فصلنامــه علمی پژوهشی دانش سرمایه‌گــذاری
سال اول/ شماره چهارم/ زمستان ۱۳۹۱

پیش بینی شاخص بورس تهران با استفاده از سری زمانی فازی بر اساس تعریف نرخ بازده

فرید رادمهر
کارشناسی ارشد مهندسی صنایع -مهندسی مالی، دانشکده مهندسی صنایع دانشگاه صنعتی امیرکبیر (مسئول مکاتبات)
f.radmehr@aut.ac.ir

ناصر شمس قارنه
استادیار دانشکده مهندسی صنایع دانشگاه صنعتی امیرکبیر
nshams@aut.ac.ir

تاریخ دریافت: ۹۱/۵/۱۸ تاریخ پذیرش: ۹۱/۹/۳۰

چکیده

در سالیان اخیر تحقیقات گسترده ای برروی مدل ها ی سری زمانی فازی انجام شده است اما در بسیاری از این تحقیقات، همواره فضای مسئله[۱] و بازه های مربوطه، بر اساس سطوح داده ها ی سری زمانی تعیین شده است. در این تحقیق با نگاهی جدید به تعیین فضای مسئله و استفاده از مفهوم نرخ بازده در بازارهای مالی، نوع جدیدی از فضای مسئله بر اساس نرخ بازده برای کاربرد در بازار های مالی و پیش بینی سری های زمانی مالی ارائه شده است. یکی از مسائل دیگر در مدل های سری زمانی فازی که تاثیر به سزایی در عملکرد آنها دارد طول بازه های مورد استفاده و نحوه ی تقسیم بندی فضای مسئله می باشد که در این زمینه تحقیقات متنوعی انجام شده است اما نتایج حاصله تا کنون راضی کننده نیست. لذا در این تحقیق با استفاده از الگوریتم شبیه سازی تبرید سعی در برطرف نمودن ایرادات مطالعات قبلی برای تعیین بازه های مناسب شده است.

حاصل تحقیق مدل RBFTS است. برای مقایسه عملکرد مدل ارائه شده و مدل های موجود در ادبیات، از دو مسئله ی بورس تایفکس[۲] و پذیرش دانشگاه آلاباما که به عنوان مرجع مقایسه ی این دسته از مدل ها هستند استفاده شده است. نتایج حاصله نشان دهنده ی برتری مدل های ارائه شده نسبت به مدل های پیشین است. در نهایت به عنوان مورد اجرایی، دو مدل نامبرده برروی شاخص بازار بورس تهران اجرا شده و نتایج تحلیل گردید.

واژه‌های کلیدی: سری زمانی فازی مرتبه ی بالا[۳]، نرخ بازده، شبیه سازی تبرید، پیش بینی.

۱- مقدمه

پیش بینی یکی از پایه های اصلی در تصمیم گیری قلمداد می شود. یک پیش بینی دقیق می تواند تا حد زیادی موجب کاهش تصمیمات غلط شده و توانایی رقابتی سازمان را ارتقا دهد. امروزه در بسیاری از شاخه های صنعت پیش بینی نقش بسزایی را ایفا می کند به طور مثال در پیش بینی تعداد مریض های مراجعه کننده به اورژانس بیمارستان، تعداد تصادفات جاده ای، پیش بینی آب وهواو پیش بینی های اقتصادی و مالی در بازار های مالی.مدل های متنوعی برای پیش بینی سری های زمانی تا کنون مطرح شده است که می توان به مدل های ARMA و ARIMA به عنوان مشهورترین مدل های کلاسیک اشاره نمود. مدل های سنتی سری های زمانی عملکرد نسبتا قابل قبولی دارند اما اشکالاتی بر این مدلها وارد است که از توانایی این مدل ها در پیش بینی سری های زمانی می کاهد. روش های سنتی سری زمانی مانند ARIMA می توانند سری های زمانی فصلی را پیش بینی کنند اما در پیش بینی مسائلی با داده های زبانی با مشکل روبرو می شوند. علاوه بر این، سری های زمانی سنتی برای پیش بینی به تعداد زیادی داده نیازمندند در حالی که در برخی مسائل واقعی داده به اندازه ی کافی موجود نیست (Jilani, 2007) و بدین ترتیب امکان اجرای مدل های سنتی برروی این سری زمانی وجود ندارد. همچنین در مدل های سنتی برای دستیابی به نتیجه ی مطلوب، پراکندگی داده ها می بایستی از تابع توزیع نرمال پیروی کند (Jilani, 2007) که این فرض نیز در بسیاری موارد نقض می شود. از سوی دیگر، متغیرهای زبانی برای بیان مشاهدات روزانه بکار می رود و در بسیاری از موارد نمی توان متغیرهای کیفی را به متغیرهای کمی تبدیل نمود. لذا برای حل چنین مسائلی مدل های سری زمانی فازی توسط سانگ و چیسام (Song and Chissom, 1993b, Song and Chissom, 1993a, Song and Chissom, 1994) مطرح گردید. نکته ی بسیار مهم در مدل های سری زمانی فازی نحوه ی تعیین مجموعه های فازی و اساس تعریف این مجموعه ها است. در قریب به اتفاق تحقیقات پیشین مجموعه های فازی براساس سطوح داده ها مطرح شده اند و تحقیقات ناچیزی در این زمینه انجام شده است، در این مقاله با تمرکز بر این مسئله تعریف جدیدی از مجموعه های فازی بر اساس نرخ بازده تدوین گردیده است همچنین در قدمی دیگر از الگوریتم شبیه سازی تبرید برای بهبود دقت بازه بندی و تعیین بازه بندی مناسب استفاده شده است که در نهایت نتایج به دست آمده نشان دهنده ی برتری مدل طراحی شده (RBFTS) نسبت به سایر مدل های موجود در ادبیات است.

سوالاتی که این تحقیق در صدد پاسخ گویی به آنها می باشد مطرح شده است. سه سوال که در این تحقیق به دنبال جواب آن بودیم را می توان به صورت زیر بیان نمود:

- آیا می توان مفهوم نرخ بهره را به عنوان یک متغیر فازی در نظر گرفت؟

- آیا اضافه نمودن این مفهوم به مدل های سری زمانی می تواند به افزایش دقت این مدل ها منجر شود؟
- آیا مدل های سری زمانی فازی کارایی لازم را در پیش بینی داده های مالی دارا هستند؟

با مطالعه ی نوع رفتار نرخ بازده در داده های بورسی دریافتیم که تغییرات این متغیر بسیار محدود تر از خود داده های قیمت می باشد؛ همچنین می توان تغییرات نرخ بازده را در مجموعه های فازی جدیدی به متغیرهای فازی تبدیل نمود. با تغییر نوع مجموعه های فازی و تعریف مجموعه های فازی جدید بر اساس نرخ بهره مدل جدیدی طراحی گردید که نتایج حاصل از اجرای این مدل برروی داده های دانشگاه آلاباما و شاخص بورس تهران نشان دهند ه ی کارایی این مسئله خواهد بود که در بخش بعدی مورد بررسی قرار گرفته است.

