

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/268687377>

Evaluation and Comparison of Forecast Performance of Linear and Non-linear Methods for Daily Returns of Tehran Stock Exchange

Article · October 2015

CITATIONS

0

READS

28

2 authors:



[Mansour Zarra-Nezhad](#)

Shahid Chamran University of Ahvaz

77 PUBLICATIONS 34 CITATIONS

[SEE PROFILE](#)



[Ali Raofi](#)

Allameh Tabatabai University

11 PUBLICATIONS 5 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Project

A New Approach to Portfolio Optimization [View project](#)



Information

- Authors GuideLine
- subscription Guideline
- Article Submission
- Contact Us
- About the Journal
- Editorial Board

Language

English

Notifications

- View
- Subscribe / Unsubscribe

User



[Home /](#)

Issue Information

Vol. 22, No. 9,(1394): اقتصاد پولی و مالی

Article Title

Evaluation and Comparison of Forecast Performance of Linear and Non-linear Methods for Daily Returns of Tehran Stock Exchange

Mansour Zarra Nezhad; Ali Raofi

pages:
1-28

Abstract

Forecasting economic and financial variables is of high significance to economic policymakers and investors; however, it is a difficult and complicated task due to the volatile and complex nature of such data.

Numerous studies have been conducted concerning different methods of forecasting macroeconomic and financial variables so far. Although frequent and sophisticated methods have been applied to forecast such variables, the nature of data under consideration has not sufficiently been taken into consideration. In terms of complexity, type and nature of data might impact the accuracy of forecasting models. In other words, linear or non-linear behaviors of data can be effective in the selection of the forecasting model. Recent research indicates that a better understanding of the generating process of variable data (linear/non-linear) leads to easier and more accurate forecasts. If, for instance, the variable follows a linear behavior, linear models, such as ARMA, will produce more acceptable accuracy. On the contrary, using more complex modeling methods such as ANN and ANFIS is more justifiable when the variable behavior is non-linear and chaotic. Using complex models for a variable with linear behavior might lead to excessive model dependency on unnecessary volatility and, in turn, reduced forecast accuracy. This paper focuses on the study of linearity, non-linearity, and/or chaotic nature of TEPIX from March 25th, 2009 to October 15th, 2011 (625 observations) using BDS test. This test was administered in three stages to determine linearity, non-linearity, or "chaotic-ness" of TEPIX: First, the test was administered on daily stock market index return; second, the test was applied to ARMA model residuals; and finally, the test was carried out for ANFIS, GARCH, and ANN residuals. The results suggest that TEPIX return variable follows a non-linear behavior. Therefore, it is expected that non-linear models are better capable of forecasting this variable. Then, different prediction techniques in ARMA linear model were compared with those of non-linear models including ANN, ANFIS, and GARCH. According to the evaluation criteria at hand (RMSE, MAE, U-Thiel, and MAPE), the accuracy of forecasts was compared. The results show that non-linear models enjoy better performance than ARMA model regarding all the criteria above. In addition, among non-linear models, ANFIS model displays the best performance in forecasting daily stock market index return. Taking the non-linear nature of data used into account, such results were predicible.

Keywords: Adaptive Neuro-Fuzzy Inference System (ANFIS), Neural Network, GARCH model, Non-linear Models, Chaos Theory, Stock Returns

JEL: G10, C52, C45, C22

key words

Adaptive Neuro-Fuzzy Inference System (ANFIS); Neural Network; GARCH model; Non-linear Models; Chaos Theory; Stock Returns

References



XML version



Full Text:

PDF

Views:

19

پیش بینی بازار روزانه بورس اوراق بهادار تهران: ارزیابی و مقایسه روش های خطی و غیر خطی

منصور زراءنژاد^۱

استاد گروه اقتصاد دانشگاه شهید چمران اهواز

علی رئوفی^۲

دانشجوی دکترا اقتصاد دانشگاه علامه طباطبایی

تاریخ پذیرش: ۹۲/۱۱/۲۱

تاریخ دریافت: ۱۳۹۲/۴/۱۹

چکیده

پیش بینی متغیرهای اقتصادی و مالی اهمیت فراوانی برای سیاست گذاران و سرمایه گذاران اقتصادی کشورها دارد، با این حال پیش بینی این متغیرها با توجه به ماهیت پرنوسان و پیچیده ای که دارند بسیار دشوار است. نوع و ماهیت داده از حیث پیچیدگی می تواند بر دقت پیش بینی مدل ها اثر گذار باشد؛ به عبارتی رفتار خطی و یا غیرخطی متغیرها می تواند بر انتخاب مدل پیش بینی اثر گذار باشد. تحقیقات اخیر نشان می دهد اگر بتوان فرآیند مولد داده های یک متغیر (خطی یا غیرخطی) را به دست آورد پیش بینی آن متغیر راحت تر و با خطای کمتری امکان پذیر خواهد بود.

در این مقاله ابتدا با استفاده از آزمون براک-دیگرت و شاینکمن (BDS)، به بررسی خطی یا غیرخطی بودن و سپس آشوبناک بودن بازده شاخص کل بورس اوراق بهادار تهران (TEPIX) طی بازه زمانی ۱۳۸۸/۰۱/۰۵ تا ۱۳۹۰/۰۷/۲۳ (۶۲۵ مشاهده) پرداخته شده است. نتایج آزمون نشان می دهد که این متغیر از یک رفتار غیرخطی تبعیت می کند. سپس با استفاده از تکنیک های مختلف پیش بینی، مدل های خطی و غیرخطی ARIMA، GARCH، ANN و ANFIS برآورد شدند و با استفاده از معیارهای دقت پیش بینی مانند RMSE، MAE، U-Thiel، MAPE، مدل ها مورد ارزیابی و مقایسه قرار گرفتند. نتایج نشان می دهد که مدل های غیرخطی نسبت به مدل ARIMA از عملکرد بهتری برخوردار بودند و در بین مدل های

غیرخطی نیز مدل ANFIS بهترین عملکرد را در پیش‌بینی بازده روزانه شاخص سهام دارا بود. در ادامه با استفاده از آماره‌ی مورگان-گرنجر-نیوبلد (MGN) معنی‌داری تفاوت دقت پیش‌بینی مدل‌های غیرخطی با مدل‌های خطی مورد آزمون قرار گرفت که نتایج نشان‌دهنده تفاوت معنی‌دار در پیش‌بینی روش‌های خطی و غیرخطی بود.

واژه‌های کلیدی: شبکه عصبی فازی ANFIS، شبکه عصبی، مدل گارچ، مدل‌های غیرخطی، نظریه آشوب، بازده سهام
طبقه بندی JEL: C22, C45, C52, G10

۱- مقدمه:

پیش‌بینی آینده در عرصه تحلیل‌های مالی و بازار سرمایه یکی از مهم‌ترین مسائل مورد بحث در بین متخصصین این حوزه است. پیش‌بینی شاخص سهام^۱ کار بسیار پیچیده‌ای است، زیرا عوامل بسیاری از جمله رویدادهای سیاسی، شرایط اقتصادی، انتظارات معامله‌گران و سایر عوامل محیطی بر روی آن اثرگذارند. علاوه بر این طبیعت شاخص سهام به طور کلی بسیار پرنوسان^۲، پویا، غیر خطی^۳، پیچیده و آشوبناک^۴ است (یودانگ و لینان، ۲۰۰۹).

در سالیان اخیر پیشرفت‌های چشم‌گیری در زمینه پیش‌بینی سری‌های زمانی صورت گرفته است. چارچوب مانایی خطی مدل‌های خود بازگشت میانگین متحرک انباشته^۶ (ARIMA) و خود بازگشت برداری^۷ (VAR) که از تکانه‌های IID^۸ استخراج و برای سال‌های متمادی اساس مدل‌سازی اقتصادسنجی محسوب شده‌اند، جای خود را به مدل‌هایی داده‌اند که ویژگی‌های

۱- Stock Index

۲- Noisy

۳- Nonlinear

۴- Chaotic

۵- Yudong & Lenan

۶- Auto Regressive Integrated Moving Average

۷- Vector Auto Regressive

۸- Independent and Identically Distributed

نامانایی و غیرخطی بسیاری از سری‌های زمانی اقتصادی و مالی را مورد بحث و بررسی قرار می‌دهند.

تحقیقات اخیر نشان می‌دهد اگر بتوان فرآیند مولد داده‌های یک متغیر (خطی یا غیرخطی) را به دست آورد پیش‌بینی آن متغیر راحت‌تر و با خطای کمتری امکان‌پذیر خواهد بود (مشیری و مروت، ۱۳۸۵). اگر چه مدل‌های خطی پیشرفته پیش‌بینی‌های مناسبی در دوره‌های زمانی کوتاه‌مدت و میان‌مدت دارند، اما بررسی‌ها در بازار سرمایه نشان داده است که رفتار سهام از یک الگوی غیرخطی تبعیت می‌کند و الگوهای خطی تنها بخشی از رفتار سهام در بازار را نشان می‌دهند (چاوشی، ۱۳۸۰). در این صورت وجود یک سیستم پویای غیرخطی در رفتار متغیر، مدل‌های خطی را عملاً دچار ابهام خواهد کرد؛ لذا پیش‌بینی داده‌هایی که از سیستم‌های غیرخطی پیروی می‌کنند، نیازمند ابزارهای هوشمند و پیشرفته‌تری مانند روش‌های محاسبات نرم^۱ است. این روش‌ها به عنوان دسته‌ای از سیستم‌های هوشمند، می‌تواند رابطه غیرخطی بین ورودی و خروجی‌ها را تشخیص دهند. در این مطالعه علاوه بر اثبات رفتار غیرخطی و همچنین آشوبناک بازده روزانه شاخص سهام بورس اوراق بهادار تهران با استفاده از آزمون BDS، مدل خطی ARIMA با روش‌های غیرخطی شبکه عصبی (ANN)^۲، شبکه عصبی فازی ANFIS^۳ و خود توضیحی واریانس ناهمسان شرطی تعمیم‌یافته (GARCH)^۴ مقایسه شده است. در ادامه با توجه به معیارهای ارزیابی، دقت پیش‌بینی مدل‌های فوق با هم مقایسه شده است. در نهایت نیز با انجام آزمون مرگان-گرنجر-نیولبد^۵ (MGN) نشان داده شده که آیا تفاوت معنی‌داری بین خطای پیش‌بینی روش‌های خطی و غیرخطی وجود دارد یا خیر.

