



Sharif University of Technology

<https://sjie.journals.sharif.edu/>



Research Article

Investigating the Performance of Two Stock Price Forecasting Models, Based on a Long Short-Term Memory Neural Network and with Two Different Approaches of Feature Selection and Time Series Analysis

Keyvan Haghighi Naeini and Mohammad Ali Rastegar*

Department of Systems and Productivity Management, Faculty of Industrial Engineering, Tarbiat Modares University, Tehran, Iran.

* corresponding author: (ma.rastegar@modares.ac.ir)

Article Info

Article history:

Received: 29 April 2024

Revised: 14 September 2024

Accepted: 8 October 2024

Keywords:

Stock price prediction,
deep learning,
genetic algorithm,
feature selection,
time series decomposition.

Abstract

In the financial literature, the capital market plays an essential role in economic growth through the financing of enterprises, the optimal allocation of resources, the improvement of liquidity of assets, the improvement of company management, and the increase of transparency in the economy. One of the most important challenges that shareholders always face in the market is to make the right decision and be in the right position with buying and selling stocks. Stock forecasting that predicts future stock movements and benefits shareholders has been an attractive research area for financial studies and researchers for the past. In this research, we present two models of neural networks that receive inputs and predict price and movement trends, and finally, the performance of these two models is compared. The data studied in this research includes the price data of the 5 largest shares of the New York market during the years 2000 to 2020, in which 80 percent of the initial samples are used as training data and the remaining 20 percent are used as test data.

In the proposed VMD-LSTM model, first, the stock price time series is decomposed using the variational mode decomposition algorithm (VMD) into the intrinsic mode functions (IMF), and then each of these IMFs is predicted by the LSTM model and After interpreting the results. In the second proposed method, available features, including price and some of the most important technical indicators, are used to predict stock prices. In the GA-LSTM model, the genetic algorithm is first used to select the best features from the entire set of features. Then, the time series of the stock closing price was predicted by an LSTM network using the selected features. The results of the research showed that because both models are very good in price prediction, the proposed GA-LSTM model, which is developed based on feature selection, has less error and more accurate performance.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

To Cite this article:

Haghighi Naeini, K. and Rastegar, M.A. 2025. Investigating the performance of two stock price forecasting models, based on a long short-term memory neural network and with two different approaches of feature selection and time series analysis, Sharif Industrial Engineering and Management Journal, 41(1), 86-97. <https://doi.org/10.24200/sj.2024.63935.2392>



بررسی عملکرد دو مدل پیش‌بینی قیمت سهام مبتنی بر شبکه‌ی عصبی حافظه‌ی کوتاه‌مدت طولانی و با دو رویکرد متفاوت انتخاب ویژگی و تجزیه‌ی سری زمانی

کیوان حقیقی نائینی و محمدعلی رستگار*

گروه مدیریت سیستم و بهره‌وری، دانشکده‌ی مهندسی صنایع، دانشگاه تربیت مدرس، تهران، ایران.

*نویسنده مسئول (ma_rastegar@modares.ac.ir)

چکیده

امروزه توسعه‌ی مدل‌های پیش‌بینی قیمت، با توجه به گسترش روزافزون بازارهای مالی، اهمیت بسیاری دارد. در پژوهش حاضر، دو مدل از ترکیب الگوریتم ژنتیک و الگوریتم تجزیه‌ی حالت متغیر به‌صورت جداگانه با شبکه‌ی عصبی LSTM ارائه شده است. در رویکرد انتخاب ویژگی، الگوریتم ژنتیک از میان تمام ویژگی‌های ورودی مدل، زیرمجموعه‌ای از ویژگی‌ها را انتخاب و شبکه‌ی عصبی LSTM، با استفاده از ویژگی‌های منتخب، قیمت سهام را پیش‌بینی می‌کند. اما در رویکرد تجزیه‌ی سری زمانی، سری زمانی قیمت سهام به‌عنوان یگانه ورودی مدل توسط الگوریتم VMD تجزیه می‌شود و شبکه‌ی عصبی LSTM با پیش‌بینی توابع تولیدشده و تجمیع نتایج، قیمت سهام را پیش‌بینی می‌کند. در پژوهش حاضر، قیمت ۵ سهم بورس نیویورک با دو مدل پیشنهادی، پیش‌بینی و عملکرد آن‌ها مقایسه شده است. نتایج تجربی، برتری مدل GA-LSTM را نسبت به سایر مدل‌های معیار نشان می‌دهد.

