



Kharazmi University

## Fracsion: New hybrid Algorithm for Predicting Industrial Index Trends in Tehran Stock Exchange

Hamid Reza Yousefzadeh<sup>1</sup>  , Amin Karrabi<sup>2</sup> , Aghileh Heydari<sup>3</sup> 

1. Department of Mathematics, Payame Noor University (PNU), Tehran, Iran, Iran.

✉E-mail: [usefzadeh.math@pnu.ac.ir](mailto:usefzadeh.math@pnu.ac.ir)

2. Department of Mathematics, Payame Noor University (PNU), Tehran, Iran, Iran.

E-mail: [Phd.karrabi.a@pnum.ac.ir](mailto:Phd.karrabi.a@pnum.ac.ir)

3. Department of Mathematics, Payame Noor University (PNU), Tehran, Iran, Iran.

E-mail: [a\\_heidari@pnu.ac.ir](mailto:a_heidari@pnu.ac.ir)

---

---

### Article Info

#### Article type:

Research Article

#### Article history:

Received:

27 October 2020

Received in revised form:

28 April 2021

Accepted:

9 May 2021

Published online:

20 June 2023

#### Keywords:

Fractal Interpolation,  
Artificial Neural Network  
(ANN),  
Support Vector Regression  
(SVR),  
Industry Index,  
Statistical Analysis R/S.

---

---

### ABSTRACT

#### Introduction

Stock market research shows that this market is a complex nonlinear dynamic system. Therefore, investing in this market is always associated with various risks, and therefore predicting the market trend is very important. With the rapid development of artificial intelligence theory and machine learning methods in recent years, some intelligent methods such as artificial neural network (ANN) and machine support vector (SVM) have been widely used to model and predict financial time series.

The SVM method is a supervised machine learning algorithm used to solve classification and regression problems. Initially, the SVM method was used only to solve classification problems, but later this algorithm was developed to work with regression or data estimation problems and was called support vector regression (SVR). On the other hand, In the economic and social sectors, owing to a large number of non-stationary and irregular data, fractal model, especially, fractal interpolation model is often used to fit the missing data and to predict the short-term trends.

In this paper, to improve the accuracy of predicting the trend of various industries in the stock market, a new hybrid Fracsion algorithm, which is a combination of two fractal interpolation algorithms and support vector regression, is proposed.

---

---

---

## Material and methods

In this paper, to improve the accuracy of index trend prediction in different industries, we propose a new algorithm that combines two fractal interpolation algorithm and support vector regression, called Fracsion algorithm. For this purpose, after confirming the fractal structure of different industries, we consider the value of the index in each fractal industry as primary data to predict the trend of the index. Then, by modifying the fractal interpolation algorithm, we will generate new data, and finally, by using the support vector regression algorithm on the obtained data, we will predict the index trend.

## Results and discussion

According to the research findings, the structure of various industries in the Tehran stock market in 2018 and 2019 is fractal. Since the computational volume, convergence, and prediction accuracy of artificial intelligence methods depend on the number of problem inputs, Using the modified fractal interpolation algorithm to generate new data (interval data) and then using the support vector regression algorithm will increase the accuracy of predicting the index trend in different industries of the stock market.

## Conclusion

The following conclusions were drawn from this research.

- the structure of various industries in the Tehran stock market in 2018 and 2019 is fractal
- The hybrid Fracsion algorithm has a higher prediction accuracy than the two methods of artificial neural network and support vector machine regression.

---

**How to cite:** Yousefzadeh, H.R., Karrabi, A., Heydari, A. (2023). Fracsion: New hybrid Algorithm for Predicting Industrial Index Trends in Tehran Stock Exchange. *Mathematical Researches*, 9 (1), 284-311.



© The Author(s).

Publisher: Kharazmi University

---



Kharazmi University

## فرکسیون: الگوریتم تلفیقی جدید پیش‌بینی روند شاخص صنایع بازار بورس و اوراق بهادار تهران

حمیدرضا یوسف‌زاده<sup>۱</sup>✉، امین کرابی<sup>۲</sup>، عقیده حیدری<sup>۳</sup>

۱. نویسنده مسئول، گروه ریاضی، دانشگاه پیام‌نور، تهران، ایران. رایانامه: [usefzadeh.math@pnu.ac.ir](mailto:usefzadeh.math@pnu.ac.ir)

۲. گروه ریاضی، دانشگاه پیام‌نور، تهران، ایران. رایانامه: [Phd.karrabi.a@pnum.ac.ir](mailto:Phd.karrabi.a@pnum.ac.ir)

۳. گروه ریاضی، دانشگاه پیام‌نور، تهران، ایران. رایانامه: [a\\_heidari@pnu.ac.ir](mailto:a_heidari@pnu.ac.ir)

### چکیده

### اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

در این مقاله به منظور بهبود دقت پیش‌بینی روند شاخص صنایع مختلف در بازار بورس، الگوریتم ترکیبی جدید فرکسیون، که تلفیقی از دو الگوریتم درونیابی فرکتال و رگرسیون ماشین بردار پشتیبان است، پیشنهاد می‌شود. برای رسیدن به این مهم، با معرفی الگوریتم درونیابی فرکتال اصلاح شده و پیاده‌سازی آن بر روی داده‌های اولیه متناظر با شاخص یک صنعت فرکتالی، به تولید داده‌های فرکتالی جدید مربوط به آن صنعت می‌پردازیم. در ادامه، با فراخوانی الگوریتم رگرسیون بردار پشتیبان بر روی داده‌های فرکتالی جدید، به پیش‌بینی روند شاخص آن صنعت خواهیم پرداخت. نتایج حاصل از پیاده‌سازی الگوریتم ترکیبی فرکسیون بر روی داده‌های متناظر با شاخص صنایع مختلف در بازار بورس و اوراق بهادار ایران و مقایسه عملکرد آن با عملکرد دو روش مرسوم یعنی شبکه عصبی مصنوعی و رگرسیون ماشین بردار پشتیبان، حاکی از برتری دقت پیش‌بینی الگوریتم پیشنهادی است.

تاریخ دریافت: ۱۳۹۹/۰۸/۰۶

تاریخ بازنگری: ۱۴۰۰/۰۲/۰۸

تاریخ پذیرش: ۱۴۰۰/۰۲/۱۹

تاریخ انتشار: ۱۴۰۲/۰۳/۳۰

### واژه‌های کلیدی:

درونیابی فرکتال،

شبکه عصبی مصنوعی (ANN)،

رگرسیون بردار پشتیبان (SVR)،

شاخص صنایع، تحلیل آماری R/S.

استناد: یوسف‌زاده، حمیدرضا؛ کرابی، امین؛ حیدری، عقیده؛ (۱۴۰۲). فرکسیون: الگوریتم تلفیقی جدید پیش‌بینی روند شاخص صنایع بازار بورس و اوراق بهادار تهران. پژوهش‌های ریاضی، ۹ (۱)، ۲۸۴-۳۱۱.



© نویسندگان.

ناشر: دانشگاه خوارزمی

## ۱. مقدمه

تحلیل سری زمانی، در دهه‌های اخیر محققان بسیاری را در علوم مختلف به سوی خود جذب کرده است. سری زمانی به بیان ساده، به مجموعه‌ای از داده‌ها که در طی زمان با فواصل منظم گردآوری می‌شوند، گفته می‌شود. هدف اصلی در این نوع تحلیل، ایجاد یک مدل آماری برای داده‌های وابسته به زمان براساس اطلاعات گذشته است، که توسط آن امکان پیش‌بینی آینده پدیده مورد نظر میسر می‌شود. به بیان دیگر تحلیل سری زمانی، ایجاد مدلی گذشته‌نگر است، تا امکان تصمیمات آینده‌نگر را فراهم سازد.

بررسی‌ها و مطالعات انجام شده در زمینه بازار بورس نشان می‌دهند که این بازار یک سیستم دینامیکی غیرخطی پیچیده است [۱۸،۲۱،۳۲]. بنابراین سرمایه‌گذاری در این بازار، همواره به دلیل وجود ریسک‌های مختلف، اهمیت پیش‌بینی روند بازار را افزون می‌کند. با پیشرفت سریع نظریه هوش مصنوعی و روش‌های یادگیری ماشین در سال‌های اخیر، برخی از روش‌های هوشمند مانند شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) به طور گسترده‌ای برای مدل‌سازی و پیش‌بینی سری‌های زمانی مالی مورد استفاده قرار گرفتند [۱۶ و ۱۷].

روش SVM یک الگوریتم نظارت‌شده یادگیری ماشین است که از آن برای حل مسائل طبقه‌بندی و رگرسیون استفاده می‌شود. در ابتدا روش SVM فقط برای حل مسائل طبقه‌بندی مورد استفاده قرار می‌گرفت، اما بعدها این الگوریتم برای کار با مسائل رگرسیون یا تخمین داده‌ها توسعه یافت و رگرسیون بردار پشتیبان (SVR) نامیده شد. از طرفی در بخش‌های اقتصادی و اجتماعی، به دلیل تعدد داده‌های ناپایستا و نامنظم، برای برازش داده‌های گم‌شده و پیش‌بینی روند داده‌ها اغلب از درونیابی فرکتال استفاده می‌شود [۲۸، ۲۹، ۳۳]. هدف اصلی این مقاله، معرفی یک الگوریتم تلفیقی جدید تحت عنوان الگوریتم فرکسیون برای پیش‌بینی روند شاخص صنایع مختلف بازار بورس است. برای این منظور ابتدا با انجام برخی از تغییرات اصلاحی در الگوریتم درونیابی فرکتال، به بهبود عملکرد این الگوریتم می‌پردازیم. سپس روند شاخص صنایع مختلف که دارای ساختار فرکتالی باشند، را توسط الگوریتم درونیابی فرکتال اصلاح شده (IFI)<sup>۴</sup> درونیابی کرده و در ادامه با استفاده از الگوریتم SVR روند شاخص را برای یک دوره کوتاه مدت یک ماهه پیش‌بینی می‌کنیم.

ساختار این مقاله به شرح زیر سازماندهی شده است. در بخش ۲ به بیان پیشینه پژوهشی کارهای انجام شده در زمینه نحوه بررسی و پیش‌بینی روند سهام پرداخته می‌شود. در بخش ۳، نظریه و روش‌های اساسی الگوریتم SVR و تحلیل آماری R/S بیان می‌گردد. الگوریتم درونیابی فرکتال اصلاح‌شده در بخش ۴ ارائه می‌شود. در بخش ۵ با معرفی الگوریتم تلفیقی جدید فرکسیون، به بررسی و تجزیه و تحلیل نتایج عددی حاصل از پیش‌بینی روند کوتاه مدت

<sup>1</sup> Artificial Neural Network (ANN)

<sup>2</sup> Support Vector Machine (SVM)

<sup>3</sup> Support Vector Regression (SVR)

<sup>4</sup> Improved Fractal Interpolation (IFI)

شاخص صنایع مختلف در بازار بورس ایران پرداخته می‌شود و عملکرد این الگوریتم با الگوریتم‌های شناخته شده نظیر ANN و SVR مورد مقایسه و ارزیابی قرار می‌گیرد و در پایان، بخش ۶ به بیان نتیجه‌گیری اختصاص داده می‌شود.