بخش‌های بعدی مقاله به شرح زیر است:

بخش دوم به ادبیات تحقیق پرداخته شده و بخش سوم مدل‌های پیش‌بینی معرفی شده‌اند. در

۱- Soft Computing

۲- Adaptive Neural Network

۳- Adaptive Neuro- Fuzzy Inference System

۴- Generalized Auto Regression Conditionally Heteroskedasticity

۵- Morgan-Granger-Newbold Test

بخش چهارم به معرفی و بررسی آماره‌های توصیفی متغیر مورد استفاده (بازده شاخص کل بورس اوراق بهادار تهران) پرداخته شد، همچنین در بخش پنجم ماهیت خطی و یا غیر خطی متغیر بازده روزانه شاخص کل با استفاده از آزمون BDS بررسی گردید. نتایج پیش بینی مدل‌های مختلف در بخش ششم قابل مشاهده است. بخش هفتم مربوط به آشنایی با انواع معیارهای ارزیابی کارایی دقت پیش‌بینی است. در بخش هشتم نتایج پیش‌بینی مدل‌های مختلف آورده شده است. بخش نهم به بررسی معنی‌داری اختلاف در پیش‌بینی‌ها با استفاده از آزمون MGN اختصاص دارد و در نهایت بخش دهم نتایج تحقیق بیان شده است.

۲- ادبیات تحقیق

مطالعات گسترده‌ای پیرامون پیش‌بینی بازده سهام صورت گرفته است. رویکرد عمده برای پیش‌بینی قیمت سهام، روش‌های آماری مبتنی بر سری‌های زمانی^۱ مانند: میانگین متحرک^۲، هموارسازی^۳، آریما^۴، رگرسیون^۵ و یا ترکیبی از آنها است. این روش‌ها بنا به ماهیت خود، در محیط‌های با تغییرات کم با تقریب خوبی قادر به پیش‌بینی هستند اما در مواردی چون پیش‌بینی بازار سهام که شرایط محیطی همواره در حال تغییر است، نمی‌توانند تقریب خوبی از تغییرات محیطی را تخمین بزنند. به این ترتیب نیاز به استفاده از ابزارها و مدل‌های نوین جهت پیش‌بینی ضروری می‌گردد؛ مدل‌هایی که بتوانند رفتار غیر خطی و گاه آشوبناک این سری‌ها را در نظر بگیرند. اگر چه سالیان زیادی نیست که مدل‌های غیر خطی و همچنین تکنیک‌های پیشرفته مورد استفاده قرار می‌گیرند، ولی در همین مدت زمان کم توانسته‌اند جایگاه خود را در علوم مختلف به ویژه اقتصاد باز کنند. بازارهای مالی نیز از این قاعده مستثنی نبوده و متخصصین سیستم‌های غیر خطی سعی در توضیح رفتار قیمت و بازده سهام و پیش‌بینی آن از طریق روش‌های پیشرفته

۱- Time series

۲- Moving average

۳- Auto regressive

۴- Auto regressive integrated Moving average

۵- Regression

غیرخطی کرده‌اند که در این میان محاسبات نرم جایگاه ویژه‌ای دارد.

دونیس و جلیلف^۱ (۲۰۰۱) با استفاده از مدل شبکه عصبی به برآورد و پیش‌بینی چهار شاخص مهم بازار سهام یعنی S&P500، FTSE100، EUROSTOXX50 و NIKKEI225 پرداخته‌اند. آن‌ها برای تخمین مدل از داده‌های روزانه برای پیش‌بینی خارج از نمونه^۲ استفاده کردند. سپس آن‌ها نتایج پیش‌بینی مدل را با پیش‌بینی روش‌های تجاری مانند خرید و نگهداری، انتظارات تطبیقی ساده^۳ و سبد میانگین متحرک^۴ (BMA) با استفاده از معیارهای مختلف مانند نسبت شارپ و آماره t مربوط به سود سهام مقایسه نموده و برتری مدل شبکه عصبی را نشان داده‌اند.

کاناس و یانوپولوس^۵ (۲۰۰۱) با استفاده از روش‌های خطی و غیر خطی اقدام به پیش‌بینی بازده ماهانه سهام بورس نیویورک کردند. نتایج آن‌ها نشان داد که مدل غیر خطی نسبت به مدل خطی از عملکرد بهتری برخوردار است.

مک‌میلان^۶ (۲۰۰۱) در مطالعه‌ای به بررسی رابطه غیر خطی بازده سهام و متغیرهایی مانند نرخ بهره و سود و نسبت تقسیم سود پرداخته است. نتایج وی وجود رابطه غیر خطی متغیرهای فوق با بازده سهام را تایید می‌کند.

ایم^۷ (۲۰۰۲) با استفاده از مدل شبکه عصبی بازده شاخص روزانه سهام برزیل را پیش‌بینی نمود و نتایج پیش‌بینی را با استفاده از معیارهای MAE، RMSE و آزمون چونگ و هندری^۸ با نتایج پیش‌بینی مدل‌های ARMA و GARCH و ساختاری مقایسه نموده و برتری مدل شبکه عصبی مصنوعی را نشان داده است.

۱- Dunis and Jalilov

۲- Out of Sample

۳- Naïve Adaptive Expectation

۴- Basket of Moving Average

۵- Kanas and Yannopoulos

۶- Mc Millan

۷- Yim

۸- Chong and Hendry encompassing test

کاریل و دیگران^۱ (۲۰۰۵) در مطالعه خود به مقایسه و بررسی مدل‌های خطی پیش‌بینی سهام (مدل فاما و فرنچ^۲) و مدل‌های غیر خطی پیش‌بینی (مدل شبکه‌های عصبی و الگوریتم ژنتیک) پرداختند که نتایج بیانگر تفاوت معنی‌دار بین پیش‌بینی مدل‌های خطی و غیر خطی بود و مدل‌های غیر خطی از مدل‌های خطی عملکرد بهتری داشتند.

عباسی و ابوییک^۳ (۲۰۰۸) به بررسی روند قیمت سهام شرکت ایران خودرو در بورس اوراق بهادار با استفاده از روش ANFIS پرداختند. یافته‌های حقیق مذکور نشان می‌دهند که روند قیمت سهام را می‌توان با مقدار خطای اندکی پیش‌بینی کرد.

از جمله مطالعات داخلی در این زمینه می‌توان به مطالعه مشیری و مروت (۱۳۸۵) اشاره کرد که با استفاده از داده‌های روزانه و هفتگی بازده شاخص کل سهام تهران (TEPIX) از سال ۱۳۷۷ تا ۱۳۸۲ و به کارگیری تکنیک‌های مختلف پیش‌بینی مانند ARIMA، ARFIMA، GARCH و شبکه عصبی مصنوعی (ANN) به پیش‌بینی این شاخص با استفاده از معیارهای ارزیابی RMSE، MAE، U-Thiel پرداخته‌اند که نتایج نشان‌دهنده برتری ANN نسبت به سایر روش‌ها بود. هم‌چنین با استفاده از آماره‌ی دیبلد-ماریانو^۴ (DM)، نشان دادند که تفاوت معنی‌داری در دقت پیش‌بینی مدل‌های مذکور وجود ندارد.

منجمی و همکاران (۱۳۸۸) با استفاده از شبکه‌های عصبی فازی و الگوریتم ژنتیک و مقایسه آن با شبکه عصبی به پیش‌بینی قیمت سهام پرداختند که نتایج آن‌ها حاکی از برتری پیش‌بینی شبکه عصبی فازی و الگوریتم ژنتیک بود.

۳- مدل‌های پیش‌بینی

۳-۱. مدل خود رگرسیون میانگین متحرک انباشته (ARIMA)

اگر متغیرهای مدل مانا باشند برای پیش‌بینی از فرآیند خود توضیحی میانگین متحرک استفاده

۱- Karyl and et al

۲- Fama and French

۳- Abbasi & Abouec

۴- Diebold-Mariano

می‌کنیم. بنا به تعریف دنباله تصادفی y_t یک فرآیند تصادفی مختلط میانگین متحرک خود توضیحی، با درجات q و p است به شرطی که داشته باشیم:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (1)$$

اما در صورت نامانایی باید از فرآیند ARIMA که مبتنی بر درجه جمعی متغیرها، d است استفاده نماییم که فرآیند فوق را به صورت ARIMA(p, d, q) نشان می‌دهند، به طوری که d, p و q به ترتیب بیانگر تعداد جملات خود توضیحی^۱، مرتبه تفاضل گیری و تعداد جملات میانگین متحرک^۲ می‌باشند. معمولاً برای تخمین الگوی ARIMA و ARMA از روش باکس-جنکینز^۳ استفاده می‌شود. روش باکس و جنکینز دارای سه مرحله شناسایی، تخمین و تشخیص دقت پردازش است (سوری، ۱۳۹۰).