۲- مرور ادبیات و پیشینه پژوهش

مبحث منطق فازی برای اولین بار توسط زاده در مجامع علمی مطرح گردید، (Zadeh, 1965) دانشمندان با استفاده از این نظریه و منطق مدل ها ی کاربردی بسیاری در زمینه های مختلف برای استفاده از این دانش ارائه کردند، که می توان از میان آنها به مدل های سری زمانی فازی اشاره نمود. برای نخستین بار مفهوم سری های زمانی فازی توسط سانگ و چیسام (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) براساس مفاهیم مجموعه های فازی برای پیش بینی پذیرش دانشگاه آلاباما ارائه گردید بعدها اشکالاتی توسط چن (Chen, 1996) در مورد این مدل مطرح گردید. او معتقد بود استفاده از این عملگرها در مدل های مذکور باعث ایجاد پیچیدگی فراوان در مسئله پیش بینی می گردد، لذا وی با ساده سازی محاسبات عددی مربوط به عملگرهای ترکیبی[۴] باعث بهبود و افزایش دقت محاسبات نسبت به مدل سانگ و چیسام گردید. عملکرد بهتر و سادگی اجرای مدل چن باعث شد تااکثریت محققان از این مدل برای طراحی مدل های خود استفاده کنند.

در اکثر قریب به اتفاق مقالات موجود در ادبیات موضوع، سطوح داده ها به عنوان مجموعه های فازی و مقادیر زبانی تعریف می شوند به گونه ای که بازه ی تغییرات داده را با روشی به چند بازه ی کوچکتر تقسیم کرده که هرکدام از بازه ها نمایانگر و معرف یک متغیر زبانی می باشد. بخش بسیارکوچکی از مطالعات انجام شده تا کنون سعی در ایجاد نو آوری در این بخش نموده و تقریبا تغییرات چندانی در این بحث اتفاق نیفتاده که در این میان می توان به مقاله ی هوانگ وهمکاران (Hwang et al., 1998) اشاره کرد که وی در این مقاله برای اولین بار به جای خود داده ها، تغییرات

داده ها را وارد مدل کرده است و سعی در بازه بندی تغییرات داده نموده است در این مقاله، متغیر های زبانی بر اساس تفاضل دو داده متوالی تعریف شده اند و مقادیرزبانی "کاهش زیاد"، " کاهش"، " بدون تغییر"، " افزایش"، "افزایش زیاد"، "افزایش شدید" برای پیش بینی داده های دانشگاه آلاباما استفاده شده است. با اینکه در این مقاله ماهیت بازه ها از سطوح داده ها به تغییرات داده ها تغییر یافت، اما همچنان استفاده از بازه های مساوی مشکل این مدل به شمار می آید. ملیکه و همکاران (Melike and Konstsntin, 2004) روش پیش بینی با استفاده از سری های زمانی مرتبه اول را معرفی نمودند، انها نیز در مقاله ی خود از تغییرات داده ها برای تعیین مقادیر زبانی استفاده کرده اند، در این مقاله نیز همانند مقاله ی قبل عمل شده است و تنها با تغییر تعداد بازه ها نتایج بهتری نسبت به مدل هوانگ (Hwang et al., 1998) به دست آورده اند. وانگ و هسو (Wang and Hsu, 2008) نیز در مقاله خود تغییرات داده را همانند دو مقاله ی پیشین وارد مدل خود کردند اما مشتق دوم تغییرات یا به عبارتی تفاضل دو تغییر متوالی را با استفاده از الگویی ابتکاری وارد مدل کرده است به گونه ای که به عنوان معیار کمکی در پیش بینی پریودهای آتی به مدل کمک کند برای نشان دادن کارائی مدل، داده‌های میزان گردشگر به تایوان از کشورهای هنگ‌کنگ، آمریکا و آلمان طی دوره زمانی سال‌های ۱۹۹۸الی ۲۰۰۰ به‌کار برده شده است. از روش ابتکاری استفاده شده در سال ۱۹۹۳ توسط سانگ و چیسام (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) برای غیر فازی‌سازی مدل استفاده شده است.

مسئله ی تعیین بازه های مناسب در مسائل پیش بینی سری های زمانی فازی برای اولین بار توسط هوارنگ (Huarng, 2001a) مورد بررسی قرار گرفت. بعدها باگسترش استفاده از مدلهای ابتکاری و فرا ابتکاری در محاسبات گوناگون، هوارنگ نیز در مقاله ی خود برای نخستین بار مدل سری زمانی فازی را با مدل فراابتکاری شبکه عصبی در آمیخت (Huarng, 2001b) لی وهمکارانش (Lee et al., 2007, Lee et al., 2008) از دو مدل فرا ابتکاری ژنتیک و شبیه سازی تبرید برای پیش بینی دمای هوا و شاخص بورس تایوان (تایفکس) استفاده کردند. چن و چانگ (Chen and Chung, 2006a, Chen and Chung, 2006b) نیز با استفاده از الگوریتم فرا ابتکاری ژنتیک، مدلهایی مرتبه اول و مرتبه ی بالا را برای پیش بینی داده های دانشگاه آلاباما ارائه کرده اند. پارک و همکاران (Park et al., 2010) مدلی دو فاکتوره و مرتبه بالا را با کمک الگوریتم فراابتکاری PSO برای پیش بینی شاخص تایفکس و کوپسی ۲۰۰[5] ارائه دادند. کوو وهمکارانش نیز در مقاله ای دیگر(Kuo et al., 2009) با استفاده از الگوریتم فرا ابتکاری PSO[6] مسئله ی بازه بندی را مورد بررسی قرار داد و توانست با این تکنیک دقت پیش بینی را برای داده های آلاباما تا حد خوبی ارتقا دهد. همچنبن کوو و همکاران در مقاله ی(Kuo et al., 2010) با تغییر مدل خود و ایجاد مدل جدیدی به نام HPSO توانستند نتایج مدل را ارتقا دهند.

با توجه به ادبیات موضوع، در این مقاله نوع جدیدی از تعریف بازه بندی بر اساس نرخ بازده تعریف می شود و در ادامه برای دستیابی به ترکیبی مناسب از بازه های پیش بینی از الگوریتم شبیه سازی تبرید برای این فرایند استفاده شده ست.

