینه اقدام به پیش‌بینی شاخص برای ۱۶۴۸ امین روز می‌گردد. در مرحله بعد میزان واقعی شاخص برای ۱۶۴۸ امین مشاهده جایگزین مقدار پیش‌بینی شده می‌شود و مجدداً به عنوان ورودی به شبکه داده می‌شود تا بعد از تعیین شبکه بهینه اقدام به پیش‌بینی متغیر تا ۱۶۴۹ امین روز می‌گردد. این فرآیند تا پیش‌بینی مقدار ۱۶۷۷ امین شاخص (۳۰ مرتبه) در هر یک از دو مدل صورت پذیرفت. سپس معیار مجذور خطا (MSE) برای هر مرتبه محاسبه شد.

با توجه به اینکه نتایج آزمون دایبولد ماریانو دال بر تفاوت معنادار میان میانگین MSE در مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده نسبت به مدل دیگر است و از آنجایی که میانگین این معیار در مدل شبکه عصبی با داده‌های نوفه‌زدایی شده کمتر از مدل دیگر است، لذا براساس معیار MSE، شبکه عصبی با داده‌های نوفه‌زدایی شده دقت بالاتری در پیش‌بینی شاخص دارد. لذا فرضیه پژوهش مبنی بر اینکه شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده نسبت به شبکه عصبی با داده‌های اولیه، عملکرد بهتری در

با مقایسه آماره محاسبه شده در جدول فوق با مقادیر بحرانی جداول Z و t-student می‌توان نتیجه گرفت که از نظر معیار میانگین مجذور خطا، تفاوت معنادار بین دو جامعه میانگین مجذور خطا مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده و شبکه عصبی با استفاده از داده‌های اولیه وجود دارد و فرضیه H_0 رد و فرضیه H_1 تأیید می‌گردد و با توجه به اینکه میانگین جامعه میانگین مجذور خطای شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده نسبت به مدل شبکه عصبی با استفاده از داده‌های اولیه کمتر است، لذا مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده عملکرد بهتری نسبت به مدل شبکه عصبی با داده‌های اولیه در پیش‌بینی شاخص سهام دارد.

نتیجه‌گیری و پیشنهادها

به منظور مقایسه عملکرد مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده نسبت به مدل شبکه عصبی با داده‌های اولیه، مقدار روزانه شاخص از ابتدای سال ۱۳۸۵ تا آخر خرداد ماه ۱۳۹۲ (۱۶۷۷ مشاهده) به عنوان نمونه انتخاب گردید. از آنجایی که هدف این پژوهش مقایسه عملکرد مدل شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده نسبت به مدل شبکه عصبی با

پیش‌بینی شاخص بازار سهام دارد، مورد تأیید قرار می‌گیرد.

نتایج این پژوهش با نتایج پژوهش‌هایتان (۲۰۰۹)، جمازی و آلویی (۲۰۱۱) و هاون و همکارانش (۲۰۱۲) در خصوص استفاده از تبدیل موجک برای نوفه‌زدایی داده‌ها و افزایش دقت شبکه عصبی در پیش‌بینی سری‌های زمانی مالی منطبق هستند. به عبارت دیگر، نوفه‌زدایی داده‌ها با استفاده از تبدیل موجک گسسته حداکثر همپوشانی سبب افزایش دقت پیش‌بینی مدل شبکه عصبی در پیش‌بینی شاخص کل بورس اوراق بهادار تهران می‌گردد.

همان‌طور که اشاره شد، سرمایه‌گذاران حقیقی و حقوقی، صندوق‌های سرمایه‌گذاری، شرکت‌های سبدگردان و شرکت‌های مشاور سرمایه‌گذاری معمولاً به دنبال پیش‌بینی شاخص کل و یا قیمت سهام شرکت‌ها هستند، اما از آنجایی که معمولاً این سری‌های زمانی دارای نویز است پیشنهاد می‌گردد که با عنایت به تأیید فرضیه پژوهش، افرادی که به پیش‌بینی شاخص سهام و یا سایر متغیرهای مالی در قالب سری زمانی اقدام می‌نمایند، بهتر است پیش از استفاده از مدل‌های پیش‌بینی، نسبت به نوفه‌زدایی داده‌ها اقدام نموده و سپس داده‌های نوفه‌زدایی شده به عنوان ورودی به مدل‌های پیش‌بینی انتقال یابد.