تعداد جملات خود توضیحی و تعداد جملات میانگین متحرک معمولاً با استفاده از توابع خود همبستگی^۴ (AC) خود همبستگی جزئی^۵ (PAC) بر اساس مراحل باکس-جنکینز محاسبه می‌شود. اما از آنجایی که ممکن است مدل‌های بهینه دیگری وجود داشته باشند که بر الگوی مذکور ترجیح داده شوند، این مدل‌ها توسط معیار اطلاعات آکائیک^۶ (AIC) یا معیار شوارتز-بیزین^۷ (SBC) بازبینی می‌شود، به گونه‌ای که مدلی مناسب است که کم‌ترین مقدار آماره آکائیک یا شوارتز-بیزین را داشته باشد.

۲-۲. مدل خود رگرسیون ناهمسان واریانس شرطی (ARCH)^۸

۱- Auto Regressive

۲- Moving Average

۳- Box-Jenkins

۴- Auto Coloration

۵- Partial Auto Coloration

۶- Akaike Information Criterion (AIC)

۷- Shwarz Bayesian Criterion (SBC)

۸- Auto Regression Conditionally Heteroskedasticity

مدت نسبتاً طولانی است که اقتصادسنجی دانان و محققان پی برده‌اند که بازدهی‌های دارایی‌های مالی دارای ویژگی خوشه‌بندی تغییرات^۱ می‌باشد؛ یعنی توزیع احتمال فراوانی آن‌ها چولگی و کشیدگی بیشتری از توزیع احتمال فراوانی نرمال دارد. اولین مدل برای توضیح وابستگی از نوع خوشه‌بندی تغییرات سری زمانی توسط انگل (۱۹۸۲)^۲ مطرح شد. وی مدل خود توضیحی واریانس ناهمسان شرطی (ARCH) را برای توجیه این نوع از وابستگی‌ها در سری زمانی مطرح نمود. بعد از وی مطالعات گوناگونی توسط سایر اقتصاددانان مانند بالرسلوف^۳ صورت گرفت و مدل‌های دیگری از نوع ARCH مانند مدل خود توضیحی واریانس ناهمسان شرطی تعمیم‌یافته (GARCH) به وجود آمدند تا بتوانند این ویژگی بازدهی‌های دارایی‌های مالی را به خوبی توضیح دهند.

از آنجایی که خوشه‌بندی تغییرات به طور ضمنی بیان می‌کند که هر تغییر ناگهانی بزرگ با هر علامتی (مثبت یا منفی) احتمال تغییرات آتی را افزایش می‌دهد، بنابراین پیش‌بینی تغییرات که این حقیقت (خوشه‌بندی تغییرات) را در نظر می‌گیرند، معمولاً دقیق‌تر از پیش‌بینی‌هایی خواهند بود که آن را در نظر نمی‌گیرند (ادگار، ۱۹۹۲).

انواع مختلفی از مدل‌های نوع ARCH در مدل‌سازی و پیش‌بینی اقتصادی استفاده می‌شود که معروف‌ترین این مدل‌ها عبارتند از: مدل ARCH(p)، مدل GARCH(p,q)، مدل ARCH-M^۴، مدل EGARCH^۵، مدل IGARCH^۶ اما بیشتر مطالعات تجربی نشان داده‌اند که مدل‌های نوع GARCH از توانایی بیشتر در مدل‌سازی و پیش‌بینی بازارهای سهام برخوردار می‌باشد. یک مدل کلی ARCH(q) به شکل زیر است (بالرسلوف، ۱۹۸۶):

۱- Volatility Clustering

۲- Engel

۳- Bollerslev

۴- ARCH in Mean Model

۵- Exponential GARCH

۶- Integrated GARCH

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + V_t \quad (2)$$

مشکلی که در عمل به هنگام استفاده از این مدل‌ها به وجود می‌آید، این است که هنگامی که q عدد بزرگی باشد، به طور معمول باعث نقض شدن فرض غیر منفی بودن و شرایط مانایی معادله‌ی واریانس می‌شود. مدل GARCH که توسط بالرسلوف (۱۹۸۶) توسعه یافت، راه‌کاری برای حل این مشکل و نیز مدلی صرفه‌جو در تعداد پارامترها است (جانستون و اسکات، ۲۰۰۰). به دلیل تعداد پارامترهای زیادی که برای یک مدل ARCH باید تخمین زده می‌شد، بالرسلوف (۱۹۸۶) این مدل را به صورت تعمیم‌یافته ارائه داد که در این مدل تعمیم‌یافته، معادله‌ی واریانس شرطی علاوه بر q توان دوم خطای گذشته، شامل p وقفه‌ی واریانس‌های شرطی گذشته نیز است. به عبارت دیگر، معادله‌ی واریانس شرطی در یک مدل GARCH(p,q) به صورت زیر می‌باشد (باچاد و همکاران^۲، ۲۰۰۱):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + V_t \quad (3)$$

۲-۳. شبکه‌های عصبی مصنوعی (ANN)

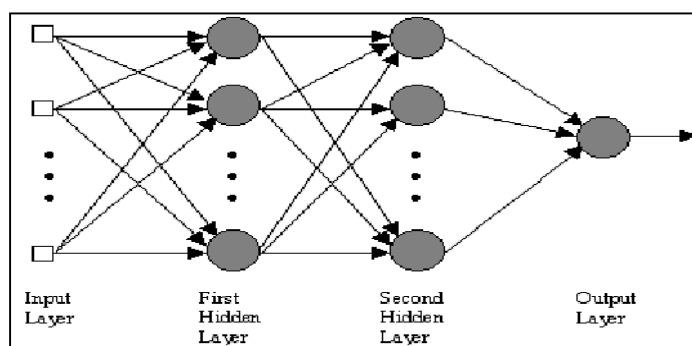
در سال‌های اخیر شاهد استفاده گسترده از شبکه‌های عصبی مصنوعی بوده‌ایم. شبکه‌های عصبی ابزار محاسباتی ساده‌ای برای آزمون داده‌ها و ایجاد مدل از ساختار داده‌هاست. داده‌هایی که برای ایجاد مدل‌ها استفاده می‌شوند، به داده‌های آموزشی مشهورند. انواع مختلفی از شبکه‌های عصبی مصنوعی با توجه به اهداف تحقیق می‌تواند مورد استفاده قرار گیرد که یکی از معروف‌ترین آن‌ها، شبکه عصبی چند لایه پیش‌خور (MFNN) است. این نوع شبکه به دلیل کاربردهای گسترده در بسیاری از ابعاد مسائل مربوط به اقتصاد، مانند پیش‌بینی اصولی، طبقه‌بندی و مدل‌سازی، بسیار مورد استقبال است. شبکه عصبی مصنوعی چند لایه پیش‌خور برای حل مسائلی که شامل یادگیری ارتباط بین یک مجموعه ورودی‌ها و خروجی‌های

۱- Johnston and Scott

۲- Bouchad and et al

مشخص هستند، مناسب است. این شبکه در حقیقت یک تکنیک آموزش با ناظر برای یادگیری ارتباط بین داده‌ها با استفاده از مجموعه داده‌های آموزش است.

شبکه عصبی مصنوعی در واقع مدل ساده شده‌ای از مغز انسان است که به صورت یک ساختار ریاضی برای نشان دادن فرآیندها و ترکیبات دلخواه غیر خطی ارتباط بین ورودی و خروجی‌های هر سیستم قابل طراحی است. این شبکه با داده‌های موجود طی فرآیند یادگیری، آموزش می‌بیند و برای پیش‌بینی در آینده مورد استفاده قرار می‌گیرد. شبکه عصبی مصنوعی از سلول‌های عصبی به نام نرون و واحدهای ارتباطی به نام اکسون^۱ تشکیل یافته است. نرون‌های شبکه عصبی در حقیقت شکل بسیار ساده‌ای از نرون‌های بیولوژیکی هستند. شکل (۱) شمای ساده‌ای از مدل شبکه عصبی را نشان می‌دهد (منهاج، ۱۳۷۹).



شکل (۱): ساختار ساده شبکه عصبی مصنوعی

همان‌طور که در شکل فوق پیداست، معماری شبکه متشکل از سه لایه است: لایه ورودی که داده‌ها را در شبکه توزیع می‌کند، لایه پنهان که داده‌ها را پردازش می‌کند و لایه خروجی که نتایج را به‌ازای ورودی‌های مشخص استخراج می‌کند.

در هر لایه تعدادی نرون به عنوان واحدهای پردازنده وجود دارد که با اتصالاتی وزن‌دار به همدیگر مرتبط می‌شوند. عملیاتی که در هر نرون صورت می‌گیرد، بدین ترتیب است که نخست

نرون تمامی ورودی‌هایی را که به سلول رسیده است، جمع می‌کند؛ سپس مقدار آستانه‌ای نرون را از آن کم می‌کند و سرانجام آن را از یک تابع محرک یا تابع فعالیت عبور می‌دهد. در نهایت خروجی نرون به دست می‌آید. طی این فرآیندها، خطای مقادیر خروجی به کم‌ترین مقدار خود می‌رسد. برای انتقال خروجی‌های هر لایه به لایه‌های بعد از توابع محرک استفاده می‌شود. توابع سیگموئیدی، خطی و آستانه‌ای از جمله توابع محرک پر کاربرد هستند. به طور کلی شبکه‌های عصبی مصنوعی به دو گروه تقسیم می‌شوند که عبارتند از شبکه‌های پیش‌خور^۱ و شبکه‌های پس‌خور^۲.