۳- مدلهای ریاضی و مدل منتخب پژوهش

۳-۱- سری های زمانی فازی

سانگ و چیسام (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) در مقاله ها ی خود اصولی را برای این مدل ها معرفی نمودند که تمامی مدل های ایجاد شده در این حیطه و همچنین مدل ارائه شده در این مقاله بر اساس سری های زمانی فازی، بر این اصول استوارند. این اصول عبارتند:

اصل۱- (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) سری های زمانی فازی. ($Y(t)(t = \ldots, 0, 1, 2, \ldots$) مقادیر متغیر در لحظه ی t) را زیر مجموعه ای از اعداد حقیقی در نظر بگیرید، و همچنین دامنه ی تغییرات متغییر توسط مجموعه های فازی $fj(t)$ افراز شده باشد، در چنین شرایط اگر $F(t)$ مجموعه ای از $f1(t), f2(t), \ldots$ باشد، در نتیجه $F(t)$ یک سری زمانی فازی برروی $y(t)$ می باشد.

اصل ۲- (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) اگر یک رابطه ی فازی $R(t - 1, t)$ به صورت $F(t) = F(t - 1) \text{ o } R(t - 1, t)$ برقرار باشد و "o" نشان دهنده ی عملگر فازی باشد و $F(t)$ و $F(t-1)$ هرکدام یک مجموعه ی فازی باشند، در نتیجه می توان گفت $F(t)$ توسط $F(t-1)$ بوجود آمده است. رابطه فازی بین $F(t)$ و $F(t-1)$ به صورت زیر نشان داه می شود.

$$F(t-1) \rightarrow F(t)$$

اصل ۳- (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) $F(t - 1)$ و $F(t)$ را برابر با A_i و A_j در نظر بگیرید. رابطه ی بین $F(t-1)$ و $F(t)$ به عنوان یک رابطه ی منطقی فازی در نظر گرفته می شود (FLR)[۷] که به صورت $A_i \rightarrow A_j$ نیز مشخص می گردد که در آن A_i به عنوان دست چپ رابطه (LHS)[۸] و A_j به عنوان دست راست (RHS)[۹] رابطه تعیین میشوند.

اصل ۴- (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) تمام روابط فازی موجود در داده ها را می توان در دسته هایی بنا بر طرف چپ هر رابطه قرار داد، به عبارت دیگر، روابطی که دارای طرف چپ یکسان باشند را می توان در یک دسته قرار داد.به

عنوان مثال دو رابطه ی $Ai \rightarrow Aj_1$ و $Ai \rightarrow Aj_2$ با توجه به دست چپ یکسان، در یک گروه قرار می گیرند.

اصل ۵- (Song and Chissom, 1993a, Song and Chissom, 1993b, Song and Chissom, 1994) فرض کنید F(t) فقط توسط F(t-1) بوجود امده باشد، و برای هر t داشته باشیم F(t).R(t-1,t) = F(t - 1)، R(t-1,t) مستقل از t باشد، در نتیجه F(t) یک سری زمانی مستقل از زمان نامیده میشود. در غیر این صورت وابسته به زمان خواهد بود.

۳-۲- استفاده از نرخ بازده برای تعیین مجموعه های فازی به عنوان متغیر تحقیق

تعریف جدید ی از بازه بندی و مجموعه های فازی بر اساس نرخ بازده[۱۰] ارائه شده است. در تمـامی مدل های پیشین از خود قیمت و یا داده برای پیش بینی استفاده شده است اما با توجه به مزیت هـای استفاده از نرخ بازده به جای خود داده، در تحقیق سعی شده تا با استفاده از این تعریف مدلی با قابلیت پیش بینی بهتر برای بازارهای مالی طراحی گردد.

با توجه به تعاریف کتاب جان هال (Hull, 2002) بازده قیمت سـهام پـس از زمـان T دارای توزیـع لاگ نرمال[۱۱] و به عبارت دیگر لگاریتم قیمت سهام در زمان T دارای توزیع نرمال است.

$$\ln S_T \sim \emptyset\left[\ln S_0 + \left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \quad (۱)$$

و یا

$$\ln\frac{S_T}{S_0} \sim \emptyset\left[\left(\mu - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T}\right] \quad (۲)$$

اگر نرخ بازده واقعی (نرخ ترکیب پیوسته[۱۲]) سالانه روی سهام τ ، قیمت سهم پس از زمان T، S_T و قیمت سهم در حال حاضر S_0 باشد، باید داشته باشیم:

$$S_T = S_0 e^{\tau T} \quad (۳)$$

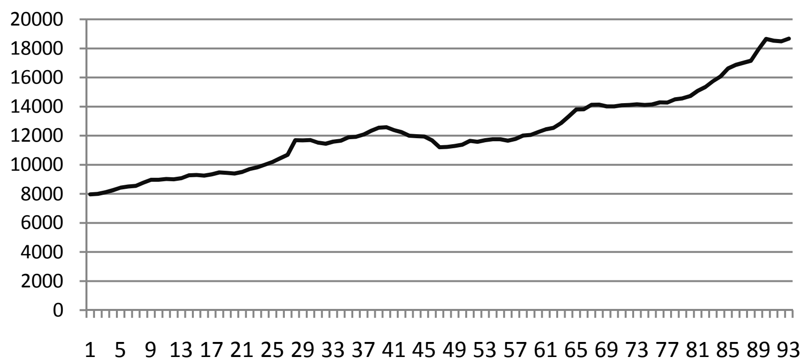
از طرفین لگاریتم گرفته و خواهیم داشت:

$$\tau = {}^{1}/_{T} \ln\frac{S_T}{S_0} \quad (۴)$$

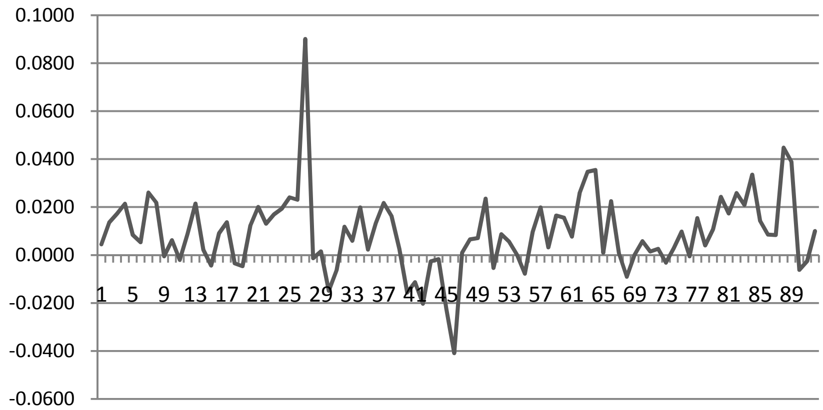
در معادله ی (۴) $1/T$ عدد ثابت است و از آنجاییکه $\ln\frac{S_T}{S_0}$ طبق معادلـه ی (۲) دارای توزیـع نرمـال است. τ نیز دارای توزیع نرمال است که میانگین و انحراف معیار آن با لحاظ ضریب $1/T$ به صـورت زیـر می باشد:

$$\tau \sim \emptyset\left[\mu-\frac{\sigma^2}{2},\frac{\sigma}{\sqrt{T}}\right] \qquad (5)$$