اطلاعات مقاله

تاریخ دریافت: ۱۴۰۳/۰۲/۱۰

تاریخ اصلاحیه: ۱۴۰۳/۰۶/۲۴

تاریخ پذیرش: ۱۴۰۳/۰۷/۱۷

واژگان کلیدی:

پیش‌بینی قیمت سهام،

یادگیری عمیق،

الگوریتم ژنتیک،

انتخاب ویژگی،

تجزیه‌ی سری زمانی.

۱. مقدمه

سرمایه‌گذاری در بازار سهام یک فعالیت محبوب و اغلب سودآور است، که هم سهامداران مبتدی و هم متخصصان حوزه‌ی مذکور از آن لذت می‌برند. سرمایه‌گذاران از راهبردهای مختلفی استفاده می‌کنند، که در دوره‌ی نگهداری و خطرپذیری متفاوت هستند. پیش‌بینی عملکرد سهام، یک هدف مهم در دنیای مالی است، زیرا یک پیش‌بینی نسبتاً دقیق این امکان را برای سرمایه‌گذار فراهم می‌کند که همواره منافع مالی بالایی را به همراه داشته باشد و همچنین ریسک‌های بازار را پوشش دهد.

پیش‌بینی بازار سهام و توسعه‌ی مدل‌های معاملاتی سودآور یکی از مسائلی است که همواره پژوهشگران، تحلیلگران اقتصادی، و معامله‌گران بازار سهام به آن توجه کرده‌اند. با این حال، ارائه‌ی مدلی که تحت شرایط مختلف بازار، عملکرد قابل اعتمادی داشته باشد، بسیار چالش‌برانگیز است. در چند دهه‌ی گذشته، به لطف پیشرفت‌های کامپیوتری و فناوری اطلاعات، مدل‌های محاسباتی هوشمند به‌عنوان جایگزین‌های مناسب برای سیستم‌های تصمیم‌گیری سنتی شروع به ظهور کرده‌اند. مدل‌های قبلی عمدتاً مبتنی بر قوانین و تحلیل‌های ایستا بوده و به دلیل فعل و انفعال‌های دستی بیش از حد، آن‌ها از احساسات انسانی مصون نبوده و در نتیجه بازدهی ناسازگار و وضعیتی

داشته‌اند.^[۱] از سوی دیگر، مدل‌های نوظهور هوش مصنوعی، مانند انواع مختلف شبکه‌های عصبی، عملکرد مناسب‌تری در پیش‌بینی قیمت ارائه کرده و در نتیجه به سرعت موردتوجه پژوهشگران حوزه‌ی مذکور قرار گرفته و مدل‌های ترکیبی بسیاری جهت پیش‌بینی قیمت سهام توسعه داده شده است.^[۲]

در همین حال، عملکرد روش‌های یادگیری ماشین سنتی تا حد زیادی به طراحی ویژگی‌ها بستگی دارد، که این امر باعث تداخل خاصی در عملکرد پیش‌بینی می‌شود. روش‌های سنتی در فرآیند پیش‌بینی نیز ممکن است با مشکلاتی مانند بیش‌برازش و کندی همگرایی مواجه شوند.^[۳] اخیراً، یادگیری عمیق، به‌عنوان یک نسخه‌ی پیشرفته از مدل‌های یادگیری ماشین سنتی، به دلیل عملکرد خوب آن در حوزه‌هایی مانند تشخیص تصویر و تشخیص گفتار، توجه زیادی را در این زمینه به خود جلب کرده است. همچنین در زمینه‌ی پیش‌بینی مالی، یک شبکه‌ی عصبی عمیق امکان افزایش دقت پیش‌بینی بازار سهام را فراهم می‌کند. مهم‌ترین مزایای روش‌های یادگیری عمیق، یادگیری خودکار ویژگی‌ها، یادگیری چندلایه‌ی ویژگی‌ها، دقت بالا در نتایج، قدرت تعمیم بالا، و شناسایی داده‌های جدید است.^[۴] بنابراین، هدف مطالعه‌ی حاضر، ارائه‌ی مدل پیش‌بینی عملکرد سهام با استفاده از شبکه‌ی عصبی عمیق