۴-۲. شبکه عصبی مصنوعی فازی

در میان روش‌های مدل‌سازی نونین، سیستم‌های فازی جایگاه ویژه‌ای را کسب کرده‌اند. این امر را می‌توان معلول توانایی پیاده‌سازی دانش بشری با استفاده از مفاهیم و عبارت‌های زبانی و قواعد فازی، روابط غیر خطی 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 مورد نظر است. به دست آوردن این قواعد مهم‌ترین و سخت‌ترین مرحله کار است و به دانش لازم و توان کافی برای پیاده‌سازی صحیح آن قواعد نیاز دارد.

به‌طور کلی، شبکه‌های عصبی مصنوعی برای اهدافی که ما از آن استفاده می‌کنیم، توانایی بالایی در توسعه یک مدل در زمانی منطقی را ندارد. از سوی دیگر، مدل‌سازی فازی برای کاربرد ادغام تصمیمات از متغیرهای متفاوت، نیازمند یک رویکردی جهت‌یادگیری از تجربیات (داده‌های جمع‌آوری شده) است. شبکه عصبی مصنوعی و مدل فازی در بسیاری از زمینه‌های کاربردی استفاده شده‌اند و در هر کدام محاسن و معایبی است. بنابراین ترکیب موفقیت‌آمیز این دو دیدگاه، مدل‌سازی شبکه‌های عصبی مصنوعی و فازی، موضوع مطالعات آینده قرار گرفته

۱- Feed Forward

۲- Feed Backward

است.

یکی از متداول‌ترین روش‌های ترکیبی سیستم استنتاج تطبیقی عصبی فازی (ANFIS) است که توسط ژانگ^۱ در سال ۱۹۹۳ معرفی گردید و حاوی مزیت‌های هر دو سیستم فازی و شبکه عصبی است (ژانگ و همکاران، ۱۹۹۷ و سرینیواسان و فیشر، ۱۹۹۵). این مدل یک سامانه فازی^۲ از نوع تاکاگی-سوگنو^۳ را در یک ساختار عصبی اجرا می‌کند و برای فرآیند آموزش از روش پس انتشار خطا یا ترکیبی از روش‌های آموزش پس انتشار خطا و حداقل مربعات خطا استفاده می‌کند. در ANFIS قوانین ثابت هستند و عامل‌های توابع عضویت بهینه‌سازی می‌شود. برای تعیین عامل‌های توابع عضویت (یا همان شکل توابع عضویت) از روند نماهای آموزشی شبکه عصبی استفاده می‌شود. نوع توابع عضویت (از قبیل مثلثی و گوسی) و تعداد توابع عضویت برای ورودی‌ها و خروجی‌ها به روش سعی و خطا تعیین می‌گردد. در ANFIS لازم است که در لایه اول نوع تابع عضویت و تعداد آن‌ها مشخص شود.

ANFIS مدلی مانند f را شناسایی می‌کند 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 که برای پیش‌بینی استفاده می‌شود و به صورت زیر ارائه می‌گردد:

$$f(x) = \frac{\sum_{i=1}^N E_t \left(\prod_{j=1}^n \mu_{A_i(j)}(E_{t-p_j}) \right)}{\sum_{i=1}^N \left(\prod_{j=1}^n \mu_{A_i(j)}(E_{t-p_j}) \right)} \quad (4)$$

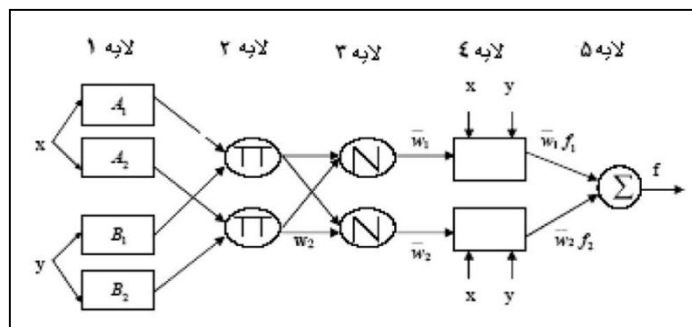
که در آن E_{t-p} ، μ ، A و i به ترتیب نشانگر وقفه‌های متغیر تا مرتبه p ، مرتبه عضویت، مجموعه فازی و قاعده فازی. هم‌چنین $\mu_{A_i(j)}(x_i)$ بیانگر مرتبه عضویت از ورودی x_i مربوطه به مقدار

۱- Jang

۲- در ادبیات منطق فازی، انواع مختلفی از سامانه‌های فازی پیشنهاد گردیده است و معروف‌ترین آن‌ها سیستم استدلال تاکاگی سوگنو (۱۹۸۵) و سیستم استدلال ممدانی (۱۹۷۷) است. این دو روش در بسیاری از جنبه‌ها یکسان می‌باشند مانند فازی سازی ورودی‌ها و عملگرهای فازی، اما تفاوت اصلی بین این دو روش در شکل خروجی آن‌ها می‌باشد (کینک و همکاران، ۲۰۰۲). در سیستم ممدانی خروجی به صورت یک مجموعه فازی است که باید دفاژی (Defuzzify) شوند، ولی در سیستم سوگنو خروجی به صورت خطی یا ثابت است (راهنمای نرم‌افزار متلب، ۲۰۱۰).

۳- Takagi-Sugeno

قاعده فازی نام $A_i^{(ji)}$ است (شکل (۲)).



شکل (۲): ساختار ANFIS با دو ورودی

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:

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(X) \quad i = 1, 2. \\ O_{2,i} &= \mu_{B_{i-2}}(Y) \quad i = 3, 4. \end{aligned} \quad (5)$$

در این لایه x و y ورودی‌های غیر فازی 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 است که خروجی گره‌ها حاصل ضرب تمامی سیگنال‌های ورودی است.

$$O_{2,i} = W_i = \mu_{A_i}(x_i) * \mu_{B_i}(y_i) \quad i = 1, 2. \quad (6)$$

هر گره خروجی بیانگر 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 \bar{w}_i نامین گره، نسبت \bar{w}_i نامین قاعده شدت برانگیختگی را برای تمامی قاعده‌های شدت برانگیختگی به صورت زیر محاسبه می‌کند:

$$O_{3,i} = \bar{W}_i = \frac{W_i}{W_1 + W_2} \quad i = 1, 2. \quad (7)$$

۱- Firing Strength

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

$$O_{4,i} = \bar{W}_i f_i = \bar{W}_i (p_i x + q_i y + r_i) \quad i = 1, 2. \quad (8)$$

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 گره هستند که به پارامترهای استنتاجی موسومند.

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

$$O_{5,i} = \sum \bar{W}_i f_i = \frac{\sum w_i f_i}{\sum w_i} \quad i = 1, 2. \quad (9)$$

داده‌ها

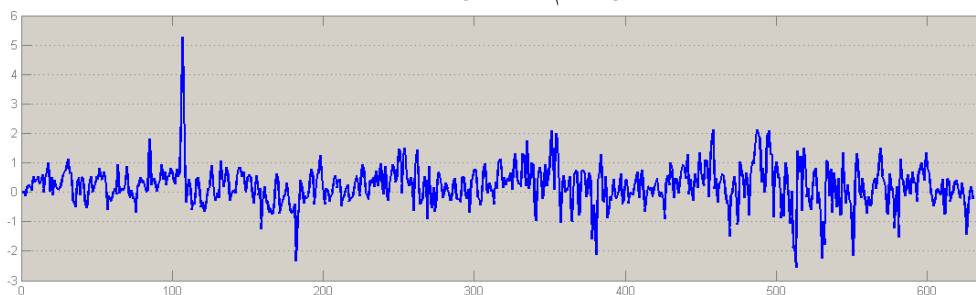
در این مطالعه از داده‌های سری زمانی بازده روزانه شاخص بورس اوراق بهادار تهران (TEPIX) در بازه زمانی ۱۳۸۸/۰۱/۰۵ تا ۱۳۹۰/۰۷/۲۳ با حجم نمونه ۶۲۵ مشاهده استفاده شده است. ۸۰ درصد از مشاهدات (۵۰۵ مشاهده) برای برآورد مدل‌ها و ۱۲۰ مشاهده دیگر برای پیش‌بینی و مقایسه دقت پیش‌بینی‌ها استفاده شده است.

بازده روزانه شاخص بورس اوراق بهادار تهران به صورت درصد تغییر در مقدار شاخص قیمت (TEPIX) تعریف و به صورت زیر محاسبه می‌شود:

$$R_t = \ln \left(\frac{P_t}{P_{t-1}} \right) * 100 \quad (10)$$

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

روند سری زمانی بازده شاخص کل سهام در شکل زیر مشاهده می‌شود.



شکل (۳): نمودار بازده شاخص کل بورس اوراق بهادار

ماخذ: محاسبات تحقیق

خلاصه ای از آماره های توصیفی برای بازده شاخص بورس تهران در جدول (۱) قابل مشاهده است. همان طور که انتظار می رود، میانگین سری مورد نظر نزدیک به صفر است. هم چنین شواهدی مبنی بر چولگی^۱ مثبت در توزیع بازده وجود دارد که نشان دهنده احتمال زیاد برای افزایش های بزرگ نسبت به کاهش ها در بازده سبد دارایی ها^۲ است. این پدیده نشان می دهد که بازده شاخص بورس تهران احتمالاً نامتقارن است. هم چنین بازده شاخص کل، دارای کشیدگی^۳ پایین است. آماره جارکو- برآ^۴ نشان می دهد که با احتمال بسیار بالا فرض صفر مبنی بر نرمال بودن توزیع بازده رد می شود، بنابراین بازده شاخص کل از توزیع نرمال تبعیت نمی کند.