بنابراین نرخ بازده واقعی سهم پس از گذشت زمان T دارای یک توزیع نرمال است که میانگین آن $\mu-\frac{\sigma^2}{2}$ و انحراف معیار آن $\frac{\sigma}{\sqrt{T}}$ است. همانطور که مشاهده می شود در بازار های مالی شناسایی رفتار بازده سهم بسیار ساده تر از مطالعه ی رفتار خود قیمت سهام است. حال با توجه به اینکه در مطالعه ی موردی این تحقیق پریودهای زمانی قیمت مورد استفاده، همه هم اندازه بوده و پریود های پیش بینی نیز به همین صورت می باشند لذا می توانیم مقدار عددی T را در محاسبه ی بازده سهام مساوی با ۱ در نظر بگیریم. شکل ۱ نیز نشان دهنده ی روند حرکت شاخص بورس تهران از تاریخ ۱۳۸۸/۱/۸ تا ۱۳۸۹/۷/۱ می باشد که شامل ۹۳ هفته است. و شکل ۲ نیز نمودار نرخ بازده را برای این شاخص طی ۹۳ هفته نشان می دهد.همانطور که مشاهده می شود میزان تغیرات و بازه ی حرکتی نرخ بازده بسیار محدودتر از روند حرکتی خود شاخص است و بیش از ۹۵٪ داده های نرخ بازده در بازه ی [۰/۰۴، ۰/۰۴-] قرار دارد و به همین دلیل انتظار می رود، پیش بینی این فاکتور با توجه تغییرات محدود آن با خطای کمتری همراه باشد.



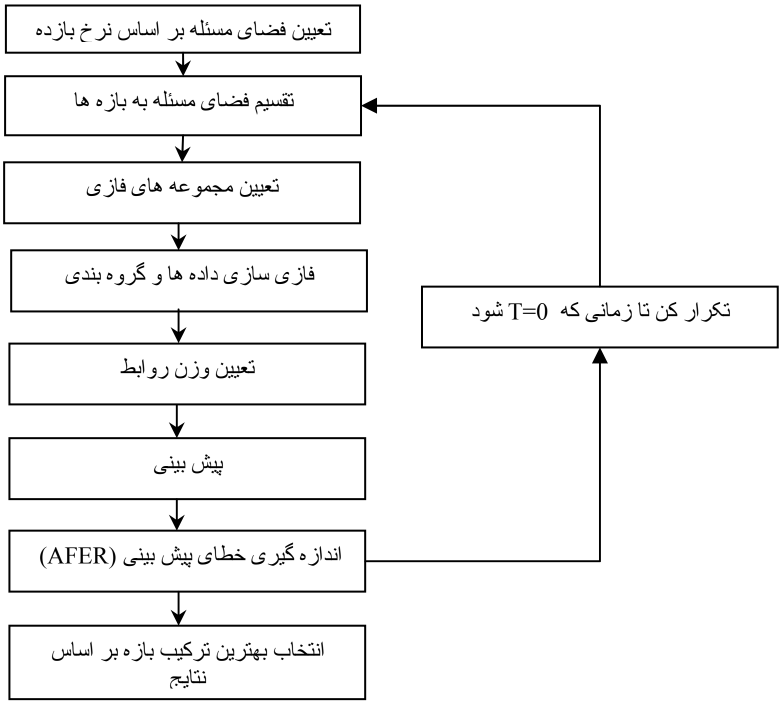
شکل ۱- نمودار قیمت شاخص بورس از ۱۳۸۸/۱/۸ تا ۱۳۸۹/۷/۱(۹۳ هفته)



شکل ۲- نمودار نرخ بازده شاخص بورس از ۱۳۸۸/۱/۸ تا ۱۳۸۹/۷/۱ (۹۳ هفته)

الگوریتم شبیه سازی تبرید

الگوریتم شبیه سازی تبرید، یک الگوریتم فرا ابتکاری برای یافتن نقاط بهینه کلی با استفاده از روابط ریاضی و احتمالی می باشد. به عبارت دیگر این الگوریتم سعی در یافتن تخمین مناسبی از جواب بهینه در فضای کلی جواب دارد. نام این الگوریتم از فرایند تبرید در مهندسی مواد گرفته شده است که شامل فرایند گرم کردن و سپس سرد کردن کنترل شده می باشد تا اندازه ی کریستال ها در ماده افزایش پیدا کند، گرما باعث می شود تا اتم ها از جای خود (بهینه ی محلی) کنده شده و به صورت تصادفی به نقاط پر انرژی تر حرکت کنند، سپس فرایند خنک کردن کنترل شده به اتم ها فرصت می دهد تا ترکیب مناسبتری را نسبت به حالت قبل بیابند (بهینه ی کلی). برای اولین بار این الگوریتم توسط متروپلیس و همکاران (Metropolis et al., 1953) و کیرک پاتریک (Kirkpatrick et al., 1983) معرفی گردید.



شکل ۳- الگوریتم مدل پیشنهادی

۴- روش شناسی پژوهش

تحقیق حاضر تحقیقی از نوع پژوهشی (اکتشافی) می باشد که با استفاده از تکنیک های بهینه سازی و با ابزار الگوریتم های تکاملی و نرم افزار Matlab انجام شده است. به طور کلی می توان تحقیق حاضر را به دو قسمت عمده تقسیم نمود که عبارتند از:

۱) تعریف بازه ها با استفاده از مفهوم نرخ بازده
۲) چگونگی تقسیم بازه ها به نحو مناسب در جهت کاهش خطای پیش بینی

در بخش نخست تحقیق از مفهوم نرخ بازده در بازارهای مالی استفاده شده است، به گونه ای که برای نخستین بار در ادبیات موضوع، تعریف نرخ بازده وارد وارد مدل سازی های سری زمانی فازی شده است. در بخش دوم تحقیق نیز روشی فرا ابتکاری بر مبنای الگوریتم شبیه سازی تبرید استفاده شده است که به عنوان ابزاری مناسب برای تعیین بازه های مناسب پیش بینی مورد استفاده قرار گرفته

است. گفتنی است که مدل ها با استفاده از نرم افزارb Matlab 2008 کد نویسی و اجرا شده است. همچنین برای اجرای مدل از دو جامعه ی آماری استفاده شده است که در ادامه توضیحات مبسوط ارائه شده است.