استناد به این مقاله:

حقیقی نائینی، کیوان و رستگار، محمدعلی، ۱۴۰۴. بررسی عملکرد دو مدل پیش‌بینی قیمت سهام مبتنی بر شبکه‌ی عصبی حافظه‌ی کوتاه‌مدت طولانی و با دو رویکرد متفاوت انتخاب ویژگی و تجزیه‌ی سری زمانی. مهندسی صنایع و مدیریت شریف، ۴۱(۱)، صص. ۸۶-۹۷. <https://doi.org/10.24200/j65.2024.63935.2392>



LSTM و الگوریتم ژنتیک جهت بهبود عملکرد سرمایه‌گذاری در بازار سهام است.

۲. ادبیات نظری و پیشینه‌ی پژوهش

قیمت سهام تحت تأثیر عوامل متعددی قرار می‌گیرد، که مستقیماً یا غیرمستقیم در عرضه و تقاضا در بازار تأثیر می‌گذارند. یکی از مهم‌ترین عوامل، عملکرد شرکت است. شرکت‌های با سودآوری و رشد بالا معمولاً افزایش قیمت سهام را تجربه می‌کنند.

وضعیت اقتصادی، مانند: نرخ بهره، تورم، و رشد اقتصادی نیز آثار مهمی در قیمت سهام دارند، زیرا عوامل مذکور در تصمیم‌های سرمایه‌گذاران اثر می‌گذارند. همچنین، رویدادهای سیاسی و جهانی، نظیر: بحران‌ها، جنگ‌ها، یا تغییرات سیاست‌های دولتی می‌توانند نوسان‌های شدیدی در بازار ایجاد کنند.

اخبار و شایعه‌های مرتبط با شرکت‌ها یا صنعت خاص نیز تأثیرگذار هستند؛ خبرهای مثبت یا منفی می‌توانند ارزش سهام را بالا یا پایین ببرند. سیاست‌های مالی و پولی دولت و بانک مرکزی، مانند تغییر در نرخ بهره یا مالیات‌ها، از دیگر عوامل مهمی هستند که در قیمت سهام اثر می‌گذارند. درنهایت، وضعیت صنعت، که یک شرکت در آن فعالیت می‌کند، نظیر: تغییرات تکنولوژیکی، رقابت و مقررات، نقش مهمی در تعیین قیمت سهام آن شرکت دارد.

یادگیری عمیق (DL)^۱، شاخه‌ای از یادگیری ماشین (ML)^۲، و هوش مصنوعی (AI)^۳ است، که امروزه به‌عنوان هسته‌ی اصلی انقلاب صنعتی چهارم در نظر گرفته می‌شود. به‌دلیل قابلیت یادگیری از داده‌ها، فناوری DL، که از شبکه‌ی عصبی مصنوعی (ANN)^۴ نشأت گرفته است، به یک موضوع داغ در زمینه‌ی محاسبات و مدل‌سازی‌های ریاضی تبدیل شده و به‌طورگسترده در زمینه‌های مختلف کاربردی، مانند: مراقبت‌های بهداشتی، تشخیص بصری، تجزیه و تحلیل متن، امنیت سایبری، پیش‌بینی بازارهای مالی، ریسک اعتباری، و استفاده شده است.