## ۲. ادبیات موضوع

در سیستم‌های غیرخطی پیچیده که مدل‌سازی، پیش‌بینی و کنترل آنها از طریق روش‌های کلاسیک دشوار و حتی در بعضی موارد غیرممکن است، از روش‌های غیرکلاسیک با ویژگی‌هایی همچون هوشمندی، مبتنی بر شناخت و خبرگی استفاده می‌شود. شبکه‌های عصبی، یکی از این روش‌های در حال تحول است که در موضوعات متنوعی از قبیل الگوسازی، شناخت الگو، خوشه‌بندی و پیش‌بینی به کار رفته و نتایج امیدبخشی به همراه داشته است.

هوارنگ<sup>۱</sup> و همکارانش در ۲۰۰۵ با استفاده از الگوریتم SVM به پیش‌بینی روند شاخص، در بازار بورس و اوراق بهادار توکیو پرداختند. نتایج بدست آمده حاکی از برتری دقت الگوریتم SVM، نسبت به سه روش تحلیل تشخیصی خطی (LDA)<sup>۲</sup>، تحلیل تشخیصی درجه دو (QDA)<sup>۳</sup> و شبکه عصبی پس انتشار المان (EBNN)<sup>۴</sup> در پیش‌بینی روند شاخص است [۱۴]. چن و چانگ<sup>۵</sup> در ۲۰۰۶ با استفاده از سری زمانی فاز و الگوریتم ژنتیک، مدل‌های مرتبه اول و مراتب بالاتر را برای پیش‌بینی داده‌های دانشگاه آلاباما ارائه کردند [۸ و ۹]. لی<sup>۶</sup> و همکارانش در ۲۰۰۷ و ۲۰۰۸ از دو الگوریتم ژنتیک و شبیه‌سازی تبریدی برای پیش‌بینی دمای هوا و شاخص بورس تایوان، تایفکس<sup>۷</sup> استفاده کردند [۲۳ و ۲۲]. کوو<sup>۸</sup> و همکارانش در ۲۰۰۹ با استفاده از الگوریتم بهینه‌سازی ازدحام ذرات (PSO)<sup>۹</sup> توانستند دقت پیش‌بینی را برای داده‌های دانشگاه آلاباما ارتقا دهند [۱۹]. آنها توانستند در ۲۰۱۰ در [۲۰] با تغییر مدل خود و ایجاد یک مدل جدید تحت عنوان الگوریتم بهینه‌سازی ازدحام ذرات ترکیبی (HPSO)<sup>۱۰</sup> نتایج قبلی خود را بهبود بخشند. پارک<sup>۱۱</sup> و همکاران در ۲۰۱۰ مدل دو عاملی با مرتبه بالا را با کمک الگوریتم PSO برای پیش‌بینی شاخص تایفکس و کوسپی<sup>۱۲</sup> ارائه دادند. روش پیشنهادی آنها دارای دقت پیش‌بینی بالاتری نسبت به روش پیشنهادی لی و همکارانش در ۲۰۰۷ بود [۳۱]. چی جی لو<sup>۱۳</sup> و همکاران در ۲۰۱۳ با ترکیب دو الگوریتم SVR و PSO به پیش‌بینی روند شاخص بورس در سه کشور چین، تایوان و هند پرداختند و نشان دادند که مدل ترکیبی آنها، روند شاخص را با خطای کمتری نسبت به مدل SVR پیش‌بینی می‌کند [۱۱]. برناردو<sup>۱۴</sup> و همکاران در ۲۰۱۸ با ترکیب دو الگوریتم SVM و ژنتیک، الگوریتم جدیدی را پیشنهاد دادند. نتایج بدست آمده حاکی از دقت بالای الگوریتم پیشنهادی آنها در پیش‌بینی بازار فارکس است [۶]. الاحماری<sup>۱۵</sup> در ۲۰۲۰ به پیش‌بینی قیمت سه ارز مجازی (دیجیتال) بیت‌کوین، ریپل و اتریوم با

<sup>1</sup> Huarng

<sup>2</sup> Linear Discriminant Analysis (LDA)

<sup>3</sup> Quadratic Discriminant Analysis (QDA)

<sup>4</sup> Elman Backpropagation Neural Networks (EBNN)

<sup>5</sup> Chen & Chung

<sup>6</sup> Lee

<sup>7</sup> TAIFEX

<sup>8</sup> Kuo

<sup>9</sup> Particle Swarm Optimization (PSO)

<sup>10</sup> Hybrid Particle Swarm Optimization (HPSO)

<sup>11</sup> Park

<sup>12</sup> KOSPI

<sup>13</sup> Chi-Jie Lu

<sup>14</sup> Bernardo

<sup>15</sup> Alahmari

استفاده از الگوریتم SVR با سه کرنل متفاوت خطی، تابع پایه شعاعی (RBF)<sup>۱</sup> و چند جمله‌ای پرداخت. نتایج حاکی از برتری دقت پیش‌بینی الگوریتم SVR با کرنل تابع پایه شعاعی نسبت به سایر کرنل‌ها در پیش‌بینی قیمت ارزهای مجازی بود [۴]. داس<sup>۲</sup> و همکاران در ۲۰۲۰ به پیش‌بینی قیمت چند سهم مختلف در بورس بنگلادش با استفاده از الگوریتم SVR با دو کرنل خطی و تابع پایه شعاعی پرداختند. آنها برای تعیین دقت پیش‌بینی هر الگوریتم از معیار مجموع مربعات خطا استفاده کردند. نتایج حاکی از برتری دقت پیش‌بینی الگوریتم SVR با کرنل خطی (با دقت حدود ۸۲٪ / ۹۶) نسبت به کرنل تابع پایه شعاعی (با دقت حدود ۹۷٪ / ۰۶) است [۱۳].

از سوی دیگر سری‌های زمانی فرکتالی به صورت گسترده از ۱۹۵۰ توسط هرست<sup>۳</sup> و در دهه ۶۰ توسط مندلیبرات<sup>۴</sup> مورد مطالعه قرار گرفت [۱۵ و ۲۶]. سری‌های زمانی فرکتالی نوعی از فرکتال‌های تصادفی‌اند که از ویژگی خود تشابهی آماری نسبت به زمان برخوردارند و به دلیل اشتراک بسیار زیاد با پدیده‌ها و فرآیندهای طبیعی وابسته به زمان، مطالعه بر روی آنها طیف گسترده‌تری از مسائل از جمله مسائل اقتصادی را در بر می‌گیرد. به عنوان نمونه خالوزاده و همکاران در سال ۱۳۷۵ ه.ش (۱۹۹۶ میلادی) به بررسی قیمت سهام شرکت شهید ایران با نماد غشید در بازار بورس ایران پرداختند و نشان دادند که سری قیمت سهام این شرکت، ماهیت تصادفی دارد و میانگین حافظه بلند مدت در این شرکت پنجاه روز است [۲]. مانتنا<sup>۵</sup> و همکارانش در سال‌های ۱۹۹۶ تا ۱۹۹۸ پژوهش‌هایی در زمینه خواص فرکتالی بازار ارز انجام دادند [۲۸]. ماتسوشیتا<sup>۶</sup> ساختار فرکتالی دلار را مورد بررسی قرار داد و نشان داد که از ساختار مثلث سرپینسکی<sup>۷</sup> تبعیت می‌کند و نوسانات قیمتی آن در کوتاه‌مدت مشابه نوسانات آن در بلندمدت است [۲۹]. سرلیتین و شینتانی<sup>۸</sup> در ۲۰۰۳ با استفاده از توان لیاپانوف، به بررسی روند شاخص داو جونز<sup>۹</sup> در دوره زمانی ۲۰۰۰-۱۹۲۸ پرداختند و نشان دادند که سری زمانی این شاخص از یک روند تصادفی پیروی می‌کند [۳۳]. توماس<sup>۱۰</sup> در ۲۰۰۷ در پژوهشی به بررسی دو بازار مهم کشور چین یعنی شانگهای و شانزن پرداخت و با استفاده از نمای هرست، تغییرات شاخص سهام در این دو بازار را با هم مقایسه کرد و به این نتیجه رسید که بازار شانگهای نسبت به بازار شانزن پایدارتر است [۳۴]. چن و ایسا<sup>۱۱</sup> در ۲۰۱۱ با استفاده از نمای هرست به بررسی بازارهای مالزی و هندوستان در طی سال‌های ۲۰۰۷-۱۹۹۰ پرداختند. آنها این دوره زمانی را به سه دوره قبل از بحران مالی آسیا (۱۹۹۶-۱۹۹۰)، دوره بحران مالی آسیا (۱۹۹۸-۱۹۹۶) و دوره بعد از بحران مالی آسیا (۲۰۰۷-۱۹۹۹) تقسیم کردند و نشان دادند که هیچ حافظه بلندمدتی در بازارهای مالزی و هندوستان پس از بحران مالی آسیا وجود ندارد [۱۰]. ژای<sup>۱۲</sup> در ۲۰۱۵ برای پیش‌بینی مقدار بار الکتریکی، از درونیایی فرکتال استفاده کرد و نشان داد که این روش دارای مزایایی از جمله دقت بالا، پیاده‌سازی آسان و تعداد نقاط درونیایی پایین است [۳۹]. وانگ<sup>۱۳</sup> و همکاران در ۲۰۱۸ به بررسی بازار شانگهای چین پرداختند. آنها ابتدا توسط

<sup>1</sup> Radial Basis Function (RBF)

<sup>2</sup> Das

<sup>3</sup> Hurst

<sup>4</sup> Mandelbrot

<sup>5</sup> Mantegna

<sup>6</sup> Matsushita

<sup>7</sup> Sierpinski Triangle

<sup>8</sup> Serletini & Shintani

<sup>9</sup> Dow Jones

<sup>10</sup> Thomas

<sup>11</sup> Chin & Isa

<sup>12</sup> Zhai

<sup>13</sup> Wang

الگوریتم SVM به پیش‌بینی داده‌ها پرداخته و سپس با استفاده از الگوریتم درونیابی فرکتال، نقاط جدید را جهت پیش‌بینی نهائی روند شاخص تولید کردند و نشان دادند که روش پیشنهادی آنها می‌تواند روند شاخص سهام بازار شانگهای را فقط در یک دوره کوتاه مدت چهار روزه تشخیص دهد [۳۸].

با توجه به پایداری نسبی بازار بورس ایران در سال ۱۳۹۸ ه.ش (۲۰۱۹ میلادی)، در این مقاله به بررسی و پیش‌بینی روند شاخص صنایع مختلف با استفاده از الگوریتم تلفیقی جدید فرکسیون خواهیم پرداخت. با توجه به این که حجم محاسبات، همگرایی و دقت پیش‌بینی روش‌های هوش مصنوعی به تعداد ورودی‌های مسئله وابسته است، لذا در این مقاله ابتدا با انتخاب یک صنعت فرکتالی و فراخوانی الگوریتم IFI، به تولید داده‌های فرکتالی (بین بازه‌ای) متناظر با آن صنعت پرداخته و سپس با پیاده سازی الگوریتم SVR، بر روی داده‌های فرکتالی بدست آمده، به پیش‌بینی روند شاخص آن صنعت می‌پردازیم. نتایج عددی حاصل از پیاده سازی الگوریتم پیشنهادی بر روی صنایع فرکتالی مختلف حاکی از افزایش دقت در پیش‌بینی روند شاخص آن صنایع نسبت به روش‌های شناخته شده ANN و SVR می‌باشد.

### ۳. نظریه روش‌های اساسی

همان‌طور که اشاره شد، هدف اصلی این مقاله، تلفیق دو الگوریتم درونیابی فرکتال و SVR و ارائه یک الگوریتم پیشنهادی جدید تحت عنوان الگوریتم فرکسیون می‌باشد. بنابراین ضروری است برای درک بهتر از چگونگی عملکرد این الگوریتم پیشنهادی، ابتدا به صورت مختصر به توصیف ساختار عملکردی دو الگوریتم SVR و درونیابی فرکتال پرداخته شود.