در این پژوهش سعی در بررسی دقت پیش‌بینی شبکه عصبی با استفاده از داده‌های نوفه‌زدایی شده در مقابل داده‌های اولیه صورت پذیرفت و نتایج حاکی از آن است که نوفه‌زدایی داده‌ها سبب افزایش دقت پیش‌بینی می‌گردد. در این پژوهش از تبدیل موجک برای نوفه‌زدایی داده استفاده شده است در حالی که روش‌های دیگری از جمله تجزیه اجزا مستقل (ICA) نیز وجود دارد. لذا پیشنهاد می‌گردد برای پژوهش‌های

آتی روش‌های دیگر نوفه‌زدایی نیز مورد بررسی قرار گیرد. به عبارت دیگر، توان روش‌های متفاوت نوفه‌زدایی با یکدیگر مورد مقایسه قرار گیرد و تعیین گردد که کدام روش نوفه‌زدایی توانایی بالاتری را در پیش‌بینی سری‌های زمانی دارا است.

منابع

- [۱] انصاری، حجت اله. (۱۳۸۶). بررسی تأثیر استفاده از مقیاس‌های زمانی متفاوت در محاسبه ارزش در معرض ریسک (VAR) با استفاده از نظریه موجک، پایان‌نامه کارشناسی ارشد، دانشکده مدیریت دانشگاه تهران.
- [۲] آذر. عادل؛ کریمی، سیروس. (۱۳۸۸). پیش‌بینی بازده سهام با استفاده از نسبت‌های حسابداری با رویکرد شبکه عصبی، مجله تحقیقات مالی، ۲۸، ۲۰-۳.
- [۳] رحیمی، محمدرضا. (۱۳۹۱). پیش‌بینی دامنه‌ی تغییرات طلا با استفاده از مدل تلفیقی ARIMA و شبکه‌ی عصبی، پایان‌نامه کارشناسی ارشد، دانشکده مدیریت دانشگاه تهران.
- [۴] شمس، شهاب الدین. (۱۳۸۷). بررسی زمان مقیاس مدل قیمت‌گذاری دارایی سرمایه‌ای از طریق تبدیل موجک، پایان‌نامه دکترا، دانشکده مدیریت دانشگاه تهران.
- [۵] عباسی‌نژاد، حسین؛ نادری، اسماعیل. (۱۳۹۱). تحلیل آشوب، تجزیه موجک و شبکه عصبی در پیش‌بینی شاخص بورس تهران، فصلنامه تحقیقات مدل‌سازی اقتصادی، ۸.
- [۶] قدیمی، محمدرضا؛ مشیری، سعید. (۱۳۸۱). مدل‌سازی رشد اقتصادی در ایران با استفاده از