- بررسی متغیر به لحاظ خطی، غیر خطی یا آشوبناک بودن

براک، دیکرت و شاینکمن^۵ در سال ۱۹۸۷ یک آزمون آماری بر مبنای فرض استقلال و توزیع یکسان (IID) اعضای سری زمانی است، ارائه داده اند که برای آزمون وضعیت ساختار سیستم از نظر خطی یا غیر خطی بودن قابل استفاده است. آماره این آزمون (DBS) به صورت زیر تعریف می شود:

$$W_{m,T}(\varepsilon) = \sqrt{T} [C_{m,T}(\varepsilon) - C_{1,T}(\varepsilon)^m] / \sigma_{m,T}(\varepsilon) \quad (11)$$

جدول (۱): آماره های توصیفی بازده روزانه

متغیر	میانگین	واریانس	جدا آفر	حداقل	انحراف معیار	چولگی	کشیدگی	جارکو-برا	مجموع	مجموع مربعات	تعداد مشاهدات
بازده	۰/۱۹۳۳	۰/۱۸۳۳	۵/۲۶۱	-۲/۵۴۵	۰/۶۸	۰/۴۴	۹/۱۴	۱۰۰۵	۱۲۰/۸۳	۲۸۹/۰۱	۶۲۵

۱- Skewness

۲- Portfolio

۳- Kurtosis

۴- Jarque-Bera

۵- Brock, Dechert and Scheinckman

روزانه										
--------	--	--	--	--	--	--	--	--	--	--

ماخذ: یافته‌های تحقیق

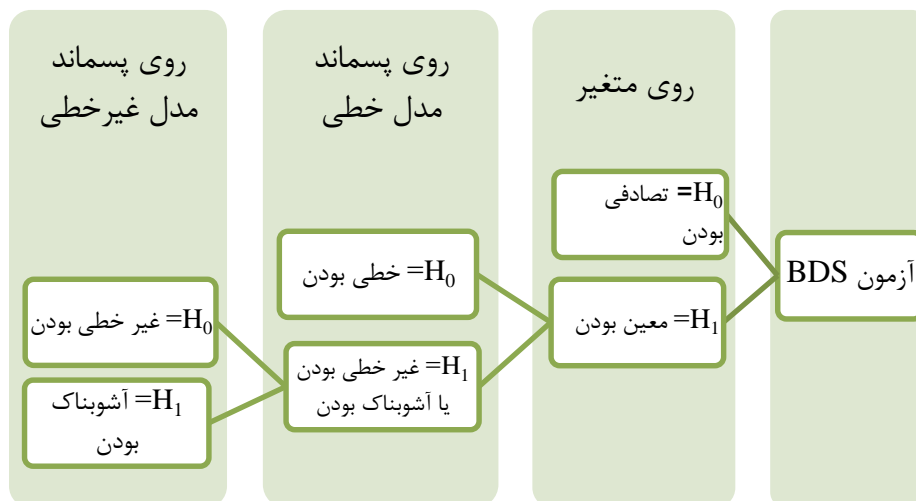
ε که در آن انحراف معیار عبارت داخل براکت، m بُعد و ε مقدار دلخواه و کوچک است.

آماره $W_{m,T}(\varepsilon)$ در نمونه‌های به حد کافی بزرگ دارای توزیع نرمال با میانگین صفر و واریانس ثابت است. بنابراین، مقدار بزرگ W منجر به رد فرضیه صفر و شاهدهی برای وجود وابستگی (از نوع خطی یا غیر خطی) در سیستم خواهد بود و مقدار نزدیک به صفر آن به معنی پذیرش فرض صفر و تاییدی بر استقلال و توزیع یکسان مشاهدات است. در این آماره، انتخاب مقدار ε دارای اهمیت بسیار است و اغلب به صورت ضربی از انحراف معیار سری زمانی انتخاب می‌شود. از آماره BDS برای آزمون خطی یا غیر خطی بودن داده‌های بازده سهام بورس اوراق بهادار، در سه مرحله استفاده شده است: مرحله اول انجام آزمون روی بازده روزانه شاخص سهام؛ مرحله دوم انجام آزمون روی پسماندهای مدل ARMA و مرحله سوم انجام آزمون روی پسماندهای مدل‌های ANN و GARCH، ANFIS.

بر اساس نمودار ۱، اگر آزمون BDS برای داده‌های خام (در این جا بازده روزانه شاخص سهام) صورت گیرد، فرضیه صفر آزمون به صورت زیر خواهد بود:

H_0 : تصادفی بودن فرآیند سری زمانی

H_1 : معین و غیر تصادفی بودن فرآیند سری زمانی



نمودار (۱): مراحل مختلف انجام آزمون BDS

چون آماره آزمون BDS دارای توزیع Z است، برای تصمیم گیری در مورد این آزمون از مقدار بحرانی آماره Z در سطح معنی داری ۵ درصد یعنی $۲/۶۴$ و سطح معنی داری ۱۰ درصد یعنی $۱/۶۴$ و مقایسه آن با آماره محاسباتی استفاده می شود. با توجه به جدول (۲)، فرضیه صفر یعنی تصادفی بودن سری زمانی رد و فرضیه مقابل آن مبنی بر معین بودن سری پذیرفته می شود. لازم به ذکر است که معین بودن سری الزاماً به معنای آشوبی بودن آن نیست.

جدول (۲): آماره BDS بر روی سری بازده روزانه شاخص بورس اوراق بهادار

	$\varepsilon/\delta = 0/5$	$\varepsilon/\delta = 1/0$	$\varepsilon/\delta = 1/5$	$\varepsilon/\delta = 2/0$
$m = 2$	۱۱/۲۷	۱۰/۱۵	۱۰/۴۳	۱۰/۸۶
$m = 3$	۱۲/۰۸	۱۰/۵۴	۱۰/۶۹	۱۰/۹۱
$m = 4$	۱۳/۲۰	۱۰/۹۵	۱۰/۷۴	۱۰/۸۳
$m = 5$	۱۴/۷۵	۱۱/۲۶	۱۰/۵۸	۱۰/۴۶
$m = 6$	۱۶/۵۱	۱۱/۷۸	۱۰/۶۱	۱۰/۲۴

مأخذ: یافته های تحقیق

در جدول فوق مقدار ε برابر با $0/5$ ، 1 و $1/5$ و 2 برابر انحراف معیار استاندارد سری در نظر

گرفته شده است. همچنین اعداد داخل جدول آماره Z را نشان می دهند.

برای انجام آزمون BDS در مرحله دوم، فرآیند خطی موجود در داده‌ها به وسیله مدل ARMA استخراج و سپس آزمون BDS بر روی پسماندهای این مدل انجام می شود. اگر سیستم مولد داده‌ها خطی باشد، فرضیه صفر مبنی بر اینکه پسماندهای مدل خطی IID هستند، پذیرفته می شود. اما اگر سیستم مولد داده‌ها غیر خطی باشد، پسماندهای مدل خطی IID نیست و در نتیجه فرضیه صفر رد خواهد شد. در این آزمون فرضیه صفر به صورت زیر خواهد بود:

خطی بودن فرایند سری زمانی: H_0

خطی غیر بودن فرایند سری زمانی: H_1

با توجه به مدل برگزیده ARIMA، با انجام آزمون BDS بر روی پسماندهای مدل نتایج زیر به دست می آید.

جدول (۳): آماره BDS بر روی پسماندهای مدل برازش شده ARMA

	$\varepsilon/\delta = 0/5$	$\varepsilon/\delta = 1/0$	$\varepsilon/\delta = 1/5$	$\varepsilon/\delta = 2/0$
$m = 2$	۳/۰۷	۲/۴۶	۱/۹۰	۲/۲۸
$m = 3$	۵/۵۸	۴/۲۵	۲/۴۰	۲/۶۵
$m = 4$	۶/۷۰	۴/۷۵	۲/۶۴	۲/۵۱
$m = 5$	۷/۹۸	۴/۹۹	۲/۷۶	۲/۴۷
$m = 6$	۷/۹۲	۵/۴۱	۲/۷۱	۲/۵۲

مأخذ: یافته‌های تحقیق

با توجه به جدول (۳)، فرضیه صفر در سطح معنی داری ۵ درصد رد می شود که نشان دهنده ساختار غیر خطی در مدل است.