روش کار مدل‌های اخیر، براساس مجموعه‌ای از الگوریتم‌هایی است که سعی می‌کنند درک بالایی از داده‌ها را با استفاده از یک نمودار عمیق با لایه‌های پردازشی متعدد، متشکل از تبدیل‌های خطی و غیرخطی متعدد مدل‌سازی کنند. فرآیند یادگیری ماشین با واردکردن داده‌های آموزشی به الگوریتم موردنظر آغاز می‌شود، تا مدل پیش‌بینی نهایی طراحی شود. سپس داده‌های ورودی جدید به الگوریتم یادگیری ماشینی وارد می‌شود و مقدار پیش‌بینی و نتایج با یکدیگر بررسی می‌شوند تا مشخص شود آیا مدل توسعه داده‌شده به درستی کار می‌کند یا خیر. اگر پیش‌بینی و نتایج مطابقت نداشته باشند، الگوریتم چند بار آموزش داده می‌شود تا زمانی که به نتیجه‌ی مطلوب برسد. این امر به الگوریتم یادگیری ماشینی امکان می‌دهد به‌طور مداوم یاد بگیرد و پاسخ بهینه را تولید کند و به‌تدریج دقت آن در طول زمان افزایش یابد.^[۵]

همان‌طور که گفته شد در سال‌های اخیر، روش‌های یادگیری ماشین و شبکه‌های عصبی مصنوعی به‌دلیل عملکرد بهترشان نسبت به مدل‌های اقتصادسنجی

سنتی، به موضوع اصلی پژوهش‌های حوزه‌ی اخیر و ابزاری کمکی برای پیش‌بینی سری‌های زمانی مالی بدل شده‌اند. تجزیه و تحلیل سری‌های زمانی برای پیش‌بینی مقادیر آینده براساس مقادیر مشاهده‌شده‌ی پیشین، یکی از بهترین ابزارها برای تحلیل روند و پیش‌بینی آینده است. برای هر متغیری که در طول زمان تغییر می‌کند، می‌توان یک سری زمانی در نظر گرفت. سری زمانی، مجموعه‌ای از نقاط داده است که به ترتیب در یک دوره‌ی زمانی رخ می‌دهند. بنابراین، هر مدل پیش‌بینی مبتنی بر داده‌های سری زمانی، زمان را به‌عنوان یک متغیر مستقل خواهد داشت و خروجی هر مدل سری زمانی، مقدار یا طبقه‌بندی پیش‌بینی‌شده در یک زمان خاص خواهد بود. مهم‌ترین مثال سری‌های زمانی در دنیای سرمایه‌گذاری، سری زمانی قیمت اوراق بهادار است، که در یک دوره‌ی زمانی مشخص و در فواصل زمانی معین قیمت سهام ثبت و ردیابی می‌شود. روش‌های پیش‌بینی با استفاده از سری‌های زمانی هم در تحلیل بنیادی و هم در تحلیل تکنیکال استفاده می‌شوند؛ که در ادامه، به بررسی اجمالی تعدادی از پژوهش‌ها در زمینه‌ی مذکور پرداخته شده است:

آدیکاری و آگروال^۵ (۲۰۱۴)، مدلی را برای پیش‌بینی نرخ ارز و شاخص‌های قیمت سهام با ادغام یک مدل پیاده‌روی تصادفی و یک شبکه‌ی عصبی مصنوعی ایجاد کرده‌اند. نیو و وانگ^۶ (۲۰۱۴)، نیز یک شبکه‌ی عصبی شعاعی پایه‌ی (RBF)^۷ مؤثر در زمان داده‌ی تصادفی ایجاد و قیمت نفت خام و شاخص‌های قیمت سهام را پیش‌بینی کرده‌اند. داس^۸ و همکاران (۲۰۱۶)،^[۸] یک مدل ترکیبی جدید SVM-TLBO را برای پیش‌بینی قیمت بسته‌شدن روزانه‌ی شاخص آتی کالا COMDEX، که در چند بورس کالای هند معامله می‌شود، پیشنهاد کرده‌اند؛ که از یک ماشین بردار پشتیبانی (SVM) و یک بهینه‌سازی مبتنی بر یادگیری (TLBO) تشکیل و عملکرد بهتر آن تأیید شده است.