در این تحقیق برای اجرا و مقایسه ی عملکرد مدل طراحی شده با سایر مدل های موجود از دو جامعه آماری استفاده شده است. از آنجا که این تحقیق مربوط به پیش بینی سری های زمانی است طبعا داده های مورد استفاده از جنس سری زمانی خواهد بود. دو سری زمانی مورد استفاده در این تحقیق عبارتست از : ۱- داده های پذیرش دانشگاه آلاباما از سالهای ۱۹۷۱ الی ۱۹۹۲، این سری زمانی اولین سری زمانی مورد استفاده در این گونه مسائل است که همچنان به عنوان یک مسئله برای مقایسه ی مدل ها مورد استفاده قرار می گیرد که در این تحقیق نیز برای مقایسه مدل با سایر مدل های موجود در ادبیات از این سری زمانی استفاده شده است. ۲- داده های شاخص بورس اوراق بهادار تهران از تاریخ ۱۳۸۸/۱/۸ تا ۱۳۸۹/۷/۱ به صورت هفتگی مورد استفاده قرار گرفته است، این جامعه آماری نیز یک سری زمانی برگرفته از داده های بورسی می باشد که برای تست مدل برروی داده های مالی استفاده شده است.

همانگونه که اشاره شد، داده های مورد استفاده در این تحقیق، داده هایی از نوع سری زمانی هستند. داده های مربوط به پذیرش دانشگاه آلاباما به عنوان یک مرجع مقایسه به شمار می رود که در بسیاری از تحقیقات انجام شده در ادبیات می توان به این داده ها دسترسی داشت. در مرحله ی بعد داده های بورس تهران مورد بررسی قرار می گیرد که در این مورد نیز داده ها از سایت بورس اوراق بهادار تهران استخراج شده است. در نهایت پس از اجرای مدل برروی هریک از داده ها، از شاخص MSE[۱۳] برای سنجش عملکرد مدل و دقت پیش بینی مدل استفاده شده است.

۵- فرضیات پژوهش

فرضیات این تحقیق را می توان به صورت زیر برشمرد:

۱) سری های زمانی فازی در شرایطی بیشتر مورد توجه قرار می گیرند که داده ها به صورت کیفی بوده و یا تعداد داده ها کم باشد.

۲) مدل سری زمانی فازی در این تحقیق از نوع مرتبه اول و مرتبه بالا می باشد.

۳) بخش فازی و توابع عضویت استفاده شده در این تحقیق از مقاله (Chen, 1996) استخراج شده است که یک مقاله مرجع به شمار می رود.

۴) نرخ بازده به صورت نرخ بازده مرکب مورد استفاده قرار می گیرد.

۵) رفتار قیمت سهام (شاخص) با استفاده از توزیع لاگ نرمال قابل تعریف است.

۶- نتایج پژوهش

در این بخش مدل حاصله برروی داده های دانشگاه آلاباما که به عنوان مرجعی برای این مدل ها محسوب می شود اجرا شده و نتایج با مدل های پیشین مقایسه می گردد و در قسمت دوم، مدل بر روی داده های شاخص بورس تهران اجرا شده و نتایج در شرایط مختلف تحلیل می گردد.

۶-۱- اجرای مدل برروی داده های پذیرش دانشگاه آلاباما

جدول۱ شامل نتایج مدل های مرتبه اول برای پیش بینی داده های دانشگاه آلاباما است. قابل توجه است که در مدلهای کوو (Kuo et al., 2009) چن و چانگ (Chen and Chung, 2006a) و RBFTS از ۱۴ بازه برای پیش بینی داده ها استفاده شده است. با توجه به نتایج بدست آمده و مقایسه ی خطای پیش بینی مدل ها (MSE) همانطور که مشهود است، مدل پیشنهادی RBFTS با داشتن کمترین خطا در پیش بینی، عملکردی فراتر از سایر مدل ها به نمایش گذاشته اند.

جدول ۱- مقایسه نتایج پیش بینی مدل‌های مرتبه اول

Year	Actual	(Chen, 2002)	(Huarng, 2001b)	(Kuo et al., 2009)	(Chen and Chung, 2006a)	RBFTS
1971	13055					
1972	13563	14 000	14 000	13 555	13 714	
1973	13867	14 000	14 000	13 994	13 714	13853
1974	14696	14 000	14 000	14 711	14 880	14812
1975	15460	15 500	15 500	15 344	15 467	15383
1976	15311	16 000	15 500	15 411	15 172	15270
1977	15603	16 000	16 000	15 411	15 467	15457
1978	15861	16 000	16 000	15 411	15 861	15707
1979	16807	16 000	16 000	16 816	16 831	16807
1980	16919	16 833	17 500	17 140	17 106	16902
1981	16388	16 833	16 000	16 464	16 380	16468
1982	15433	16 833	16 000	15 505	15 464	15438
1983	15497	16 000	16 000	15 411	15 172	15520
1984	15145	16 000	15 500	15 411	15 172	15084
1985	15163	16 000	16 000	15 344	15 467	15289
1986	15984	16 000	16 000	16 018	15 467	16024
1987	16859	16 000	16 000	16 816	16 831	16982
1988	18150	16 833	17 500	18 060	18 055	17912
1989	18970	19 000	19 000	19 014	18 998	18999
1990	19328	19 000	19 000	19 340	19 300	19376
1991	19337	19 000	19 500	19 340	19 149	19457
1992	18876	19 000	19 000	19 014	19 149	18547
	MSE	407507	226611	22965	35324	9692

در مرحله بعد، مدل های با تعداد مرتبه ی بالا با یکدیگر مقایسه شده اند. جدول ۲ شامل خطای پیش بینی هریک از مدل ها در تعداد مرتبه ی مختلف می باشد. قابل ذکر است که در تمامی مدل ها از ۷ بازه برای پیش بینی استفاده شده است. در این بخش نیز مدل RBFTS در بین سایر مدل ها دارای بهترین عملکرد(کمترین خطای پیش بینی MSE) است.

جدول ۲- مقایسه ی خطای پیش بینی مدل ها بر اساس تعداد ۷ بازه و با مرتبه ۲ تا ۹

order	(Chen, 2002)	(Chen and Chung, 2006a)	(Singh, 2007)	(Kuo et al., 2009)	RBFTS
2	89093	67834		67123	7837
3	86694	31123	133700	31644	3606
4	89376	32009		23271	3821
5	94539	24948		23534	2697
6	98215	26980		23671	2882
7	104056	26969		20651	2708
8	102179	22387		17106	2397
9	102789	18734		17971	2359

همانگونه که مشاهده می شود، مدل RBFTS در این بخش در شرایط مختلف برروی داده های دانشگاه آلاباما مورد آزمون قرار گرفت که در نهایت نتایج حاصله نشان داد که مدل RBFTS با بهره بردن از تعریف نرخ بازده در تعیین بازه های پیش بینی، عملکرد بسیار بهتری نسبت به مدل های موجود در ادبیات داشته است.