#### ۳-۱. رگرسیون بردار پشتیبان (SVR)

الگوریتم SVR مستقیماً از نظریه یادگیری آماری وپنیک به نام SVM استخراج شده است [۳۶]. این روش جزء روش‌های آموزش نظارت‌شده است که در آن همانند بسیاری از روش‌های یادگیری ماشین، فرآیند ساخت مدل شامل دو مرحله آموزش و آزمایش است. در انتهای مرحله آموزش، قابلیت تعمیم مدل آموزش دیده با استفاده از داده‌های آزمایش مورد ارزیابی قرار می‌گیرد.

فرض کنید  $S = \{(x_i, y_i) : x_i \in R^d, y_i \in R, i = 0, 1, 2, \dots, N\}$  مجموعه داده‌های آموزشی باشد که در آن  $R^d$  نشان دهنده فضای ورودی ( $d$  بعد فضا)،  $x_i$  برداری در فضای ورودی و  $y_i$  مقدار خروجی متناظر با آن است. هدف الگوریتم SVR یافتن تابع  $f(x)$  به نحوی است که به ازای تمامی نمونه‌های آموزشی  $x_i$ ، فاصله  $f(x_i)$  از  $y_i$  حداکثر برابر  $\epsilon$  شود.

فرض کنید

(۱)

$$f(x) = w^T \cdot x + b \quad x, w \in R^d, b \in R$$

در رابطه (۱)، متغیرهای  $w$  و  $b$  باید به نحوی طراحی شوند که نه تنها به ازای تمامی نمونه‌های آموزشی  $x_i$ ، انحراف مقدار  $f(x_i)$  از  $y_i$  حداکثر به میزان  $\epsilon$  باشد بلکه میزان پیچیدگی محاسباتی مربوطه کمینه شود. برای تحقق

این هدف، به کمینه‌سازی  $\|w\|$  یا  $\|w\|^2 = w^T \cdot w$  می‌توان اشاره کرد [۳]. به عبارت دیگر، هدف در SVR کمینه‌سازی نرم اقلیدسی  $w$  است به طوری که انحراف مقدار  $f(x)$  از مقادیر داده خروجی  $y_i$  کمتر از  $\varepsilon$  باشد. این مسأله، به صورت مسأله کمینه‌سازی

$$\min \frac{\|w\|^2}{2}$$

$$s.t \quad \begin{cases} y_i - f(x_i) \leq \varepsilon & i = 0, 1, 2, \dots, N \\ f(x_i) - y_i \leq \varepsilon & i = 0, 1, 2, \dots, N \end{cases}$$

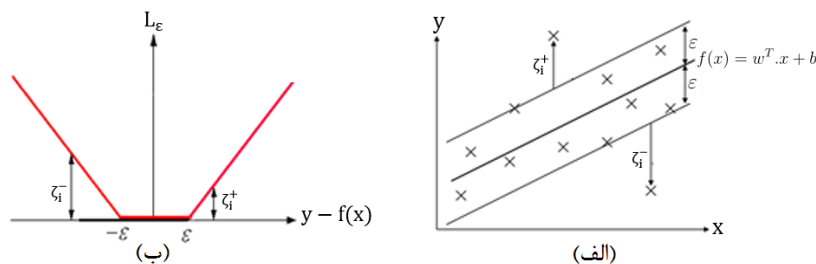
مطرح می‌گردد. فرض ضمنی در مسأله (۲) ناتهی بودن فضای جواب شدنی مسأله است و این که تابع  $f(x)$  همواره می‌تواند تمامی داده‌های آموزشی  $(x_i, y_i)$  را با دقت  $\varepsilon$  تخمین بزند. اما در بعضی مواقع، ممکن است میزان انحراف برخی از داده‌های آموزشی از مقدار حاشیه مجاز به فاصله  $\varepsilon$  بیشتر شود که برای پوشش قرار دادن تمام داده‌های آموزشی و لذا ناتهی شدن فضای شدنی مسأله (۲)، تابع زیان  $L_\varepsilon(y_i, f(x_i))$  را به صورت

$$L_\varepsilon(y_i, f(x_i)) = \begin{cases} 0 & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon & \text{در غیر صورت} \end{cases}$$

تعریف می‌کنیم. اکنون برای داده‌هایی که  $|y_i - f(x_i)| > \varepsilon$ ، متغیرهای کمکی  $\xi_i^+$  و  $\xi_i^-$  را به صورت زیر تعریف می‌کنیم (شکل ۱ مشاهده شود).

$$\begin{aligned} \xi_i^+ &= y_i - f(x_i) - \varepsilon & y_i - f(x_i) &\geq \varepsilon \\ \xi_i^- &= f(x_i) - y_i - \varepsilon & y_i - f(x_i) &\leq -\varepsilon \end{aligned}$$

تعریف ۱. ابرصفحه‌هایی با معادله  $f(x) = w^T \cdot x + b + \varepsilon$  و یا  $f(x) = w^T \cdot x + b - \varepsilon$  که فاصله داده‌های واقع بر آنها از تابع  $f(x)$  دقیقاً برابر  $\varepsilon$  است، ابرصفحه‌های حاشیه‌ای نامیده می‌شوند.



شکل ۱- تابع رگرسیون خطی و تابع خسارت

در شکل ۱ الف، تابع  $f(x)$  به همراه ابرصفحه‌های حاشیه‌ای آن، و در شکل ۱ ب نمودار تابع زیان با توجه به رابطه (۳) نشان داده شده است. با در نظر گرفتن تابع زیان و متغیرهای کمبود  $\xi_i^+$  و  $\xi_i^-$ ، مسأله کمینه‌سازی (۳) را می‌توان به صورت مسأله کمینه‌سازی

(۵)

$$\min \varphi(w, \xi_i^+, \xi_i^-) = \frac{\|w\|^2}{2} + C \left( \sum_{i=0}^N \xi_i^+ + \sum_{i=0}^N \xi_i^- \right)$$

s.t

$$\begin{aligned} y_i - f(x_i) &\leq \varepsilon + \xi_i^+ & i = 0, 1, 2, \dots, N \\ y_i - f(x_i) &\leq \varepsilon + \xi_i^- & i = 0, 1, 2, \dots, N \\ \xi_i^+, \xi_i^- &\geq 0 \end{aligned}$$

بازنویسی کرد که در آن،  $\|w\|$  بیانگر نرم بردار وزن و  $C$  یک پارامتر تنظیمی است که بسته به اهمیت کاهش انحراف، توسط کاربر به سیستم اعمال می‌شود. برای یافتن مقادیر  $w$  و  $b$  در مسأله بهینه‌سازی (۵)، از روش تابع لاگرانژ استفاده می‌کنیم. لذا با در نظر گرفتن ضرایب  $\alpha_i^+$  و  $\alpha_i^-$  و  $\mu_i^+$  و  $\mu_i^-$  به عنوان ضرایب لاگرانژ متناظر با محدودیت‌های اول تا چهارم مسأله، تابع لاگرانژ به صورت زیر تعریف می‌شود.

(۶)

$$\begin{aligned} L_P = & \frac{\|w\|^2}{2} + C \sum_{i=0}^N (\xi_i^+ + \xi_i^-) - \sum_{i=0}^N \alpha_i^+ (\varepsilon + \xi_i^+ + (w^T \cdot x_i + b) - y_i) - \\ & \sum_{i=0}^N \alpha_i^- (\varepsilon + \xi_i^- - (w^T \cdot x_i + b) + y_i) - \sum_{i=0}^N \xi_i^+ \mu_i^+ - \sum_{i=0}^N \xi_i^- \mu_i^- \end{aligned}$$

می‌توان نشان داد که جواب متناظر با تابع لاگرانژ (۶) نسبت به متغیرهای اولیه و دوگان یک نقطه زینی است [۲۴]. با توجه به شرط لازم بهینگی در نقطه زینی، مشتقات جزئی تابع لاگرانژ  $L_P$  نسبت به متغیرهای  $\xi_i^+$ ،  $\xi_i^-$ ،  $b$  و  $w$  را برابر صفر قرار می‌دهیم. لذا خواهیم داشت

(۷)

$$\begin{aligned} \frac{\partial L_P}{\partial b} = 0 &\rightarrow - \sum_{i=0}^N \alpha_i^+ + \sum_{i=0}^N \alpha_i^- = 0 \rightarrow \sum_{i=0}^N (\alpha_i^+ - \alpha_i^-) = 0 \\ \frac{\partial L_P}{\partial w} = 0 &\rightarrow w - \sum_{i=0}^N \alpha_i^+ x_i + \sum_{i=0}^N \alpha_i^- x_i = 0 \rightarrow w = \sum_{i=0}^N (\alpha_i^+ - \alpha_i^-) x_i \\ \frac{\partial L_P}{\partial \xi_i^+} = 0 &\rightarrow -\alpha_i^+ - \mu_i^+ + C = 0 \rightarrow \mu_i^+ = C - \alpha_i^+ & i = 0, 1, 2, \dots, N \\ \frac{\partial L_P}{\partial \xi_i^-} = 0 &\rightarrow -\alpha_i^- - \mu_i^- + C = 0 \rightarrow \mu_i^- = C - \alpha_i^- & i = 0, 1, 2, \dots, N \end{aligned}$$

با بازنویسی تابع  $L_P$  بر حسب روابط (۷)، مسأله بهینه‌سازی صرفاً وابسته به متغیرهای دوگان خواهد شد.

(۸)

$$L_D = -\varepsilon \sum_{i=0}^N (\alpha_i^+ + \alpha_i^-) + \sum_{i=0}^N y_i (\alpha_i^+ - \alpha_i^-) - \frac{1}{2} \sum_{i=0}^N \sum_{j=0}^N (\alpha_i^+ - \alpha_i^-) (\alpha_j^+ - \alpha_j^-) (x_i \cdot x_j)$$

اکنون با بیشینه‌کردن تابع  $L_D$  و در نظر گرفتن محدودیت‌های (۹)، مقادیر ضرایب  $\alpha^+$  و  $\alpha^-$  بدست می‌آیند.

$$\sum_{i=0}^N \alpha_i^+ = \sum_{i=0}^N \alpha_i^-$$

$$0 \leq \alpha_i^+ \leq C \quad i = 0, 1, 2, \dots, N$$

$$0 \leq \alpha_i^- \leq C \quad i = 0, 1, 2, \dots, N$$

مسئله بهینه‌سازی (۸) یک مسئله برنامه‌ریزی درجه دوم (QP) است. پس از حل مسئله بهینه‌سازی (۸) با در نظر گرفتن محدودیت‌های (۹) و یافتن مقادیر بهینه  $\alpha_i^+$  و  $\alpha_i^-$ ، مقدار بهینه بردار وزن را با استفاده از رابطه (۷) به صورت زیر محاسبه می‌کنیم.