در مرحله آخر برای آزمون ساختار آشوبی، ساختار غیر خطی مدل با استفاده از مدل‌های ANN، ANFIS و GARCH استخراج می شود و سپس آزمون BDS بر روی پسماندهای این مدل انجام می شود. فرضیه صفر این آزمون به صورت زیر است:

وجود فرآیند غیر خطی در مدل: H_0

وجود فرآیند آشوبی در مدل: H_1

جدول (۴): آماره BDS بر روی پسماندهای مدل برازش شده ANFIS

	$\varepsilon/\delta = 0/5$	$\varepsilon/\delta = 1/0$	$\varepsilon/\delta = 1/5$	$\varepsilon/\delta = 2/0$
$m = 2$	۳/۳۱	۲/۹۸	۳/۲۲	۲/۴۷
$m = 3$	۲/۸۲	۳/۴۵	۳/۷۰	۲/۵۴
$m = 4$	۲/۹۸	۳/۹۹	۴/۲۶	۳/۰۴
$m = 5$	۱/۵۷	۳/۹۹	۴/۳۷	۳/۱۸
$m = 6$	۱/۱۷	۳/۵۶	۳/۹۷	۲/۵۷

مأخذ: یافته‌های تحقیق

جدول (۴): آماره BDS بر روی پسماندهای مدل برازش شده GARCH

	$\varepsilon/\delta = 0/5$	$\varepsilon/\delta = 1/0$	$\varepsilon/\delta = 1/5$	$\varepsilon/\delta = 2/0$
$m = 2$	۴/۲۹	۴/۶۰	۵/۱۷	۵/۶۳
$m = 3$	۵/۱۱	۵/۶۰	۶/۳۰	۶/۹۵
$m = 4$	۵/۸۰	۶/۴۶	۷/۱۶	۷/۶۶
$m = 5$	۵/۶۱	۶/۸۱	۷/۵۴	۷/۹۲
$m = 6$	۵/۲۵	۷/۰۵	۷/۹۱	۸/۱۷

مأخذ: یافته‌های تحقیق

جدول (۵): آماره BDS بر روی پسماندهای مدل برازش شده ANN

	$\varepsilon/\delta = 0/5$	$\varepsilon/\delta = 1/0$	$\varepsilon/\delta = 1/5$	$\varepsilon/\delta = 2/0$
$m = 2$	۳/۳۱	۲/۹۸	۳/۲۲	۲/۴۷
$m = 3$	۲/۸۲	۳/۴۵	۳/۷۰	۲/۵۴
$m = 4$	۲/۹۸	۳/۹۹	۴/۲۶	۳/۰۴
$m = 5$	۱/۵۷	۳/۹۹	۴/۳۷	۳/۱۸
$m = 6$	۱/۱۷	۳/۵۶	۳/۹۷	۲/۵۷

مأخذ: یافته‌های تحقیق

با توجه به سه جدول اخیر فرضیه صفر در سطح معنی داری ۵ و ۱۰ درصد رد می‌شود که نشان

دهنده رفتار آشوبناک در متغیر مورد بررسی است.

۵- برآورد مدل

۱-۵. پیش‌بینی بازده روزانه توسط مدل ARMA

برای پیش‌بینی بازده سهام بورس اوراق بهادار به روش باکس-جنکینز، باید مدل مناسب را از میان مدل‌های خود بازگشتی (AR)، میانگین متحرک (MA) و خود بازگشت میانگین متحرک (ARIMA) انتخاب کنیم. اولین مرحله در این انتخاب بررسی مانایی سری زمانی است. بدین منظور از آزمون دیکی-فولر تعمیم یافته^۱ استفاده می‌شود. نتایج این آزمون نشان داد که سری مورد استفاده در سطح ماناست. که در جدول زیر نمایش داده شده است ((۰,۰۰۰)، (۰,۰۰۰) (-۱۵/۷۷).

جدول (۶): انتخاب بهترین مدل ARMA با استفاده از معیار آکائیک

AR/MA	0	1	2	3	4	5	6	7	8	9	10
0	1.98446	1.73831	1.73646	1.73830	1.74074	1.73616	1.73712	1.73831	1.72229	1.72466	1.72626
1	1.75260	1.73505	1.72185	1.72575	1.72869	1.72950	1.73335	1.73737	1.72220	1.72560	1.72906
2	1.74319	1.73676	1.72578	1.72653	1.71210	1.73344	1.73474	1.71618	1.72507	1.72904	1.71154
3	1.73012	1.72487	1.72812	1.71124	1.71584	1.72046	1.71103	1.72014	1.72911	1.73261	1.72243
4	1.73138	1.72833	1.73132	1.71307	1.71905	1.72115	1.72588	1.71200	1.70936	1.71333	1.71731
5	1.73119	1.73409	1.73514	1.71161	1.71169	1.72892	1.71293	1.71007	1.73749	1.70950	1.72363
6	1.73487	1.73800	1.71583	1.71264	1.71567	1.73234	1.70781	1.70452	1.70786	1.69472	1.71133
7	1.72880	1.72876	1.72669	1.72918	1.73264	1.72322	1.70958	1.69038	1.69408	1.70866	1.69599
8	1.72445	1.72217	1.72542	1.72868	1.69755	1.69299	1.69044	1.70968	1.685657	1.69684	1.69662
9	1.72365	1.72574	1.72914	1.73161	1.72740	1.73232	1.69422	1.69747	1.69197	1.69624	1.68985
10	1.72511	1.72914	1.72917	1.71898	1.72255	1.71932	1.68982	1.69140	1.71463	1.71314	1.69008

مأخذ: یافته‌های تحقیق

سپس باید تعداد وقفه‌های بهینه برای مدل‌های AR و MA انتخاب شود. برای آغاز ۱۰ وقفه برای مدل‌های خود بازگشتی و میانگین متحرک در نظر گرفته و کد لازم در ایویوز ۷ نوشته و اجرا شد. بر اساس معیار آکائیک، ۸ وقفه از مدل خود بازگشتی و ۸ وقفه از مدل میانگین متحرک به عنوان وقفه‌های بهینه انتخاب شد و همان‌طور که در جدول زیر نشان داده شده است، کم‌ترین

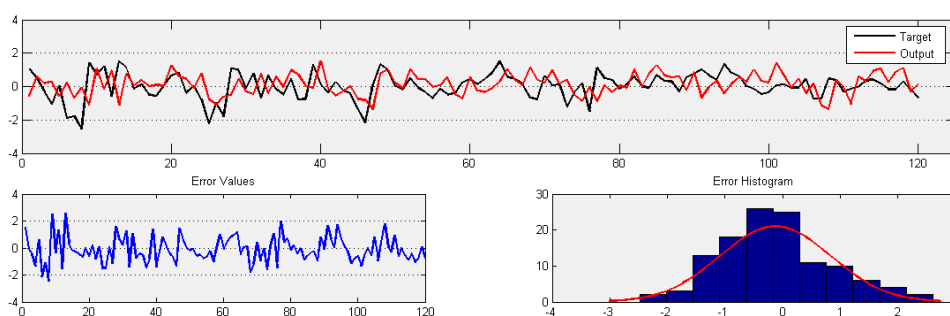
۱- Augmented Dickey-Fuller tes

۲- EViews

مقدار آماره آکایک در بین ۱۲۱ معادله تخمین زده شده در وقفه‌های منتخب رخ می‌دهد. بنابراین مدل نهایی به صورت $ARMA(8, 8)$ است.

برای اطمینان از اعتبار نتایج مدل فوق به بازبینی مدل از طریق بررسی مانایی باقی‌مانده‌ها پرداختیم که نتیجه کاملاً رضایت بخش است؛ زیرا آزمون ADF برای باقی‌مانده‌های مدل $ARIMA(8, 8)$ برابر با $-14/95$ با سطح معنی‌داری $0/000$ است.

در شکل زیر مقادیر واقعی متغیر، مقادیر برازش شده از طریق پیش‌بینی شبکه، سری خطاها و هیستوگرام جزء خطا و هم‌چنین مقایسه آن با توزیع نرمال نشان داده شده است.



شکل (۴): مقادیر پیش‌بینی شده توسط مدل ARMA

ماخذ: یافته‌های تحقیق

۲-۵. پیش‌بینی بازده روزانه توسط مدل GARCH

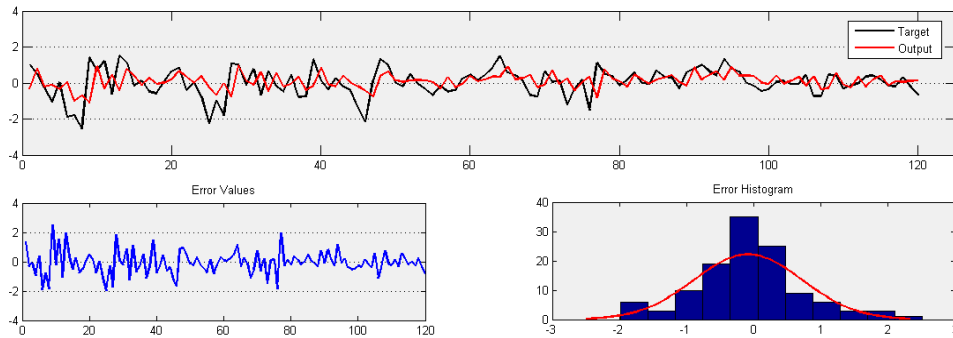
به منظور پیش‌بینی بازده روزانه شاخص بورس اوراق بهادار با مدل GARCH ابتدا مدل $GARCH(1, 1)$ بر بازده روزانه شاخص برازش و سپس مقدار این شاخص‌ها با استفاده از این مدل پیش‌بینی شد. دقت پیش‌بینی این مدل با استفاده از معیارهای مختلف با سایر مدل‌های پیش‌بینی مورد استفاده در این تحقیق مقایسه شد. نتایج مدل $GARCH(1, 1)$ به صورت زیر است:

در شکل فوق مقادیر واقعی متغیر، مقادیر برازش شده از طریق پیش‌بینی شبکه، سری خطاها و هیستوگرام جزء خطا و هم‌چنین مقایسه آن با توزیع نرمال نشان داده شده است.