۶-۲- اجرای مدل بر روی شاخص بورس اوراق بهادار تهران

۶-۲-۱- مبانی شاخص بورس تهران

بورس اوراق بهادار تهران از فروردین ماه ۱۳۶۹ اقدام به محاسبه و انتشار شاخص قیمت خود با نام تپیکس نمود. این شاخص، نماگر تغییرات قیمت کل بازار است و به صورت میانگین وزنی و با استفاده از فرمول زیر محاسبه می‌گردد.

$$\mathrm{TEPIX} = \frac{\sum_{i=1}^{n} p_{it} q_{it}}{D_t} \times 100 \qquad (۶)$$

Pit = قیمت شرکت iام در زمان t

qit = تعداد سهام منتشره شرکت iام در زمان t

Dt = عدد پایه در زمان t که در زمان مبداء برابر $\sum p_{io} q_{io}$ بوده است

Pio = قیمت شرکت iام در زمان مبدأ

qio = تعداد سهام منتشره شرکت iام در زمان مبدأ

n = تعداد شرکت‌های مشمول شاخص

شاخص قیمت بورس تهران، سهام تمام شرکت‌های پذیرفته شده در بورس را در برمی‌گیرد و در صورتی‌که نماد شرکتی بسته باشد یا برای مدتی معامله نشود، قیمت آخرین معامله آن در شاخص لحاظ می‌گردد. همان‌گونه که از فرمول مشخص است، تعداد سهام منتشره ی شرکت‌ها، معیار وزن‌دهی در شاخص مزبور است.

۶-۲-۲- نتایج اجرای مدل

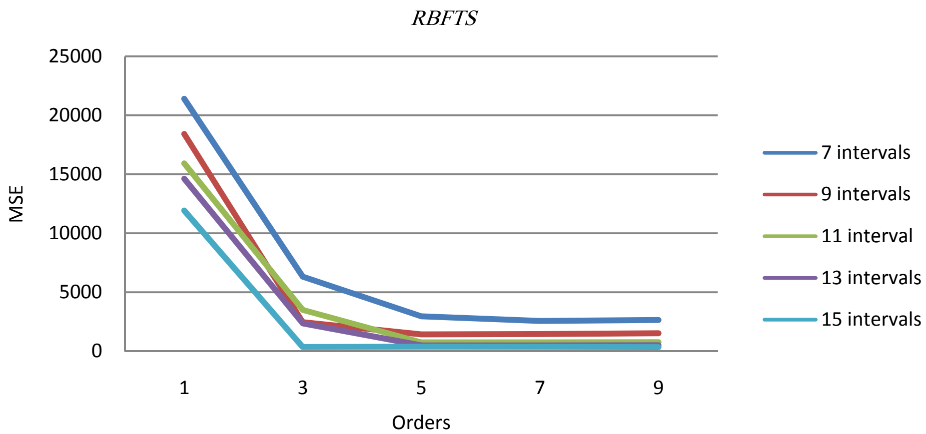
برای اجرای مدل، مقادیر شاخص بورس تهران از تاریخ ۱۳۸۸/۱/۸ تا ۱۳۸۹/۷/۱ به صورت هفتگی در نظر گرفته شده است که شامل ۹۳ هفته است ، در این بخش برای پیش بینی داده های شاخص بورس از مدل RBFTS استفاده شده است به گونه ای ۳۰ هفته ی اول به عنوان داده های اولیه به مدل داده شد و سپس در هر هفته شاخص هفته ی آینده پیش بینی شده است. مدل RBFTS با در نظر گرفتن تعداد بازه و تعداد مرتبه ی مختلف برروی داده ها ی شاخص بازار بورس تهران اجرا شد که نتایج بدست آمده در جدول ۳ است.

جدول ۳- نتایج پیش بینی مدل RBFTS در تعداد بازه و مرتبه ی مختلف

interval / order	7	9	11	13	15
1	21406.99	18421.13	15925.54	14631.69	11933.34
3	6317.507	2454.326	3502.997	2369.02	360.0131
5	2949.444	1424.938	753.854	497.0049	389.7111
7	2560.497	1450.791	753.6775	498.7177	379.5185
9	2644.038	1520.59	765.2533	507.192	347.0526

شکل ۴ نمایانگر تغییرات خطای پیش بینی در افزایش مرتبه ی مدل است. آن چنان که مشاهده می شود در مدل RBFTS در تعداد بازه ی ثابت، افزایش تعداد مرتبه ی مدل از ۱ به ۳، به طور کاملا چشمگیری دقت مدل را افزایش می دهد اما افزایش مرتبه از ۵ تا ۹ تاثیر چندانی بر عملکرد مدل ندارد لذا می توان در یافت که انتخاب مدل با تعداد مرتبه ی ۳ یا ۵ برای پیش بینی داده های شاخص بورس مناسب است. همانطور که در شکل های اشاره شده مشخص است افزایش تعداد مرتبه ی مدل

می تواند به افزایش کارایی مدل کمک شایانی بنماید به گونه ای که در جدول ۴ میزان بهبود تابع هدف در هر مدل و تحت تعداد بازه های مشخص مورد بررسی قرار گرفته است. همانگونه که در جدول ۴ مشاهده می شود با افزایش تعداد بازه ها، تاثیر افزایش مرتبه ی مدل در کارایی مدل بیشتر می شود. به گونه ای که در مدل با ۷ بازه، افزایش تعداد مرتبه مدل موجب بهبود ۸۷ درصدی مدل می شود و در مدل با ۱۵ بازه این مقدار برابر ٪۹۷ می شود. البته باید به این مسئله توجه داشت که اگر چه افزایش مرتبه ی مدل موجب افزایش دقت پیش بینی می شود، اما این امر موجب افزایش بیش از حدحساسیت مدل شده و مدل در مواجهه با داده های جدید و رفتارهای جدید قیمت، قابلیت خود را از داست داده و امکان انجام پیش بینی را نخواهد داشت. لذا با این تفاسیر می توان دریافت که افزایش حساسیت لزوما موجب افزایش کارایی قیمت در تمامی داده ها نخواهد شد. در نتیجه می توان گفت در این مسئله تعداد مرتبه ی ۳ تا ۵ برای پیش بینی داده های بورسی مناسب به نظر می رسد.