(۱۰)

$$w = \sum_{i=0}^N (\alpha_i^+ - \alpha_i^-) x_i$$

برای محاسبه مقدار متغیر  $b$  از داده‌های موجود روی ابر صفحه‌های حاشیه‌ای استفاده می‌کنیم. از آن‌جا که انحراف این داده‌ها از تابع  $f(x)$  دقیقاً برابر  $\varepsilon$  است، لذا داریم

(۱۱)

$$y_i - f(x_i) = \varepsilon \rightarrow y_i - (w^T \cdot x_i + b) = \varepsilon \rightarrow b = y_i - w^T \cdot x_i - \varepsilon$$

$$y_i - f(x_i) = -\varepsilon \rightarrow y_i - (w^T \cdot x_i + b) = -\varepsilon \rightarrow b = y_i - w^T \cdot x_i + \varepsilon$$

از آن‌جا که به ازای داده‌های روی ابر صفحه‌های حاشیه‌ای، مقادیر مختلفی از  $b$  حاصل می‌شود، لذا در نهایت میانگین این مقادیر را به عنوان مقدار نهائی  $b$  تعریف می‌کنیم (برای مطالعه بیشتر به مرجع [۱۲] مراجعه شود).

### ۳-۲. معیار تحلیل R/S

بازارهای سهام نقشی اساسی در تامین و تخصیص بهینه منابع مالی و در نتیجه در رونق اقتصادی کشورها ایفا می‌نمایند. بنابراین پایداری این بازارها برای همه عوامل بازار، خصوصاً سرمایه‌گذاران و سیاستمداران دارای اهمیت فراوانی است. به طور کلی می‌توان گفت که حساسیت و ایجاد نوسانات مختلف اقتصادی یا غیراقتصادی در یک بازار پایدار نسبت به یک بازار ناپایدار کمتر می‌باشد که باعث می‌شود با توجه به ریسک کم سرمایه‌گذاری در بازارهای پایدار، سرمایه‌گذاران تمایل بیشتری به فعالیت و سرمایه‌گذاری در این بازارها داشته باشند.

<sup>1</sup> Quadratic Programming (QP)

از زمان پیدایش بازارهای سهام، روش‌های متنوعی برای شناسایی بازارهای پایدار مطرح شده‌اند که یکی از این روش‌ها، نظریه بازار فرکتالی است که توسط پیترز<sup>۱</sup> مطرح شد و در ادامه به شرح آن خواهیم پرداخت. با توجه به این نظریه، اجسام دارای ساختار فرکتالی، پایدارتر هستند.

برای استفاده از درونیابی فرکتال ابتدا باید از وجود پایداری در بازار مالی مورد بحث اطمینان حاصل نمود. برای این منظور در این مقاله از شاخصی تحت عنوان "نمای هرست" استفاده می‌شود. نمای هرست معیاری است که به اندازه‌گیری پایداری یا حافظه بلندمدت در سری‌های زمانی می‌پردازد. برای محاسبه مقدار نمای هرست روش‌های مختلفی ارائه شده که یکی از آنها، استفاده از معیار تحلیل R/S است. این معیار نخستین بار بر مبنای مطالعات هرست، در مورد مطالعه‌ی پدیده‌های طبیعی مانند خصوصیات هیدرولوژیکی حوزه نیل، در ۱۹۵۱ مطرح شد [۱۵]. در بازارهای مالی از این معیار برای تمییز دادن سیستم‌های تصادفی از غیرتصادفی، پایداری روندها و طول دوره چرخه‌ها استفاده می‌شود [۱]. مقدار دامنه R در این معیار برابر اختلاف مقادیر کمینه و بیشینه انحراف از میانگین تجمعی سری‌های زمانی است. هرست برای نرمال سازی مقدار دامنه R نسبت به نوسانات ورودی‌های سری زمانی، از انحراف معیار سری زمانی استفاده کرد و معیار تحلیل R/S را در یک دوره زمانی مشخص تعریف کرد [۲۵].

فرآیند محاسبه نمای هرست (H) با استفاده از معیار تحلیل R/S در قالب الگوریتم ۱ معرفی می‌شود (برای آشنایی با جزئیات بیشتر به مراجع [۱۵ و ۲۶] مراجعه کنید).

#### الگوریتم ۱. محاسبه نمای هرست (H) با معیار R/S

**ورودی:** سری زمانی به طول M

**خروجی:** تعیین مقدار نمای هرست (H)

**گام ۱-** سری زمانی به طول M را به  $d$  زیرمجموعه به طول یکسان  $\tau$  تقسیم و هر زیرمجموعه را با  $Z_{\tau,m}$  ( $m = 1, 2, \dots, d$ ) نشان دهید.

به ازاء هر  $m$  از ۱ تا  $d$  گام‌های ۲ تا ۵ را تکرار کنید.

**گام ۲-** میانگین  $E_{\tau,m}$  و انحراف معیار  $S_{\tau,m}$  متناظر با زیرمجموعه  $Z_{\tau,m}$  را محاسبه کنید.

**گام ۳-** قرار دهید

(۱۲)

$$X_{\tau,m}(i) := Z_{\tau,m}(i) - E_{\tau,m}$$

**گام ۴-** سری تجمعی  $Y_{\tau,m}$  متناظر با  $X_{\tau,m}$  را محاسبه کنید.

(۱۳)

$$Y_{\tau,m}(j) = \sum_{i=1}^j X_{\tau,m}(i), \quad m = 1, 2, \dots, d, \quad j = 1, 2, \dots, \tau$$

<sup>1</sup> Peters

گام ۵- اختلاف بین بیشترین و کمترین مقدار  $Y_{\tau,m}$  را محاسبه کنید.

(۱۴)

$$R_{\tau,m} = \max_{i=1,2,\dots,\tau} \{Y_{\tau,m}(i)\} - \min_{i=1,2,\dots,\tau} \{Y_{\tau,m}(i)\}$$

گام ۶- به ازاء هر  $Z_{\tau,m}$ ، نسبت  $(\frac{R_{\tau,m}}{S_{\tau,m}})$  را محاسبه کنید و سپس به ازاء هر مقدار  $\tau$ ،  $(R/S)_{\tau}$  را برابر میانگین مقادیر  $(\frac{R_{\tau,m}}{S_{\tau,m}})$  بدست آمده قرار دهید.

گام ۷- مقادیر  $(R/S)_{\tau}$  را بر حسب  $\tau$  در یک نمودار لگاریتمی رسم کنید.

گام ۸- با استفاده از روش رگرسیون حداقل مربعات، مجموعه نقاط بدست آمده را توسط یک خط برازش دهید و سپس شیب خط برازش شده را به عنوان نمای هرست ( $H$ ) در نظر بگیرید.

شایان ذکر است که دامنه تغییرات  $H$  در بازه  $(0,1)$  است که این بازه به سه بخش زیر تقسیم می‌شود (برای مطالعه بیشتر به مراجع [۱۲ و ۲۳] مراجعه شود).

الف)  $H=0.5$  به این معنی است که هر جمله از این سری مستقل از جمله بعدی است و لذا سری زمانی کمیت مورد مطالعه غیرقابل پیش‌بینی است.

ب) اگر  $H \in (0,0.5)$ ، سری را ناپایدار گوئیم که مشخصه اصلی آن نوسانات پرتلاطم است یعنی ممکن است در بخشی از سری، روند مثبت و در بخش دیگر روند منفی مشاهده گردد.

ج) اگر  $H \in (0.5,1)$ ، سری را پایدار می‌گوئیم که مشخصه اصلی آن حافظه بلندمدت است، به این معنی که نوسانات سری در زمان حال، با نوسانات آن در آینده مطابقت دارد. برای مثال در مورد  $H=0.6$  به میزان ۶۰ درصد این احتمال وجود دارد که اگر حرکت قبلی در جهت مثبت بوده حرکت بعدی نیز در جهت مثبت باشد.

مثال ۱. در این مثال نحوه محاسبه نمای هرست  $H$  را برای هشت داده تصادفی زیر با توجه به الگوریتم ۱ بیان کنیم.

$$\{0.04, 0.02, 0.05, 0.08, 0.021, -0.17, 0.05, 0.0001\}$$

گام ۱- سری زمانی را به  $d = 1.2.4$  زیرمجموعه  $Z_{\tau,m}$  با طول یک‌سان  $\tau = 8,4,2$  تقسیم می‌کنیم (ستون اول جدول ۱ ملاحظه شود).

گام ۲- میانگین  $E_{\tau,m}$  و انحراف معیار  $S_{\tau,m}$  متناظر با هر زیرمجموعه  $Z_{\tau,m}$  را به دست می‌آوریم (ستون ۲ و ۳ در جدول ۱ ملاحظه شود).

جدول ۱- میانگین و انحراف معیار زیرمجموعه  $Z_{\tau,m}$  در مثال ۱

زیرمجموعه $Z_{\tau,m}$	$E_{\tau,m}$	$S_{\tau,m}$
$Z_{8,1} = \{0.04, 0.02, 0.05, 0.08, 0.021, -0.17, 0.05, 0.0001\}$	0.011	0.08
$Z_{4,1} = \{0.04, 0.02, 0.05, 0.08\}$ $Z_{4,2} = \{0.021, -0.17, 0.05, 0.0001\}$	0.047 -0.024	0.03 0.1
$Z_{2,1} = \{0.04, 0.02\}$ $Z_{2,2} = \{0.05, 0.08\}$ $Z_{2,3} = \{0.021, -0.17\}$ $Z_{2,4} = \{0.05, 0.0001\}$	0.03 0.065 -0.074 0.025	0.01 0.02 0.13 0.04

گام ۳- با توجه به رابطه (۱۲) میانگین اعضای هر زیرمجموعه  $Z_{\tau,m}$  را از عناصر آن زیرمجموعه کم می‌کنیم (جدول ۲).

جدول ۲- محاسبه مقادیر  $X_{\tau,m}$  متناظر با هر زیرمجموعه در مثال ۱

$X_{\tau,m}(i) := Z_{\tau,m}(i) - E_{\tau,m}$	$X_{8,1} = \{0.029, 0.009, 0.039, 0.069, 0.01, -0.181, 0.039, -0.011\}$
	$X_{4,1} = \{-0.008, -0.028, 0.003, 0.033\}$ $X_{4,2} = \{0.046, -0.145, 0.075, 0.024\}$
	$X_{2,1} = \{0.01, -0.01\}$ $X_{2,2} = \{-0.015, 0.015\}$ $X_{2,3} = \{0.096, -0.096\}$ $X_{2,4} = \{0.025, -0.025\}$

گام ۴- با توجه به رابطه (۱۳) سری تجمعی هر زیرمجموعه  $X_{\tau,m}$  را محاسبه می‌کنیم (جدول ۳ ملاحظه شود).

جدول ۳- سری تجمعی اعضای زیرمجموعه‌های  $X_{\tau,m}$  در جدول ۲

$Y_{\tau,m}(j) = \sum_{i=1}^j X_{\tau,m}(i)$	$Y_{8,1} = \{0.029, 0.037, 0.076, 0.144, 0.154, -0.027, 0.011, 0\}$
	$Y_{4,1} = \{-0.008, -0.035, -0.033, 0\}$ $Y_{4,2} = \{0.046, -0.099, -0.024, 0\}$
	$Y_{2,1} = \{0.01, 0\}$ $Y_{2,2} = \{-0.015, 0\}$ $Y_{2,3} = \{0.096, 0\}$ $Y_{2,4} = \{0.025, 0\}$

گام ۵- دامنه  $R_{\tau,m}$  را بر اساس اختلاف بین بیشینه و کمینه مقدار  $Y_{\tau,m}(i)$  محاسبه می‌کنیم.

گام ۶- برای هر زیرمجموعه  $Z_{\tau,m}$ ، نسبت  $\left(\frac{R_{\tau,m}}{S_{\tau,m}}\right)$  را محاسبه می‌کنیم و سپس به ازاء هر مقدار  $\tau$ ،  $(R/S)_{\tau}$  را برابر میانگین مقادیر  $\left(\frac{R_{\tau,m}}{S_{\tau,m}}\right)$  به ازای  $m$  های مختلف قرار می‌دهیم (جدول ۴ ملاحظه شود).