۳-۵. پیش‌بینی بازده روزانه توسط مدل ANN

به منظور پیش‌بینی بازده روزانه شاخص به روش شبکه‌های عصبی مصنوعی از کدنویسی در نرم افزار متلب استفاده شد. برای آموزش شبکه، روش آزمون و خطا به کار گرفته شد و از ده نرون در

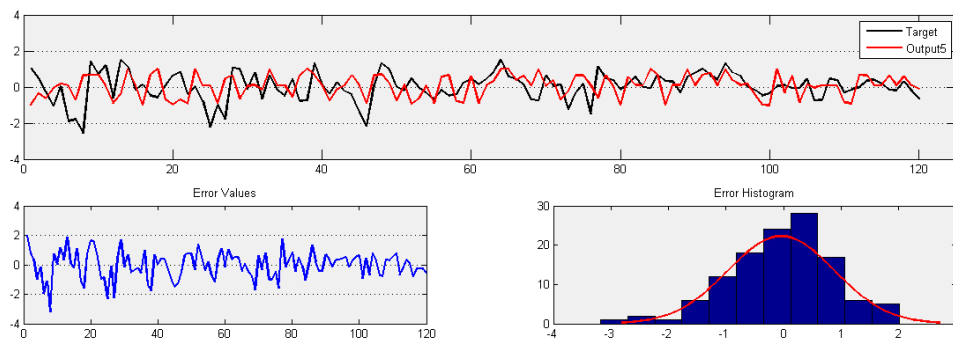
لایه ورودی، چهار نرون در لایه میانی و توابع فعال‌سازی تانژانت سیگموید^۱ و از تابع همانی^۲ در لایه خروجی برای یافتن مدل بهینه استفاده شد. در این ارتباط از ده وقفه اول سری زمانی متغیر مورد بررسی برای مدل‌سازی استفاده شد.



شکل (۵): مقادیر پیش‌بینی شده توسط مدل $GARCH(1,1)$

ماخذ: یافته‌های تحقیق

در شکل زیر مقادیر واقعی متغیر، مقادیر برآزش شده از طریق پیش‌بینی شبکه، سری خطاها و هیستوگرام جزء خطا و همچنین مقایسه آن با توزیع نرمال نشان داده شده است.



شکل (۶): مقادیر پیش‌بینی شده توسط مدل ANN

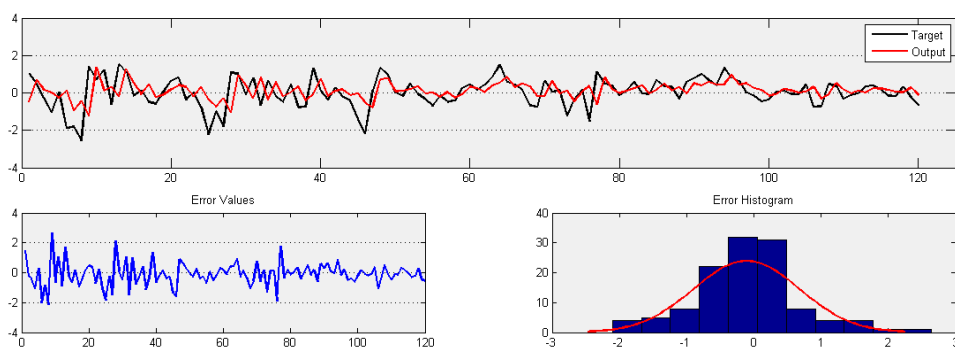
ماخذ: یافته‌های تحقیق

۱- Tangent Sigmoid

۲- Identity Function

۴-۵. پیش بینی بازده روزانه توسط مدل ANFIS

برای آموزش و پیش بینی شبکه عصبی فازی از کد نویسی در نرم افزار متلب نسخه ۲۰۱۰ استفاده شد. مدل مورد استفاده یک سامانه استنتاج فازی سوگنو با تابع عضویت ورودی «گوسی»^۱ و توابع خروجی خطی را در یک ساختار عصبی اجرا می کند و برای فرآیند آموزش از ترکیبی از روش های آموزش پس انتشار خطا و حداقل مربعات خطا استفاده می کند. بدین صورت موفق به پیاده سازی یک سیستم عصبی فازی شده ایم که قابلیت یادگیری داشته باشد که بدین صورت عمل می کند؛ در هر دور آموزش، هنگام حرکت رو به جلو خروجی های گره ها به صورت عادی تا لایه آخر محاسبه می شوند و سپس پارامترهای نتیجه توسط روش کم ترین مجموع مربعات خطا محاسبه می شوند. در ادامه پس از محاسبه خطا در بازگشت رو به عقب نسبت خطا بر روی پارامترهای شرط، پخش شده و با استفاده از روش شیب نزولی خطا مقدار آنها تصحیح می شود. در این مطالعه ۸۰ درصد داده ها برای آموزش و ۲۰ درصد باقی مانده برای پیش بینی در نظر گرفته شده است. برای مدل سازی و آموزش شبکه از طریق تغییر مداوم توابع عضویت مختلف، تعداد توابع عضویت و تعداد وقفه های متغیر، از طریق آزمون و خطا شبکه ای با شش تابع عضویت «گوسی» برای ورودی و توابع خطی برای متغیر خروجی شبکه ای با وقفه های ۱، ۴، ۸ و ۱۲ ساخته شد. در شکل زیر مقادیر واقعی متغیر، مقادیر برازش شده از طریق پیش بینی شبکه، سری خطاها و هیستوگرام جزء خطا و هم چنین مقایسه آن با توزیع نرمال نشان داده شده است.



شکل (۷): مقادیر پیش بینی شده توسط مدل ANFIS

ماخذ: یافته‌های تحقیق

۶- معیارهای ارزیابی کارایی مدل‌ها پیش‌بینی

هایکین^۱ (۱۹۹۴)، برای بررسی کارایی مدل‌های پیش‌بینی سری‌های زمانی، معیارهایی را معرفی کرد که نام و فرمول برخی از مهم‌ترین آن‌ها در زیر آمده است. یکی از این معیارها، معیار ریشه‌ی میانگین مجذور خطا^۲ (RMSE) اشاره کرد که یکی از متداول‌ترین معیارهای ارزیابی به شمار می‌رود.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} \quad (12)$$

از دیگر معیارهای متداول در ارزیابی پیش‌بینی‌ها می‌توان به میانگین مطلق خطا^۳ MAE و معیار معیار درصد میانگین مطلق خطاهای پیش‌بینی^۴ (MAPE) اشاره کرد. امتیاز استفاده از شاخص MAPE این است که می‌توان برای مقایسه پیش‌بینی سری‌هایی که دارای مقیاس متفاوت هستند از آن استفاده کرد، زیرا این معیار وابسته به مقیاس نیست. این شاخص به صورت زیر تعریف می‌شود.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| * 100 \quad (13)$$

شاخص میانگین مطلق خطا نیز به صورت زیر است.

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (14)$$

ضریب نابرابری تایل^۵ (TIC) یکی دیگر از معیارهایی است که برای مقایسه عملکرد مدل‌های

۱- Hykin

۲- Root Mean Squared Error (RMSE)

۳- Mean Absolute Error (MAE)

۴- Mean Absolute Percentage Error (MAPE)

۵- Theil inequality coefficient (TIC)

پیش بینی به کار برده می شود. این معیار RMSE را به گونه ای تعدیل می کند که همواره بین صفر و یک قرار گیرد. هر چه مقادیر این شاخص ها پایین تر باشد، پیش بینی مطلوب تر خواهد بود.

$$TIC = \frac{\sqrt{\sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n}}}{\sqrt{\sum_{i=1}^n \hat{Y}_i/n} + \sqrt{\sum_{i=1}^n Y_i/n}} \quad (15)$$

که در آن n داده های مورد استفاده؛ Y ، مقادیر واقعی و \hat{Y} ، مقادیر پیش بینی شده است.

۷- ارزیابی دقت مدل های پیش بینی

برای مقایسه و ارزیابی قدرت پیش بینی مدل ها از معیارهای ارزیابی کارایی مدل های پیش بینی RMSE، MAE، MAPE و TIC استفاده شده است. نتایج در جدول زیر قابل مشاهده است:

جدول ۷: مقایسه عملکرد مدل های پیش بینی

TIC	MAPE	MAE	RMSE	روش پیش بینی
۰/۶۷	۳۶۸/۷	۰/۷۵	۰/۹۶	ARMA (۸ و ۸)
۰/۶۶	۲۶۵/۵۲	۰/۵۹	۰/۸۰	GARCH
۰/۶۸	۳۰۶/۱۱	۰/۶۱	۰/۸۳	ANN
۰/۴۵	۲۴۸/۵۲	۰/۳۷	۰/۵۲	ANFIS

ماخذ: یافته های تحقیق

همان طور که در جدول فوق مشخص است مدل ANFIS از نظر تمامی معیارهای ارزیابی عملکرد، نسبت به سایر روش ها برتری دارد. بعد از آن، مدل های GARCH و ANN از لحاظ معیارهای عملکرد دقت بیشتری داشته اند.

۸- بررسی معنی داری اختلاف در پیش بینی ها

هر چند معیارهای ارزیابی مذکور از جمله معیارهای کاربردی و مفید در زمینه بررسی قدرت پیش بینی روش های مختلف هستند اما هیچ کدام از آنها قادر نیستند برتری یک روش را به صورت آماری بررسی کنند. بنابراین، برای این که از نظر آماری نیز فرضیه برابری صحت پیش بینی ها در مدل های رقیب مورد آزمون قرار گیرند، از آزمون مرگان-گرنجر-نیولند (MGN)

استفاده شد (دیبولد و ماریانو، ۲۰۰۲).

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

$$\begin{aligned} S_t &= e_{1,t} + e_{2,t} \\ D_t &= e_{1,t} - e_{2,t} \\ e_{1,t} &= x_m - x_{1p} \\ e_{2,t} &= x_m - x_{2p} \end{aligned} \quad (16)$$

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 می‌باشند. 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.

$$MGN = \frac{\hat{\rho}_{sd}}{\sqrt{\frac{1 - \hat{\rho}_{sd}^2}{N - 1}}} \quad (17)$$

که در آن ρ_{sd} ضریب همبستگی بین S_t و D_t و N تعداد مشاهدات است. مقدار MGN محاسبه شده با t جدول با درجه آزادی $N-1$ مقایسه می‌شود (دیبولد ماریانو، ۲۰۰۲).