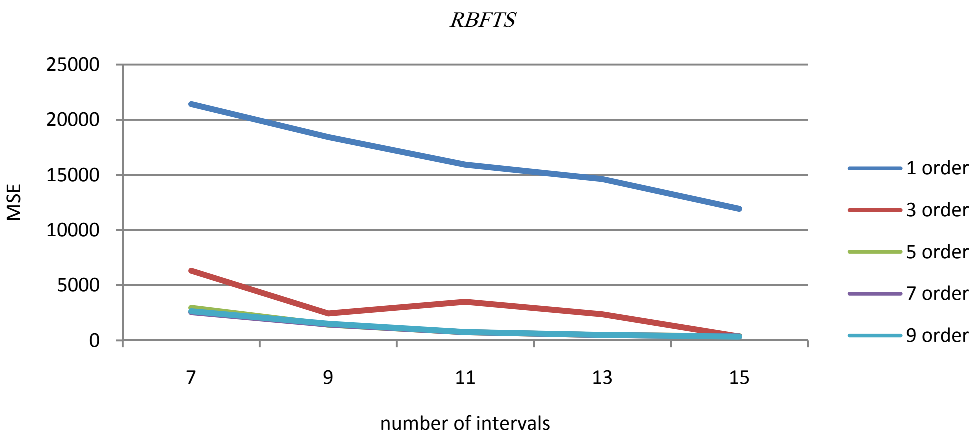


شکل ۴- تغییرات خطای پیش بینی در اثر تغییر مرتبه ی مدل (RBFTS)

جدول ۴- اثر تغییر مرتبه مدل در تابع هدف

Model / Intervals	RBFTS
7	65%.87
9	75%.91
11	19%.95
13	53%.96
15	09%.97

شکل ۵ اثر افزایش تعداد بازه ها در میزان دقت پیش بینی مورد بررسی قرار گرفته است. و جدول ۵ نیز نشان دهنده ی میزان بهبود تابع هدف در افزایش تعداد بازه ها با تعداد مرتبه ی مشخص است.



شکل ۵- تغییرات خطای پیش بینی در اثر تغییر تعداد بازه ها (RBFTS)

جدول ۵- اثر تغییرات تعداد بازه در تابع هدف

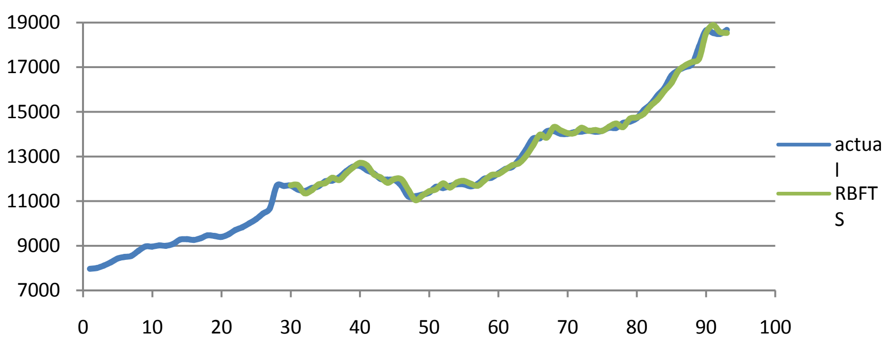
Model / Order	RBFTS
1	60%.77
3	38%.82
5	37%.87
7	08%.83
9	75%.83

با مشاهده ی شکل ۵ می توان دریافت که افزایش تعداد بازه ها ی پیش بینی در تمامی تعداد مرتبه های مدل موجب بهبود کار کرد مدل می شود. در جدول ۵ نیز میزان بهبود کارایی مدل در ازای افزایش بازه ها در تعداد مرتبه های مختلف مورد بررسی واقع شده است. همانگونه که مشاهده می شود تقریبا در تمامی تعداد مرتبه ها، افزایش بازه های مدل موجب افزایش ۸۰ درصدی کارایی مدل می شود.

همانگونه که از مقایسه دو جدول ۴ و جدول ۵ بر می آید، افزایش تعداد مرتبه ی مدل در RBFTS تاثیر بیشتری در مقایسه با افزایش تعداد بازه ها دارد و در نهایت می توان استنباط نمود که انتخاب

این مشخصات موجب می شود تا مدلی کارا با قابلیت اطمینان بالایی در اختیار داشته باشیم، از آنجایی که داده های بورسی همواره دارای تنش و آشفتگی خاصی می باشد لذا همواره باید مدل را به گونه ای در نظر گرفت که قابلیت مواجهه با این تنش ها و آشوب ها را داشته باشد و کارایی مدل در این شرایط صدمه نبیند. لذا برای نیل به چنین هدفی از تحلیل های انجام شده در فوق می توان استنباط نمود که تعداد مرتبه ی ۳ تا ۵ و تعداد بازه ی ۱۵ برای پیش بینی شاخص بورس تهران با استفاده از مدل RBFTS مناسب است و در این شرایط می توان مدلی کارا با قابلیت اطمینان مناسب در اختیار داشت.

در ادامه در شکل ۶ نمونه ای از پیش بینی مدل RBFTS برروی داده های شاخص بورس نشان داده شده است، گفتنی است که این نتیجه از مدل RBFTS با ۷ بازه و مرتبه ی اول حاصل شده است که در حقیقت پایه ای ترین حالت پیش بینی برای این مدل به شمار می رود. مقدار R^2 محاسبه شده برای این مدل برابر ۰/۹۷ است که می توان نتیجه گرفت مدل RBFTS با داشتن R^2 برابر ۰/۹۷، در پایه ای ترین حالت نتایج قابل قبولی را از خود ارائه داده است.



شکل ۶- نمودار پیش بینی مدل RBFTS مرتبه اول و ۷ بازه

۷- نتیجه‌گیری و بحث

هدف از انجام این تحقیق ارائه مدلی برای پیش بینی سری های زمانی مالی با کمک مفهوم سری های زمانی فازی می باشد. در این تحقیق سعی شده است که با افزودن مفاهیم مالی به سری های زمانی فازی علاوه بر افزایش کارایی مدل، نوع جدید ی از این مدل ها را با هدف پیش بینی سری های زمانی مالی ارائه داد. همانگونه که در ابتدای تحقیق ذکر شد، قریب به اتفاق مدل های سری زمانی فازی بر اساس مدل پایه ی ارائه شده توسط چن (Chen, 1996) طراحی و بسط داده شده اند در تحقیقات مختلف گذشته همواره دو عامل اصلی مدل مد نظر قرار گرفته است: ۱- چگونگی تقسیم بازه