جدول ۴-نسبت‌های  $\left(\frac{R_{\tau,m}}{S_{\tau,m}}\right)$  محاسبه شده در مثال ۱

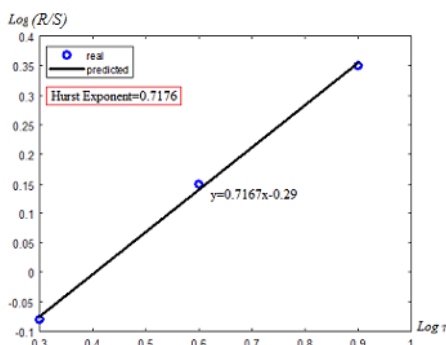
$\tau$	زیرمجموعه $Z_{\tau,m}$	$R_{\tau,m}$	$S_{\tau,m}$	$\left(\frac{R_{\tau,m}}{S_{\tau,m}}\right)$	$(R/S)_{\tau}$
8	$Z_{8,1}$	$R_{8,1} = 0.181$	$Z_{8,1} = 0.08$	2.2	2.2
4	$Z_{4,1}$	$R_{4,1} = 0.035$	$Z_{4,1} = 0.03$	1.3	1.4
	$Z_{4,2}$	$R_{4,2} = 0.145$	$Z_{4,2} = 0.1$	1.5	
2	$Z_{2,1}$	$R_{2,1} = 0.01$	$Z_{2,1} = 0.01$	1	0.8
	$Z_{2,2}$	$R_{2,2} = 0.015$	$Z_{2,2} = 0.02$	1	
	$Z_{2,3}$	$R_{2,3} = 0.096$	$Z_{2,3} = 0.13$	0.7	
	$Z_{2,4}$	$R_{2,4} = 0.025$	$Z_{2,4} = 0.04$	0.5	

گام ۷- مقادیر  $(R/S)_{\tau}$  را بر حسب  $\tau$  در یک نمودار لگاریتمی رسم کنید.

گام ۸- با استفاده از روش رگرسیون حداقل مربعات، مجموعه نقاط به دست آمده در جدول ۵ را توسط یک خط برازش دهید و سپس شیب خط برازش شده را به عنوان نمای هرست ( $H$ ) معرفی می‌کنیم (شکل ۲).

جدول ۵-نمونه محاسبه لگاریتم مقادیر  $(R/S)_{\tau}$  و  $\tau$  در مثال ۱

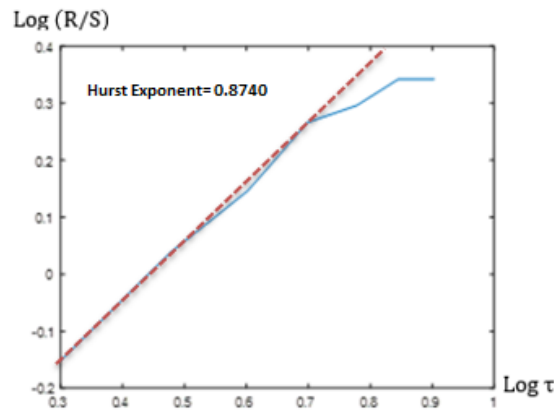
$\tau$	$(R/S)_{\tau}$	$\text{Log } \tau$	$\text{Log } (R/S)_{\tau}$
8	2.2	0.9	0.35
4	1.4	0.6	0.15
2	0.8	0.3	-0.08



شکل ۲- محاسبه مقدار نمای هرست در مثال ۱ با استفاده از الگوریتم ۱

با توجه به شکل ۲ ملاحظه می‌شود که نمای هرست برابر با  $H = 0/7167$  می‌باشد.

**مثال ۲.** برای محاسبه نمای هرست  $H$  برای یک نمونه واقعی، مقادیر شاخص صنعت بانکداری در سال ۹۸ (از تاریخ ۱۳۹۸/۰۱/۰۵ لغایت ۱۳۹۸/۱۲/۲۸ برگرفته شده از سایت شرکت مدیریت فناوری بورس تهران<sup>۱</sup> را به عنوان ورودی در نظر می‌گیریم. با استفاده از الگوریتم ۱، نمای هرست ( $H$ ) محاسبه شده این صنعت و خط برازش شده مربوطه در شکل ۳ آورده شده است.



شکل ۳- برآورد نمای هرست متناظر با شاخص "صنعت بانکداری" در بازار بورس ایران

در این شکل مقدار  $H$  متناظر با شاخص صنعت بانکداری برابر با  $H = 0.8740 > 0.5$  می‌باشد که بیانگر فرکتالی بودن شاخص این صنعت است.

#### ۴. الگوریتم درونیابی فرکتال اصلاح شده

در بازارهای مالی مانند بازار سهام، قیمت‌ها همواره به صورت لحظه‌ای تغییر می‌کند و به همین دلیل نیاز به تحلیل، ارزیابی و پیش‌بینی بازارهای مالی از دغدغه‌های اصلی دانشمندان حوزه‌های مالی، اقتصادی و به ویژه ریاضیات کاربردی است. در بخش‌های اقتصادی، به دلیل تعدد داده‌های نایستا و نامنظم، برای برازش داده‌های گم‌شده و پیش‌بینی روند کوتاه مدت داده‌ها اغلب از درونیابی فرکتال استفاده می‌شود. مفهوم درونیابی فرکتال برای اولین بار توسط برنزی<sup>۲</sup> و بر پایه سیستم‌های توابع تکرار شونده معرفی شد [۵].

در الگوریتم درونیابی فرکتال برای یافتن نقاط درونیابی، روش‌های مختلفی نظیر الگوریتم مازل و هایس<sup>۳</sup>، استفاده از اکستریم محلی [۷]، اکستریم محلی مقید شده [۴۰] و غیره وجود دارد که در این مقاله برای یافتن نقاط درونیابی از الگوریتم مازل و هایس استفاده می‌شود.

پس از یافتن نقاط درونیابی، انتخاب تابع درونیاب از اهمیت بسیاری برخوردار است. در الگوریتم درونیابی فرکتال از توابع درونیابی مختلفی نظیر توابع خطی، توابع چند جمله‌ای، توابع کسری خطی، توابع آفین و توابع آفین برشی می‌توان

<sup>1</sup> <http://www.fipiran.com>

<sup>2</sup> Barnsley

<sup>3</sup> Mazel & Hayes

<sup>4</sup> Shear Affine Function

استفاده کرد (برای مطالعه بیشتر به مرجع [۵] مراجعه شود)، که در این مقاله از توابع آفین برشی به عنوان تابع درونیاب استفاده می‌شود.

تعریف ۲. تابع آفین برشی دو بعدی  $F: R^2 \rightarrow R^2$  به صورت

$$F(X) = F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = AX + B,$$

تعریف می‌شود که در آن پارامترهای  $a, d, c, e, f$  همگی اعداد حقیقی هستند.

فرض می‌کنیم مجموعه داده‌های یک مسأله به صورت  $\{(x_i, y_i) \in R^2 \mid i = 0, 1, 2, \dots, N\}$  باشند که در آن  $x_0 < x_1 < \dots < x_N$  و همچنین مجموعه نقاط درونیابی که چگونگی یافتن آنها در الگوریتم ۲ توضیح داده خواهد شد، به صورت  $\{(x_n^*, y_n^*) \in R^2 \mid n = 1, 2, \dots, N\}$  و دستگاه توابع تکرار به صورت  $\{w_n, n = 1, 2, \dots, M\}$  تعریف گردد به گونه‌ای که توابع تعریف شده در این دستگاه از نوع توابع آفین برشی دو بعدی و به صورت رابطه (۱۵) باشند.

(۱۵)

$$w_n \begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} a_n & 0 \\ c_n & d_n \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} e_n \\ f_n \end{pmatrix} \quad w_n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{n-1}^* \\ y_{n-1}^* \end{pmatrix}$$

(۱۶)

$$w_n \begin{pmatrix} x_N \\ y_N \end{pmatrix} = \begin{pmatrix} x_n^* \\ y_n^* \end{pmatrix}$$

که در معادلات فوق  $w_n$ ،  $n$  آمین تابع آفین برشی،  $\begin{pmatrix} x_i \\ y_i \end{pmatrix}$  مختصات نقطه‌ای در فضای دو بعدی، پارامتر  $d_n \in [-1, 1]$  به عنوان عامل انقباض یا عامل مقیاس قائم،  $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$  و  $\begin{pmatrix} x_N \\ y_N \end{pmatrix}$  به ترتیب مختصات اولین و آخرین نقطه در مجموعه داده‌های کمیت مورد نظر،  $\begin{pmatrix} x_n^* \\ y_n^* \end{pmatrix}$  و  $\begin{pmatrix} x_{n-1}^* \\ y_{n-1}^* \end{pmatrix}$  مختصات دو نقطه درونیابی متوالی در مجموعه داده‌ها و ضرایب حقیقی  $a_n, c_n, e_n, f_n$  پارامترهای تابع آفین برشی متناظر با  $w_n$  هستند. اکنون با توجه به (۱۵) و در نظر گرفتن شرایط مرزی (۱۶) دستگاه معادلات

(۱۷)

$$\begin{aligned} a_n x_0 + e_n &= x_{n-1}^* \\ a_n x_N + e_n &= x_n^* \\ c_n x_0 + d_n y_0 + f_n &= y_{n-1}^* \\ c_n x_N + d_n y_N + f_n &= y_n^* \end{aligned}$$

حاصل می‌شود. نکتهٔ حائز اهمیت در این الگوریتم نحوهٔ محاسبهٔ عامل مقیاس قائم  $d_n$  است. برای محاسبهٔ عامل مقیاس قائم در ادبیات موضوع، روش‌های مختلفی ارائه شده است. در اکثر این روش‌ها مقدار  $d_n$  به صورت تقریبی و پس از حل یک معادله غیرخطی پیچیده بدست آمده که برای حل آن باید از روش‌های عددی استفاده کرد [۲۴ و ۳۰]. در این مقاله با معرفی یک روش جدید برای محاسبهٔ مقدار  $d_n$  در صدد رفع این مشکل می‌باشیم. بنابراین پس از

<sup>1</sup> Vertical Scaling Factor

محاسبه ضریب  $d_n$  با استفاده از رویکردی که در الگوریتم ۲ توضیح داده خواهد شد، چهار مجهول دیگر دستگاه (۱۷) به صورت زیر محاسبه می‌شوند.

(۱۸)

$$\begin{aligned} a_n &= \frac{x_n^* - x_{n-1}^*}{x_N - x_0} \\ e_n &= \frac{x_N x_{n-1}^* - x_0 x_n^*}{x_N - x_0} \\ c_n &= \frac{y_n^* - y_{n-1}^* - d_n(y_N - y_0)}{x_N - x_0} \\ f_n &= \frac{x_N y_{n-1}^* - x_0 y_n^* - d_n(x_N y_0 - x_0 y_N)}{x_N - x_0} \end{aligned}$$

با محاسبه مجهولات موجود در (۱۵)، می‌توان توابع درونیاب را در فاصله هر دو نقطه درونیاب  $(x_{n-1}^*, y_{n-1}^*)$  و  $(x_n^*, y_n^*)$  به راحتی محاسبه نمود.