جدول (۸): نتایج آزمون MGN

ANFIS	ANN	GARCH	ARIMA	
			-----	ARIMA
		-----	-۳,۹۵۰۱	GARCH
	-----	-۲,۰۱۱۲۰	-۱,۶۹۰۰۸	ANN
-----	-۲,۳۳۹۷۸	-۱,۰۹۳۱۴	-۴,۰۳۴۲۱	ANFIS

منبع: یافته‌های تحقیق

با مقایسه‌ی آماره‌های به دست آمده در جدول فوق با جدول توزیع نرمال مشاهده می‌شود که روش‌های غیرخطی به صورت معنی‌داری نسبت به روش خطی ARIMA پیش‌بینی دقیق‌تری داشته‌اند.

۹- جمع بندی و نتیجه گیری

معمولاً برای پیش بینی از روش های آماری چون رگرسیون خطی یا چند جمله ای، مدل های ساختاری، مدل های سری زمانی مانند خود توضیحی میانگین متحرک (ARIMA)، نمو هموار و از این قبیل روش ها استفاده می کنند. اما این مدل ها با ضعف هایی همراه اند که به محقق اجازه نمی دهد تا عوامل پیچیده و غیرخطی موثر بر پیش بینی را در نظر بگیرد. علاوه بر این اخیراً اثبات شده که بسیاری از مشاهدات سری زمانی اقتصادی غیرخطی بوده و تخمین مدل های خطی برای مسائل پیچیده دنیای واقعی همیشه رضایت بخش نیست. هم چنین پیش بینی بر اساس مدل های غیرخطی اقتصادسنجی نیز با محدودیت های زیادی همراه است. از این رو استفاده از تکنیک های غیرخطی برای پیش بینی سری های زمانی روز به روز در حال گسترش است. شبکه های عصبی مصنوعی، منطق فازی و ترکیب شبکه های عصبی مصنوعی و منطق فازی که شبکه های عصبی فازی نامیده می شود جدیداً به عنوان روش هایی برای تقریبی غیرخطی استفاده می شوند و دارای محاسن زیادی در پیش بینی، طبقه بندی، خوشه بندی و سایر تکنیک های بهینه سازی می باشند. در این مطالعه، ابتدا با استفاده از آزمون BDS به بررسی غیرخطی و آشوبناک بودن سری بازده روزانه شاخص سهام پرداخته شده است. از این آزمون برای خطی یا غیرخطی بودن داده های بازده سهام بورس اوراق بهادار، در سه مرحله استفاده شده است: مرحله اول انجام آزمون روی بازده روزانه شاخص سهام؛ مرحله دوم انجام آزمون روی پسماندهای مدل ARMA و مرحله سوم انجام آزمون روی پسماندهای مدل های ANFIS، GARCH و ANN. نتایج نشان می دهد که متغیر بازده بورس اوراق بهادار تهران از یک رفتار غیرخطی تبعیت می کند. بر این اساس انتظار می رود مدل های غیرخطی توانایی بهتری در پیش بینی این متغیر داشته باشند.

سپس از روش شبکه عصبی فازی ANFIS، مدل شبکه عصبی مصنوعی و مدل GARCH، که انعطاف پذیری لازم در پیش بینی روابط خطی و غیرخطی را دارا است، برای مقایسه با مدل ARIMA، که یک روش صرفاً خطی است، استفاده شد تا کارایی این مدل ها در پیش بینی بازده روزانه شاخص بورس سنجیده شود. با استفاده از معیارهای ارزیابی RMSE، MAE، MAPE و TIC به مقایسه نتایج مدل ها پرداخته شد که در تمام این شاخص ها مدل های غیرخطی نسبت به

مدل ARMA از عملکرد بهتری برخوردار بودند که این امر با نتایج آزمون BDS که بیان کننده رفتار غیرخطی متغیر بود، سازگار است. هم‌چنین با استفاده از آزمون مورگان-گرنجر-نیوبلد (MGN) به بررسی معنی دار بودن اختلاف در قدرت پیش‌بینی مدل‌ها پرداخته شد که نتایج نشان دهنده برتری معنی دار پیش‌بینی برای مدل‌های غیرخطی نسبت به مدل‌های خطی بود. در پایان توصیه می‌شود برای پیش‌بینی متغیرهایی مانند قیمت طلا، نرخ ارز و بازار سهام که عوامل مهم تصمیم‌گیری در بازارهای مالی هستند از روش‌هایی شبکه‌های عصبی فازی، الگوریتم ژنتیک و مدل‌های ترکیبی اقتصاد سنجی و منطق فازی مانند خود توضیحی میانگین متحرک انباشته فازی که قدرت پیش‌بینی متغیرهای غیرخطی را دارند استفاده شود.

References

- [1] Abbasi E., & Abouec, A. (2008). Stock price forecast by using neuro-fuzzy inference system. *Proceedings of World Academy of Science, Engineering and Technology*, 36, 320–323. bonds. *Journal of Financial Economics*, 33, 3–56
- [2] Bollerslev T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Economics*, 31: 307-327.
- [3] Bouchad J. P., Matacz A. (2001). Potters M. The Leverage Effect in Financial Markets: Retarded Volatility and Panic. *Physical Review Letters*, 87(22) : 228701-4.
- [4] Box, G.P.F; Jenkins, G.M. (1978). *Time Series Analysis: Forecasting and Control*. 3rd Edn., Holden Day, San Francisco.
- [5] Brock, W. A., W. Dechert, & J. Scheinkman. (1987). A test for independence based on the correlation dimension. Working paper, University of Wisconsin at Madison, University of Houston, and University of Chicago.
- [6] Chavoshi, K. (2001). The Behavior of Stock Prices in the Tehran Stock Exchange, M.S. thesis, The University of Imam Sadegh. (in Persian)
- [7] Diebold FX, and Mariano RS. (2002). Comparing predictive accuracy. *Journal of Business and Economic Statistics* 20, 134-144
- [8] Dunis C.L., J. Jalilov (2001). *Neural Network Regression and Alternative Forecasting Techniques for Predicting Variables*. Working Paper, Liverpool Business School and CIBEF.
- [9] Engle F. R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 987-1007.
- [10] Fama E., & French, K. (1993). *Common risk factors in the returns on stocks and*
- [11] Haykin S. (1994), *Neural Networks: a Comprehensive Foundation*. Macmillan, New York.
- [12] Jang J. S. R. (1993). ANFIS Adaptive network-based fuzzy inference systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 23, 665–685.
- [13] Jang J.S.R., Sun C.T, Mizutani E. (1997) *Neuro-Fuzzy and Soft Computing A*

- Computational Approach to Learning and Machine Intelligence, Prentice Hall, 640pp.
- [14] Johnston k., & Scott, E. (2000). GARCH Models and the Stochastic Process Underlying Exchange Rate Price Change. *Journal of Financial and Strategic Decisions*, Vol. 2. pp. 13-24.
- [15] Kanas A. & Yannopoulous, A. (2001). Comparing linear and nonlinear forecasts for stock returns. *International Review of Economics and finance*, Vol. 10, pp 383-398
- [16] Karyl Q.C, et al (2005). A Comparision between Fama and Frenchs Model and Artificial Neural Network in Predicating the Chinese stock market. *Computer and operations research*, vol 32, pp 2499-2512
- [17] Kiink, T. Vesterstroem J. S. & Riget J. (2002). Particle swarm optimization with spatial particle extension, *Proceedings of the IEEE Congress on Evolutionary Computation*, , pp. 1474-1479.
- [18] MATLAB user's guide, (2010). Fuzzy logic Toolbox, by the math works Inc.
- [19] McMillan, D. G. (2001). Non-linear predictability of Stock market returns. *international Review of Economics and Finance*, 10, 353-368.
- [20] Menhaj, M. (2000). *Foundations of Artificial Intelligence*, Volume 1 and 2, Amirkabir University of Technology Publishing. (in Persian)
- [21] Monajemi, A., Abzari, M., & Shavazi, A. (2009). Stock price forecast in stock market using fuzzy neural network and genetic algorithm and the comparison with artificial neural network, *Quarterly Journal of quantitative Economics*, No. 3, pp 1-26. (in Persian)
- [22] Moshiri, S., & Morovat, H. (2006). TEPIX forecast using linear and non-linear models, *Quarterly Journal of Business*, No. 41, pp 245-275. (in Persian)
- [23] Roger J. S. Jang, C. T. Sun and Mizutani E. (1996). *Neuro-Fuzzy and Soft Computing: a computational approach to learning and machine intelligence*, to be published by Prentice-Hall
- [24] Souri, A. (2011). *Econometrics with Eviews 7 application*, 2nd publishing, Tehran: Frahang Shenasi Publishing. (in Persian)
- [25] Srinivasan K., Fisher D. (1995). Machine Learning Approaches to Estimating Software Development Effort. *IEEE Transactions on Software Engineering*. 21(2): 126–137.
- [26] Takagi, T. & Sugeno, M. (1983). Derivation of fuzzy control rules from human operator's control actions. In *Proc. IFAC symp. Fuzzy Inform. Knowledge representation and decision analysis* (pp. 55–60).
- [27] Yim, Juliana. (2002). A comparison of neural networks with time series models for forecasting returns on a stock market index. Working paper, school of Economics and Finance.
- [28] Yudong, Z., & Lenan, W. (2009). Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications*, 36(5), 8849–8854.