- Time-Scales, *International Review of Financial Analysis*, 15, 203-219.
- [15] Gency. Ramazan, Selcuk. Faruk, Witcher. Brandon. (2002). An Introduction to Wavelet and Other Filtering Methods in Finance and Economics, Academic press.
- [16] Harvery, David I. (1997). The Evaluation of Economic Forecasts, PhD Thesis.
- [17] Haven. Emmanuel, Liu. Xiaoquan, Shen. Liya. (2012). De-nising Option Prices with the Wavelet Method, *European Journal of Operational Research*, 222.
- [18] Jammazi. Rania, Aloui., Chaker. (2011). Crude Oil Price Forecasting: Experimental Evidence from Wavelet Decomposition and Neural Network Modeling, *Energy Economics xxx*, 1-14.
- [19] Johnstone. I. M, Silverman. B. W. (1997). Wavelet Threshold Estimators for Data with Correlated Noise, *Journal of the Royal Statistical Society*, 95, 1-37.
- [20] Kaastra. I, Boyd. M. (1996). Designing a Neural Network for Forecasting Financial Economic Time Series, *Neurocomputing*, 10, 215-236.
- [21] Li. Xia, He. Kaijian, Lai. King Keung, Zou. Yingchao. (2014). Forecasting Crude Oil Price with Multiscale Denoising Ensemble Model, *Mathematical Problems in Engineering*, 1-19.
- [22] Misiti, Michel. Misiti, Yves. Oppenheim, Georges. Poggi, Jean-Michel. (2009). Wavelet Toolbox 4, User's Guide. Mathworks.
- [23] Ramsey. James B, Lampart. Camille. (1998). The Decomposition of Economic Relationships by Time Scale Using Wavelet: Expenditure and income, *Studies in nonlinear dynamics and econometrics*, 3, 23-42.
- [24] Ruey. H. (2003). Time series forecast with neural network and wavelet techniques, Submitted for the Degree of Bachelor of Engineering.
- [25] S. Maia. Andre Luis, A.T.de. Carvalho. Francisco, B. Ludermir. Teresa. (2008). Forecasting Models for Interval-Valued Time Series, *Neurocomputing*, 71, 3344-3352.
- [26] Tan, Chong. (2009). Financial Time Series Forecasting Using Improved Wavelet Neural Network, Master's Thesis.
- شبکه‌های عصبی مصنوعی (ANN)، فصلنامه پژوهش‌های اقتصادی ایران، ۱۲، ۳۳-۱.
- [۷] قنبری، علی؛ خضری، محسن؛ ترکی سمایی، رقیه. (۱۳۸۸). تخمین ریسک سیستماتیک در مقیاس‌های زمانی مختلف با استفاده از آنالیز موجک برای بورس اوراق بهادار تهران، فصلنامه اقتصاد مقداری، ۴، ۲۹-.
- [۸] محمدعلی‌زاده، آر.ش. (۱۳۸۸). بررسی رابطه بین بازده سهام و تورم در بورس اوراق بهادار تهران در زمان مقیاس‌های مختلف با استفاده از تبدیل موجک، پایان‌نامه کارشناسی ارشد، دانشکده مدیریت دانشگاه تهران.
- [۹] محمدی، شاپور؛ عباسی‌نژاد، حسین. (۱۳۸۴). تحلیل سیکل‌های تجاری ایران با استفاده از نظریه موجک‌ها، مجله تحقیقات اقتصادی ایران، ۷۵، ۲۰-۱.
- [۱۰] نمازی، محمد؛ کیامهر، محمد مهدی. (۱۳۸۶). پیش‌بینی بازده روزانه سهام شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران با استفاده از شبکه‌های عصبی مصنوعی، مجله تحقیقات مالی، ۴۴، ۱۳۴-۱۱۵.
- [11] Brooks. Chris. (2008). Introductory Econometrics for Finance, Cambridge University Press.
- [12] Cao. Qing, Leggio. B. Kary, Schniederjans. J. Marc. (2005). A Comparison between Fama and French's Model and Artificial Neural Network in Predicting the Chinese Stock Market, *Computers & Operational Research*, 32, 2499-2512.
- [13] Donoho. D.L, Johnstone. I. M. (1995). Adapting to Unknown Smoothness by Wavelet Shrinkage, *Journal of the American Statistical Association*, 90, 613-627.
- [14] Fernandez. Viviana. (2006). The CAPM and Value at Risk at Different

- the Art, *International Journal of Forecasting*, 14, 35-62.
- [29] Zhang. G. Peter, Qi. Min. (2005). Neural Network Forecasting for Seasonal and Trend Time Series, *European Journal of Operational Research*, 160, 501-514.
- [27] Wong. Bo.K, Selvi. Yakup. (1998). Neural Network Application in Finance: A Review and Analysis of Literature, *Information & Management*, 34, 129-139.
- [28] Zhang. Guoqiang, Patuwo. B. Eddy, Y. Hu. Michael. (1998). Forecasting with Artificial Neural Networks: The State of