در شبیه‌سازی داده‌ها با استفاده از هندسه فرکتال، دستگاه توابع تکراری به عنوان توابع تکراری بهینه به گونه‌ای انتخاب می‌شود که اختلاف میان مجموعه داده‌های اندازه‌گیری شده و مجموعه داده‌های حاصل از شبیه‌سازی با آن دستگاه توابع تکرار، کمینه شود. در تحقیقات مختلف برای محاسبه فاصله بین دو مجموعه از روش‌های مختلف نظیر فاصله منهن  $(L_1)$ ، فاصله اقلیدسی  $(L_2)$  و یا فاصله هاسدرف<sup>۱</sup> استفاده می‌شود. در این مقاله برای محاسبه فاصله دو مجموعه از یکدیگر از فاصله هاسدرف که در ذیل بیان می‌شود، استفاده می‌کنیم (برای مطالعه بیشتر به مرجع [۳۰] مراجعه شود). شایان ذکر است که با توجه به اینکه در این مقاله، هدف ارائه یک رویکرد جهت پیش‌بینی روند شاخص یک صنعت است، لذا اهمیت انتخاب نوع معیار فاصله در این‌جا مورد بررسی قرار نخواهد گرفت.

**تعریف ۴.** فرض کنید  $d(x, y)$  بیانگر فاصله اقلیدسی بین دو نقطه  $x$  و  $y$  باشد. فاصله هاسدرف مجموعه  $Z$  از مجموعه  $W$  با رابطه

(۱۹)

$$D(Z, W) = \max\{d'(x, W) \mid x \in Z\}$$

محاسبه می‌شود که در آن

(۲۰)

$$d'(x, W) = \min\{d(x, y) \mid y \in W\} \quad \forall x \in Z$$

حال از آن‌جا که فاصله هاسدرف فاقد خاصیت جابه‌جایی است، یعنی  $D(Z, W) \neq D(W, Z)$ ، لذا فاصله هاسدرف دو مجموعه  $Z$  و  $W$  برابر با  $\max\{D(Z, W), D(W, Z)\}$  تعریف می‌شود.

<sup>۱</sup> Hausdorff

اکنون پس از تعریف فاصله هاسدرف، به بیان الگوریتم درونیابی فرکتال اصلاح شده (IFI) می‌پردازیم. برای این منظور فرض کنید که  $L_{ij}$  خط واصل بین دو نقطه  $(x_i, y_i)$  و  $(x_j, y_j)$  باشد و فاصله نقطه  $(x_k, y_k)$  از خط  $L_{ij}$  را با  $d_{i,j}(x_k, y_k)$  نشان دهیم. مراحل الگوریتم IFI در قالب الگوریتم ۲ به شرح زیر می‌باشد.

### الگوریتم ۲. درونیابی فرکتال اصلاح شده (IFI)

**ورودی:** مجموعه نقاط Points شامل N نقطه داده‌ای

**خروجی:** تعیین مجموعه نقاط درونیابی FPoints و توابع درونیابی متناظر با مجموعه Points

**گام ۱-** قرار دهید  $FPoints := \emptyset$ ,  $Points := \{(x_i, y_i) \in R^2 \mid i = 0, 1, 2, \dots, N\}$  و  $k := 0$

**گام ۲-**  $FPoints := FPoints \cup \{(x_0, y_0), (x_N, y_N)\}$  را بروز کنید.

**گام ۳-**  $\mu$  را به صورت  $\mu := \{d_{0,N}(x_i, y_i) \mid i = 0, 1, 2, \dots, N-1\}$  محاسبه کنید.

**گام ۴-** قرار دهید  $n := k + 2$  و  $z := 1$ .

**گام ۵-**  $u_n$  را به صورت  $u_n := \max\{d_{k,j}(x_i, y_i) \mid k < i < j\}$  محاسبه کنید.

**گام ۶-** عامل مقیاس قائم  $d_n$  را مطابق رابطه  $d_n = \frac{u_n}{\mu}$  محاسبه کنید.

۶-۱- اگر  $|d_n| \geq 1$  آن‌گاه  $z := z + 1$  و به گام ۵ بازگردید.

۶-۲- اگر  $|d_n| < 1$  آنگاه بنابر روابط (۱۸)، سایر پارامترهای تابع آفین را در فاصله بین نقاط  $k$  و  $z$  محاسبه کنید.

**گام ۷-** تابع آفین بدست آمده را بر روی تمامی نقاط مجموعه Points اعمال کنید.

**گام ۸-** انحراف بین نقاط اندازه‌گیری شده  $y_i$  و نقاط شبیه‌سازی شده (خروجی گام ۷) را محاسبه کنید.

**گام ۹-** قرار دهید  $z := z + 1$ . اگر  $z == N$  به گام ۱۰ بروید و در غیر این صورت به گام ۵ بروید.

**گام ۱۰-** کمترین انحراف بدست آمده را انتخاب و نقطه  $(x_i, y_i)$  متناظرش را به مجموعه FPoints اضافه کنید.

**گام ۱۱-** اگر  $i == N$  به گام ۱۲ بروید در غیر این صورت قرار دهید  $k := i$  و  $n := n + 1$  و  $z := k + 2$  به گام ۵ بروید.

**گام ۱۲-** توابع درونیابی را بر روی مجموعه نقاط Points اعمال کنید و داده‌های جدید را تولید کنید.

با توجه به اینکه حجم محاسبات، همگرایی و دقت پیش‌بینی در روش‌های هوش مصنوعی مانند الگوریتم SVR، وابسته به تعداد ورودی‌های مسئله است و از آنجا که برای پیش‌بینی روند شاخص متناظر با صنایع مختلف، تغییرات شاخص هر صنعت در پایان هر روز معاملاتی به عنوان داده‌های ورودی در نظر گرفته می‌شود، لذا پس از تشخیص فرکتال بودن ساختار صنایع، ابتدا با استفاده از الگوریتم IFI (الگوریتم ۲)، به تولید داده‌های جدید فرکتالی (داده‌های بین بازه‌ای) پرداخته و سپس با فراخوانی الگوریتم SVR بر روی داده‌های بدست آمده، به پیش‌بینی روند شاخص آن صنعت خواهیم پرداخت. فرآیند پیشنهادی اشاره شده جهت پیش‌بینی روند شاخص یک صنعت را در قالب الگوریتم تلفیقی فرکسیون (الگوریتم ۳) به صورت زیر بیان می‌کنیم.

### الگوریتم ۳. الگوریتم تلفیقی فرکسیون

**ورودی:** داده‌های مربوط به شاخص یک صنعت در یک بازه زمانی خاص

**خروجی:** پیش‌بینی روند شاخص صنعت مربوطه در یک بازه زمانی کوتاه مدت یک ماهه

- گام ۱- بررسی پایداری و ساختار فرکتالی داده‌های مربوط به یک صنعت با محاسبه نمای هرست (الگوریتم ۱).
- گام ۲- تولید نقاط فرکتالی (داده‌های بین بازه‌ای) برای صنعت با ساختار فرکتالی (الگوریتم ۲).
- گام ۳- فراخوانی الگوریتم SVR برای آموزش نقاط فرکتالی گام ۲.
- گام ۴- پیش‌بینی روند کوتاه مدت شاخص صنعت بر اساس تابع رگرسیون بدست آمده از گام ۳.

## ۵. نتایج عددی

در این بخش ابتدا به بررسی ساختار فرکتالی سری زمانی شاخص هر یک از صنایع به صورت مجزا می‌پردازیم. پس از تایید ساختار فرکتالی هر صنعت، با استفاده از الگوریتم فرکسیون پیشنهادی (الگوریتم ۳)، روند کوتاه مدتی شاخص آن صنعت را پیش‌بینی می‌کنیم. در پایان به منظور ارزیابی کارایی الگوریتم پیشنهادی در پیش‌بینی روند شاخص صنایع، به مقایسه عملکرد آن با دو روش مرسوم ANN و SVR خواهیم پرداخت.

### ۵-۱. پیش‌بینی کوتاه مدت شاخص صنایع مختلف بورس

در این بخش با پیاده‌سازی الگوریتم فرکسیون پیشنهادی، روند شاخص صنایع مختلف را با دید کوتاه‌مدت یک‌ماهه پیش‌بینی کنیم. همانطور که اشاره شد، پیش از فراخوانی الگوریتم فرکسیون، ضروری است فرکتالی بودن سری زمانی متناظر با شاخص هر یک از صنایع به صورت مجزا مورد بررسی قرار گیرد.

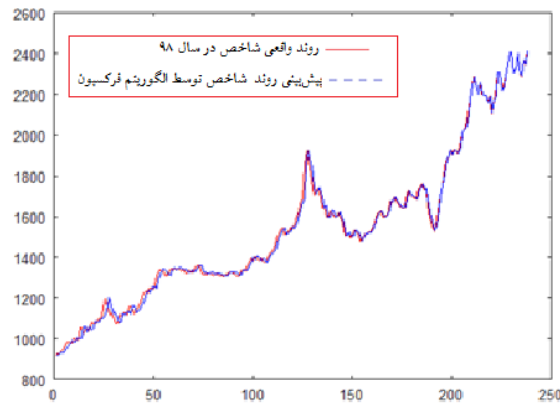
برای این مهم، روند شاخص ۹ صنعت فعال بورس ایران که حدود ۸۳/۸ درصد از کل ارزش بازار بورس ایران را تشکیل می‌دهند را در سال‌های ۹۷ و ۹۸ مورد تجزیه و تحلیل قرار می‌دهیم. نتایج مربوط به ضریب هرست به تفکیک هر صنعت در سال‌های ۹۷ و ۹۸ در جدول ۶ آورده شده است.

جدول ۶- درصد سهم صنایع مختلف از بازار بورس به همراه ضریب هرست هر صنعت در سال‌های ۹۷ و ۹۸

نام صنعت	تعداد شرکت	درصد از کل بازار	ضریب هرست در سال ۹۷	ضریب هرست در سال ۹۸
بانکها و موسسات اعتباری	۱۰	۱۷/۵۶	۰/۶۸۸۱	۰/۸۷۴۰
فلزات اساسی	۲۶	۱۶/۶۴	۰/۷۸۱۰	۰/۸۱۸۶
محصولات شیمیایی	۳۰	۱۱/۲۷	۰/۵۷۲۰	۰/۶۷۷۸
چند رشته‌ای صنعتی	۴	۹/۲۳	۰/۷۳۹۱	۰/۶۷۶۲
مخابرات	۲	۹/۱	۰/۵۳۱۷	۰/۵۴۴۰
استخراج کانه های فلزی	۹	۸/۶۸	۰/۷۹۷۹	۰/۸۱۲۶
فرآورده های نفتی	۷	۴/۷۵	۰/۷۳۷۹	۰/۶۳۶۹
صنعت خودرو	۳۱	۳/۶۴	۰/۶۱۱۰	۰/۶۵۴۹
صنعت سیمان	۳۲	۲/۹۱	۰/۸۴۵۳	۰/۷۲۷۲

بر اساس داده‌های جدول ۶، از آنجا که در هر صنعت  $H > 0.5$  می‌باشد، لذا صنایع مورد بحث، در سال‌های ۹۷ و ۹۸ دارای ساختار فرکتالی بوده و روند حرکتی شاخص هر صنعت کمیتی قابل پیش‌بینی است. بنابراین استفاده از مدل فرکتالی برای تجزیه و تحلیل نوسانات آنها تایید می‌شود. به عنوان مثال در صنعت بانکها و موسسات اعتباری با

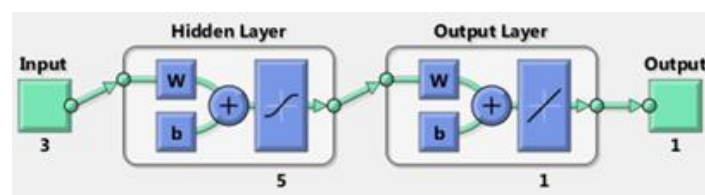
پس از اعمال الگوریتم فرکسیون بر روی داده‌های مربوطه (مقدار شاخص صنعت بانکداری در سال ۹۸)، مشاهده می‌شود که روند نوسانی شاخص این صنعت بر اساس این الگوریتم به خوبی قابل توصیف و لذا قابل پیش‌بینی است (شکل ۴ ملاحظه شود).