ها به نحو مناسب در جهت کاهش خطای پیش بینی ۲- تنظیم توابع عضویت متغیرهای زبانی. در این مدلها از سطوح داده ها برای تعریف مجموعه های فازی استفاده شده است اما در این مقاله برای نخستین بار نوع جدیدی از سری های زمانی فازی با تعریف بازه بندی بر اساس نرخ بازده معرفی گردید و در ادامه برای بهبود دقت بازه بندی از الگوریتم شبیه سازی تبرید استفاده شد. در نهایت داده های دانشگاه آلاباما به عنوان مرجعی برای سنجش اعتبار مدل استفاده شد که نتایج به دست آمده نشان دهنده ی برتری مدل RBFTS در مقایسه با سایر مدل ها بود. در ادامه به عنوان یک مورد اجرایی، داده های شاخص بورس تهران مورد استفاده قرار گرفت و نتایج حاصل تحت شرایط مختلف مورد بررسی قرار گرفت. با توجه به تفاوت ماهیت داده های بورسی نسبت به داده های دانشگاه آلاباما، افزایش تعداد مرتبه ی مدل در پیش بینی داده های بورسی همواره موجب افزایش کارایی مدل نمی شود. داده های بورسی دارای یک بی نظمی و اغتشاش خاصی می باشند که با بالا رفتن حساسیت مدل به تدریج از قابلیت اطمینان مدل کاسته می شود و مدل پیش بینی در هنگا بروز اغتشاشات و بی نظمی ها در داده ها نمی تواند عملکرد مناسبی از خود بروز دهد. لذا با انجام تحلیلهای مختلف برروی نتایج حاصله برروی داده های بورس، مدل سری زمانی RBFTS با تعداد مرتبه ی ۳ یا ۵ و تعداد بازه ی ۱۵ برای پیش بینی شاخص بورس تهران پیش نهاد شد. گفتنی است که در این بخش نیز شاهد پیش بینی و نتایج مطلوب مدل ارائه شده بودیم.

مدل RBFTS یک مدل تک فاکتوره است برای تحقیقات آتی می توان این مدل را در مسائل دو فاکتوره اجرا نمود که به عنوان مثال در مورد بازار سرمایه می توان بی نظمی بازار را به صورت فاکتور دوم فازی ورد مدل کرد. که در این صورت مدل با در نظر گرفتن روند قیمتی و بی نظمی بازار تحلیل دقیقتری ازشرایط بازار را ارائه خواهد نمود.. همچنین حجم معاملات بازار را نیز به عنوان فاکتور دوم برای مدل در نظر گرفت که در این زمینه تحقیقات وسیعی در حال انجام است.

فهرست منابع

1) CHEN, S. M. 1996. Forecasting enrollments based on fuzzy time series. Fuzzy Sets and Systems, 81, 311-319.
2) CHEN, S. M. 2002. Forecasting enrollments based on high-order fuzzy time series. Cybernetics and Systems: An International Journal, 33, 1-16.
3) CHEN, S. M. & CHUNG, N. Y. 2006a. Forecasting enrollments of students by using fuzzy time series and genetic algorithms. International Journal of Information and Management Sciences, 17, 1-17.
4) CHEN, S. M. & CHUNG, N. Y. 2006b. Forecasting enrollments using high-order fuzzy time series and genetic algorithms: Research Articles. International Journal of Information and Management Sciences, 21, 485-501.

5) HUARNG, K. 2001a. Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets and Systems, 123, 387-394.
6) HUARNG, K. 2001b. Heuristic models of fuzzy time series for forecasting. Fuzzy Sets and Systems, 123, 369-386.
7) HULL, J. C. 2002. Fundamentals of futures and option markets, New Jersey, Prentice Hall.
8) HWANG, J., CHEN, S. M. & LEE, C. H. 1998. Handling forecasting problems using fuzzy time series. Fuzzy sets and Systems, 100, 217-228.
9) JILANI, T. A. Year. Fuzzy metric approach for fuzzy time series forecasting based on frequency density based partitioning. In: world academy of science, engineering and technology conf, 2007.
10) KIRKPATRICK, S., GELLAT, C. D. & VECCHI, M. P. 1983. Optimization by simulated annealing. Science, 220, 671-680.
11) KUO, H., HORNG, H., KAO, S. J., LIN, T. W., LEE, T. L. & PAN, C. L. 2009. An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. Expert Systems with Applications, 36, 6108-6117.
12) KUO, H., HORNG, H., KAO, S. J., LIN, T. W., LEE, T. L. & PAN, C. L. 2010. Forecasting TAIFEX based on fuzzy time series and particle swarm optimization. Expert Systems with Applications, 37, 1494-1502.
13) LEE, L. W., WANG, H. F. & CHEN, S. M. 2008. Temperature prediction and TAIFEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques. Expert Systems with Applications, 34, 328-336.
14) LEE, L. W., WANG, L. H. & CHEN, S. M. 2007. Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms. Expert Systems with Applications, 33, 539-550.
15) MELIKE, S. & KONSTSNTIN, Y. D. 2004. Forecasting enrollment model based on first-order fuzzy time series. In Proceedings of international conference on computational intelligence. Istanbul, Turkey.
16) METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H. & TELLER, E. 1953. Equations of state calculations by fast computing machines. Journal of Chemical Physics, 21, 1087-1092.
17) PARK, J. I., LEE, D. J., SONG, C. K. & CHUN, M. G. 2010. TAIFEX and KOSPI 200 forecasting based on two-factors high-order fuzzy time series and particle swarm optimization. Expert Systems with Applications, 37, 959-967.
18) SINGH, S. R. 2007. A simple method of forecasting based on fuzzy time series. Applied Mathematics and Computation, 186, 330-339.
19) SONG, Q. & CHISSOM, B. S. 1993a. Forecasting enrollments with fuzzy time series -- Part I. Fuzzy Sets and Systems, 54, 1-9.
20) SONG, Q. & CHISSOM, B. S. 1993b. Fuzzy time series and its models. Fuzzy Sets and Systems, 54, 269-277.
21) SONG, Q. & CHISSOM, B. S. 1994. Forecasting enrollments with fuzzy time series-part II. Fuzzy Sets and Systems, 62, 1-8.

22) WANG, C. H. & HSU, L. C. 2008. Constructing and applying an improved fuzzy time series model: Taking the tourism industry for example. Expert Systems with Applications, 34, 2732–2738
23) ZADEH, L. A. 1965. Fuzzy sets. Information and Control, 8, 338-353

یادداشت‌ها

[1] Universe of discourse
[2] TAIFEX (TAIWAN FUTURE EXCHANGE)
[3] High order fuzzy time series
[4] composition operation
[5] KOPSI 200
[6] Particle Swarm Optimization
[7] Fuzzy logical relationship
[8] Left hand side
[9] Right hand side
[10] Rate Of Return
[11] اصطلاحا اگر لگاریتم یک متغیر تصادفی دارای توزیع نرمال باشد تابع توزیع متغیر را لاگ نرمال می نامند
[12] Continuously compounding Rate of Return Per Annum
[13] Mean Square Error