شکل ۴- داده‌های واقعی و داده‌های درونیابی شده متناظر با شاخص صنعت بانکداری در سال ۹۸

#### ۲-۵. معیارهای دقت پیش‌بینی

به منظور ارزیابی کارایی الگوریتم پیشنهادی در پیش‌بینی روند شاخص صنایع، به مقایسه عملکرد آن با دو روش مرسوم ANN و SVR می‌پردازیم. شبکه عصبی استفاده شده در این مقاله از نوع شبکه عصبی پرسپترون سه لایه است. تعداد نرون‌های لایه ورودی برابر با سه نرون که معادل مقدار شاخص در روز معاملاتی قبل، مقدار شاخص در دو روز معاملاتی قبل و مقدار شاخص در چهار روز معاملاتی قبل است. تعداد نرون‌های لایه پنهان برابر با پنج نرون و تعداد نرون‌های لایه خروجی برابر با یک نرون که معادل مقدار شاخص در روز معاملاتی بعد است. همچنین توابع فعالساز استفاده شده در این شبکه پرسپترون در لایه پنهان از نوع تانژانت هیپربولیک و در لایه خروجی از نوع خطی می‌باشد. ساختار شبکه عصبی استفاده شده در این مقاله، در شکل ۵ آورده شده است.



شکل ۵- ساختار شبکه عصبی مصنوعی مورد استفاده در این مقاله

در این مقاله برای مقایسه سه الگوریتم پیشنهادی فرکسیون، ANN و SVR از معیارهای معتبر شناخته شده مانند میانگین مربعات خطا (MSE)<sup>۱</sup> و مربع خطای نسبی (RSE)<sup>۲</sup> و میانگین قدر مطلق خطا (MAE)<sup>۳</sup> که به ترتیب در روابط (۲۱) تعریف شده‌اند، استفاده شده است.

<sup>۱</sup> Mean Squared Error (MSE)

<sup>۲</sup> Relative Squared Error (RSE)

<sup>۳</sup> Mean Absolute Error (MAE)

(۲۱)

$$MSE = \frac{1}{N+1} \sum_{i=1}^N (\hat{x}_i - x_i)^2$$

$$RSE = \frac{\sum_{i=1}^N (\hat{x}_i - x_i)^2}{\sum_{i=1}^N (\bar{x}_i - x_i)^2}$$

$$MAE = \frac{1}{N+1} \sum_{i=1}^N |\hat{x}_i - x_i|,$$

که در آن  $\bar{x}_i$ ،  $\hat{x}_i$  و  $x_i$  به ترتیب برابر با میانگین داده‌های اندازه‌گیری شده، مقدار داده‌های پیش‌بینی شده و مقدار واقعی داده‌های اندازه‌گیری شده می‌باشند. در این مقاله به منظور سهولت در مقایسه عملکرد روش‌ها، از انحراف نسبی هر یک از این معیارها استفاده می‌کنیم.

به عنوان مثال فرض می‌کنیم که مقدار معیار RSE در سه الگوریتم فرکسیون، SVR و ANN به ترتیب برابر  $\alpha$  و  $\beta$  و  $\gamma$  باشد. با تعریف  $m = \min\{\alpha, \beta, \gamma\}$  مقدار انحراف نسبی RSE نسبت به بهترین مقدار بدست آمده RSE در سه الگوریتم که در این مقاله آن را با نماد  $\overline{RSE}$  مشخص می‌کنیم، برای هر روش به صورت زیر محاسبه می‌شود.

(۲۲)

$$\overline{RSE} = \frac{\alpha - m}{m} \quad \text{الگوریتم فرکسیون}$$

$$\overline{RSE} = \frac{\beta - m}{m} \quad \text{الگوریتم SVR}$$

$$\overline{RSE} = \frac{\gamma - m}{m} \quad \text{الگوریتم ANN}$$

محاسبه انحراف نسبی متناظر با دو معیار دیگر یعنی MSE و MAE نسبت به بهترین مقدار بدست آمده در سه الگوریتم که با  $\overline{MAE}$  و  $\overline{MSE}$  نمایش می‌دهیم، به طور مشابه انجام می‌شود.

### نرمال‌سازی داده‌ها

با توجه به این که کار کردن با داده‌ها به صورت خام باعث کاهش سرعت و دقت پیش‌بینی می‌شود، لذا ابتدا بایستی ورودی‌ها و در بعضی از موارد خروجی‌ها نرمال شوند. برای نرمال کردن داده‌های ورودی از رابطه (۲۳) استفاده می‌کنیم.

(۲۳)

$$x_N = \frac{x - x_{min}}{x_{max} - x_{min}},$$

در این رابطه  $x_{min}$  و  $x_{max}$  به ترتیب مقدار کمینه و بیشینه داده‌های ورودی و  $x_N$  بیانگر داده نرمال شده متناظر با ورودی  $x$  می‌باشد.

در ادامه به منظور تجزیه و تحلیل دقیق‌تر عملکرد الگوریتم فرکسیون پیشنهادی در میزان دقت پیش‌بینی روند شاخص صنایع، عملکرد این الگوریتم با عملکرد دو الگوریتم دیگر یعنی ANN و SVR بر روی داده‌های مربوط به شاخص ۹ صنعت مختلف در سال‌های ۹۷ و ۹۸ با توجه به سه معیار ارزیابی اشاره شده به ترتیب در جداول ۷ و ۸، مورد بررسی و مقایسه قرار می‌گیرد.

جدول ۷- مقایسه دقت پیش‌بینی روند شاخص‌های مختلف با استفاده از سه الگوریتم در سال ۹۷

الگوریتم فرکسیون			SVR الگوریتم			ANN الگوریتم			نام صنعت
$\overline{MSE}$	$\overline{RSE}$	$\overline{MAE}$	$\overline{MSE}$	$\overline{RSE}$	$\overline{MAE}$	$\overline{MSE}$	$\overline{RSE}$	$\overline{MAE}$	
۰	۰	۰/۱۳	۰/۲۲	۰/۵۱	۰	۳/۱۶	۱/۳۷	۰/۶۹	بانک‌ها و مؤسسات اعتباری
۰	۰	۰	۰/۲۷	۱/۱۵	۰/۲۳	۰/۴۵	۱/۹۸	۰/۵۳	فلزات اساسی
۰	۰/۵۶	۰/۵۸	۰/۶۱	۰	۰	۰/۵۵	۰/۵۹	۰/۲۰	محصولات شیمیایی
۰	۰	۰	۰/۵۰	۰/۱۶	۰/۱۶	۰/۵۰	۰/۱۷	۰/۱۸	چند رشته‌ای صنعتی
۰	۰	۰	۰/۲۹	۰/۴۷	۰/۴۴	۰/۶۱	۱/۲۹	۰/۶۲	مخابرات
۰	۰	۰	۰/۴۰	۰/۲۷	۰/۵۹	۰/۶۰	۰/۶۱	۰/۲۶	استخراج کانه‌های فلزی
۰	۰	۰	۰/۳۶	۰/۴۹	۰/۱۴	۰/۹۱	۱/۷۵	۰/۳۴	فرآورده‌های نفتی
۰/۴۴	۰/۵۴	۰/۱۷	۰	۰	۰	۰/۱۳	۰/۳۷	۰/۱۹	صنعت خودرو
۰	۰	۰	۰/۳۵	۰/۹۴	۰/۳۵	۰/۸۳	۱/۵۱	۰/۷۳	صنعت سیمان
۰	۰	۰	۰/۳۹	۰/۳۷	۰/۱۸	۱/۵۲	۱/۳۰	۰/۷۲	میانگین انحرافات

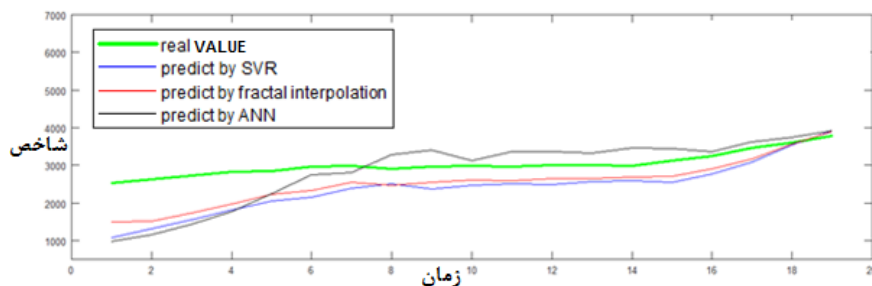
نتایج عددی جدول ۷ نشان می‌دهد که الگوریتم فرکسیون پیشنهادی جز در صنعت خودرو و ساخت قطعات در سایر صنایع از عملکرد بهتری در میزان دقت پیش‌بینی روند شاخص نسبت به دو الگوریتم دیگر، یعنی ANN و SVR برخوردار است.

جدول ۸- مقایسه دقت پیش‌بینی روند شاخص‌های مختلف با استفاده از سه الگوریتم در سال ۹۸

الگوریتم فرکسیون			الگوریتم SVR			الگوریتم ANN			نام صنعت
$\overline{MSE}$	$\overline{RSE}$	$\overline{MAE}$	$\overline{MSE}$	$\overline{RSE}$	$\overline{MAE}$	$\overline{MSE}$	$\overline{RSE}$	$\overline{MAE}$	
۰	۰	۰	۰/۰۲	۰/۳۳	۰/۰۵	۱/۶۵	۲/۴۳	۰/۵۹	بانکها و مؤسسات اعتباری
۰	۰	۰	۰/۱	۰/۰۹	۰/۰۵	۳/۸۷	۳/۹	۱/۱۴	فلزات اساسی
۰	۰/۰۰۴	۰	۰	۰	۰/۰۵۵	۸/۸۷	۶/۵۵	۱/۸	محصولات شیمیایی
۰	۰	۰	۰/۰۰۷	۰/۱	۰/۰۶	۳/۱	۳/۳	۰/۸۷	چند رشته‌ای صنعتی
۰	۰	۰	۰/۲۳	۰/۶۶	۰/۱۶	۲/۹	۳/۲۷	۰/۳۷	مخابرات
۰	۰	۰	۰/۹۵	۱/۱۸	۰/۳۴	۲۲/۳	۱۱/۷	۳/۰۶	استخراج کانه‌های فلزی
۰	۰	۰	۰/۰۲	۰/۰۹	۰/۰۹	۰/۰۴	۰/۱۲	۰/۱۲	فرآورده‌های نفتی
۰	۰	۰	۰/۴	۰/۷۲	۰/۱۸	۵/۲	۱۱/۴	۳/۶۸	صنعت خودرو
۰	۰	۰	۰/۱۴	۰/۳۴	۰/۱	۰/۸۴	۴/۹	۱/۵۸	صنعت سیمان
۰	۰	۰	۰/۲	۰/۳۹	۰/۱۱	۵/۴۱	۵/۲۸	۱/۴۷	میانگین انحرافات

نتایج عددی جدول ۸ حاکی از برتری عملکرد الگوریتم پیشنهادی نسبت به دو الگوریتم دیگر، صرف نظر از نوع صنعت مورد مطالعه، در میزان دقت پیش‌بینی روند شاخص صنایع در سال ۹۹ است.

بر این اساس، به منظور درک بهتر از نحوه عملکرد هر یک از سه الگوریتم، به عنوان نمونه میزان دقت پیش‌بینی روند شاخص صنعت بانکداری توسط سه الگوریتم مورد بررسی دقیق‌تر قرار می‌گیرد. با توجه به نتایج جدول ۸، در شکل ۶، پیش‌بینی روند شاخص صنعت بانکداری در یک ماهه نخست سال ۹۹ توسط سه الگوریتم ANN، SVR و الگوریتم فرکسیون پیشنهادی نسبت به روند واقعی شاخص این صنعت ترسیم شده‌اند. همانطور که مشاهده می‌شود، الگوریتم فرکسیون روند آینده شاخص این صنعت را دقیق‌تر از دو روش مرسوم ANN و SVR پیش‌بینی می‌کند.



شکل ۶- مقادیر واقعی و تخمین زده شده شاخص صنعت بانکداری در ماه نخست ۱۳۹۹

## ۶. نتیجه‌گیری

با توجه به ساختار دینامیکی و نوسانات غیرخطی بازار سهام، پیش‌بینی دقیق روند این بازار با استفاده از روش‌های قدیمی دشوار است. با پیشرفت سریع نظریه هوش مصنوعی و روش‌های یادگیری ماشین در سال‌های اخیر، برخی از روش‌های هوشمند، مانند شبکه عصبی مصنوعی (ANN) و رگرسیون بردار پشتیبان (SVR) به طور گسترده‌تری برای پیش‌بینی سری‌های زمانی مالی مورد استفاده قرار گرفتند. از آنجا که حجم محاسبات، همگرایی و دقت پیش‌بینی در روش‌های هوش مصنوعی به تعداد ورودی‌های مسئله بستگی دارد، لذا در این مقاله، مقدار شاخص هر صنعت در پایان هر روز معاملات، به عنوان داده‌های اولیه برای پیش‌بینی روند شاخص در نظر گرفته شد. پس از تشخیص ساختار فرکتالی روند شاخص صنایع مورد بررسی، با استفاده از الگوریتم درونیابی فرکتال اصلاح شده، داده‌های جدید فرکتالی (داده‌های بین بازه‌ای) تولید شد و در پایان با فراخوانی الگوریتم SVR بر روی داده‌های فرکتالی بدست آمده، به پیش‌بینی روند شاخص صنایع مورد نظر پرداخته شد. در این مقاله به منظور ارزیابی کارایی الگوریتم فرکسیون پیشنهادی، با توجه به داده‌های سال ۹۷ و ۹۸ به پیش‌بینی روند شاخص ۹ صنعت بزرگ و فعال در بازار بورس و اوراق بهادار تهران به ترتیب در یک ماه نخست سال‌های ۹۸ و ۹۹ بر اساس این الگوریتم پیشنهادی پرداخته شد. نتایج بدست آمده صرف نظر از نوع صنعت مورد مطالعه، حاکی از برتری عملکرد الگوریتم پیشنهادی نسبت به دو الگوریتم شناخته شده ANN و SVR در میزان دقت پیش‌بینی روند شاخص صنایع می‌باشد. در تحقیقات آتی می‌توان به بررسی کارایی روش فرکسیون در پیش‌بینی روند قیمت طلا، ارزهای دیجیتال و یا سهام پرداخت. همچنین می‌توان به دنبال راهکاری جهت استفاده از نتایج الگوریتم فرکسیون در نشانگرها یا نوسانگرها به منظور تعیین نقاط برگشت بازار، سطوح حمایت و مقاومت یک شاخص بود.

## References

۱. مروت ح.، آزمون فرضیه بازار فرکتالی در بورس اوراق بهادار تهران، بورس اوراق بهادار، شماره ۱۹، (۱۳۹۱)، ص ۲۵-۵.
۲. خالوزاده، ح.، خاکی صدیق، ع.، لوکس، ک.، آیا قیمت سهام در بازار بورس تهران قابل پیش بینی است؟ (کاربرد موردی تحلیل R/S برای سهام شهد). مجله تحقیقات مالی. دانشکده مدیریت دانشگاه تهران، شماره ۱۱ (۱۳۷۵)، ص ۳۷-۴۶.
3. Abaszade, M. & Effati S. Support Vector Regression with random output variable and probabilistic constraints, *Iranian Journal of Fuzzy Systems*, vol.14, (2017), pp.43-60.
4. Alahmari, S. Predicting the Price of Cryptocurrency using Support Vector Regression Methods, *Journal of Mechanics of Continua and Mathematical Sciences*, vol.19, (2020), pp.313-322.
5. Barnsley, M. *Fractal everywhere*. 1ed New York, Academic Press,(1988).
6. Bernardo J.A., Rui Ferreira N. & Nuno Horta, Combining Support Vector Machine with Genetic Algorithms to optimize investments in Forex markets with high leverage, *Applied Soft Computing*, vol. 64, (2018) , pp. 596–613.
7. Chuanzhen, L., Xiangdong, G., & Shuping Sh. A speedup method for fractal encoding of digital signals, *Signal Processing, 5th International Conference on WCCC-ICSP*, (2000),pp.1115-1111.
8. Chen, S.M. & Chung, N.Y. Forecasting enrollments of students by using fuzzy time series and genetic algorithms, *International Journal of Information and Management Sciences*, vol.17, (2006), pp.1-17.
9. Chen, S. M. & Chung, N. Y. Forecasting enrollments using high-order fuzzy time series and genetic algorithms: Research Articles. *International Journal of Information and Management Sciences*, vol.21, (2006), pp.485-501.
10. Chin, C. & Isa, Z. A short range dependence adjusted hurst exponent evaluation for Malaysian and Indonesian financial markets. *African Journal of Business Management*, Vol.5, (2011) , pp.2644-2653.

11. Chi-Jie, Lu . Hybridizing nonlinear independent component analysis and support vector regression with particle swarm optimization for stock index forecasting, *Neural Applied & Soft computing*, vol.40, (2013), pp.164-178.
12. Cortes, C. & Vapnik V., Support vector networks, *Machine Learning*, vol.20, (1995), pp.273–297.
13. Das S., Arman M. S., Hossain, S.S. & Islam, S. Bangladeshi Stock Price Prediction and Analysis with Potent Machine Learning Approaches, *Cyber Security and Computer Science*, (2020), pp.230-240.
14. Huang, W., Nakamori, Y. & Wang, S.Y. Forecasting stock market movement direction with support vector machine, *Computers and Operations Research*, vol.32, (2005), pp.2513-2522.
15. Hurst, H.E. Long term storage capacities of reservoirs. *Transactions of the American Society of Civil Engineers*, vol.116, (1951), pp.770-799.
16. Kara, Y., Boyacioglu, M.A. & Baykan, O.K. Predicting direction of stock price index movement using artificial neural networks and support vector machines: the sample of the Istanbul Stock Exchange, *Expert Systems with Applications*, vol.38, (2011), pp.5311-5319.
17. Kim, K. Financial time series forecasting using support vector machines, *Neurocomputing* 55, (2003) , pp.307-319.
18. Kumagai, Y. Fractal structure of financial high frequency data, *Fractals*, vol.10(1), (2002), pp.13-18.
19. Kuo, H., Horng, H., Kao, S.J., Lin, T.W., Lee, T.L. & Pan, C.L. An improved method for forecasting enrollments based on fuzzy time series and particle swarm optimization. *Expert Systems with Applications*, vol.36, (2009), pp.6108-6117.
20. Kuo, H., Horng, H., Kao, S.J., Lin, T.W., Lee, T.L. & Pan, C.L. Forecasting TAIEX based on fuzzy time series and particle swarm optimization. *Expert Systems with Applications*, vol.37, (2010), pp.1494-1502.

21. Kwapien, J. & Drozd, S. Physical approach to complex systems, *Physics Reports*, vol.515, (2012), pp.115-226.
22. Lee, L.W., WANG, L.H. & CHEN, S.M. Temperature prediction and TAIEX forecasting based on fuzzy logical relationships and genetic algorithms. *Expert Systems with Applications*, vol.33, (2007), pp.539-550.
23. Lee, L.W., Wang, H.F. & Chen, S.M. Temperature prediction and TAIEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques. *Expert Systems with Applications*, vol.34, (2008), pp.328-336.
24. Li G.Z. , Huang J.B. & Huang H.Y. Calculating method of contraction operators in fractal interpolation based on the B-spline, *Journal of Ordnance Engineering College* ,vol.18(2), (2006) , pp.76-78.
25. Li J., Rescaled range (R/S) analysis on seismic activity parameters, *Acta Seismologica Sinica*, vol.14(2), (2001), pp. 148-155.
26. Mandelbrot, B. & Wallis, J. Robustness of the rescaled range R/S in the measurement of non-cyclic long run statistical dependence. *Water Resources Research*, vol.5, (1969), pp.967-988.
27. Mangasarian O.L., *Nonlinear Programming*, McGraw-Hill., New York, (1969).
28. Mantegna, R.N. & Stanley, H.E. Turbulence and Financial Markets. *Nature*, vol.383, (1996), pp.587-588.
29. Matsushita, M. & Fukiwara, H. Fractal growth in financial markets formation, *physical sciences and biology*, vol.35, (1993), pp.637-548.
30. Mazel, D. S. & Hayes, M. H. Using iterated function systems to model discrete sequences, in *IEEE Transactions on Signal Processing*, vol.40,(1992) , pp. 1724-1734.
31. Park, J. I., Lee, D. J., Song, C. K. & Chun, M.G. TAIEX and KOSPI 200 forecasting based on two-factors high-order fuzzy time series and particle swarm optimization. *Expert Systems with Applications*, vol.37, (2010), pp.959-967.

32. Peters, E.E. Fractal Market Analysis: Applying Chaos Theory to Investment and Economics, New York: John Wiley and Son Inc, vol.24, (1994).
33. Serletin. A. & Shintani M. No evidence of chaos but some evidence of dependence in US stock market, Chaos, solitons and fractals, Vol.17, (2003), pp.449-459.
34. Thomas A.T. An Empirical Analysis of the Fractal Dimension of Chinese Equity Returns, doctoral dissertation. (2007).
35. Vandebei R.J., LOQO users manual-version 3.10, Optimization Methods and Software, vol.11, (1997), pp.485-514.
36. Vapnik V. The Nature of Statistical Learning Theory. 2th ed. Springer-Verlag, (1995).
37. Wang, H.Y. & Wang, T.T. Multifractal analysis of the Chinese stock, bond and fund markets, Physica A: Statistical Mechanics and its Applications, vol.512, (2018), pp.280-292.
38. Wang H.Y. , Li H. & Shen J.Y. A Novel Hybrid Fractal Interpolation-SVM Model for Forecasting Stock Price Indexes, worldscientific, vol.27 , (2018), NO.04.
39. Zhai M.Y. A new method for short-term load forecasting based on fractal interpolation and wavelet analysis, Electrical Power and Energy Systems, vol.69, (2015), pp.241-245.
40. Ziaei, A. N., Keshavarzi, A. R., & Emdad, H. Fractal scaling and simulation of velocity components and turbulent shear stress in open channel flow, Chaos, Solitons and Fractals, vol. 24, (2005), pp.1031-1045.