اگرچه شبکه‌های عصبی ذکرشده، عملکرد پیش‌بینی نسبتاً خوبی داشته‌اند، اما دقت آن‌ها در زمانی که سری‌های زمانی پویا و غیرخطی بوده‌اند، همچنان رضایت‌بخش نبوده است. با برقراری ارتباط بین واحدهای لایه‌ی پنهان، وابستگی داده‌ها در نقاط مختلف زمان را می‌توان با شبکه‌های عصبی بازگشتی (RNN)^۹ توضیح داد. ساختار داده‌ی مرتبط قبل و بعد، RNN‌ها را به شکل ویژه‌ای برای پیش‌بینی داده‌های سری زمانی مناسب می‌کند.^[۹] و ادبیات موضوع شاهد کاربرد گسترده‌ی آن‌ها در پیش‌بینی سری‌های زمانی مالی بوده است. با این حال، اطلاعات تاریخی نگهداری‌شده در یک RNN با افزایش بازه‌ی زمانی افزایش می‌یابد و منجر به ناپدیدشدن گرایان یا افزایش ناگهانی گرایان می‌شود و نتایج پیش‌بینی RNN را مخدوش می‌کند. با ظهور شبکه‌ی حافظه‌ی کوتاه‌مدت طولانی (LSTM)، ساختار "سلول حافظه" برای جایگزینی واحدهای لایه‌ی پنهان RNN معرفی شده است. هر سلول حافظه از یک مکانیزم "دروازه" ویژه تشکیل شده است، که باعث می‌شود LSTM اطلاعات تاریخی مهم را حفظ و موارد بی‌اهمیت را فیلتر کند. براساس برتری ویژه‌ی مکانیزم شبکه‌ی عصبی LSTM، از شبکه‌ی عصبی اخیر به‌طورگسترده در پیش‌بینی سری‌های زمانی و به ویژه برای پیش‌بینی سری زمانی قیمت سهام استفاده می‌شود. به‌عنوان مثال، دینگ و کین^{۱۰} (۲۰۲۰)،^[۱۰] یک مدل شبکه‌ی عصبی عمیق بازگشتی مرتبط با چندین ورودی و خروجی مبتنی بر LSTM را

⁶ Niu & Wang

⁷ Radial Basis Function

⁸ Das

⁹ Recurrent Neural Network

¹⁰ Ding & Qin

¹ Deep Learning

² Machine Learning

³ Artificial intelligence

⁴ artificial neural network

⁵ Adhikari & Agrawal

تجزیه‌ی VMD، که از طریق آن می‌توان سری پیچیده‌ی اصلی را به یک سری زیرمجموعه‌ی محدود تجزیه کرد، ارائه شده است. زیرمجموعه‌های تجزیه‌شده توسط الگوریتم VMD از طریق شبکه‌ی LSTM پیش‌بینی شده و نتیجه‌ی پیش‌بینی سری زمانی اصلی از تجمیع نتایج پیش‌بینی زیرمجموعه‌ها به‌دست آمده است. نتایج به‌دست‌آمده، برتری مدل پیشنهادی را نسبت به مدل پایه و مدل‌های مبتنی بر الگوریتم تجزیه‌ی حالت تجربی (EMD) نشان می‌دهند.

کیو^{۱۶} و همکاران (۲۰۲۰)،^[۱۶] یک مدل پیش‌بینی جهت پیش‌بینی قیمت باز شدن سهام ارائه کرده‌اند؛ که در آن، ابتدا داده‌های سهام از طریق الگوریتم تبدیل موجک (WT) پردازش و سپس یک شبکه‌ی عصبی مبتنی بر توجه LSTM برای پیش‌بینی قیمت باز شدن سهام توسعه داده شده است.

خان و همکاران (۲۰۲۳)،^[۱۷] یک روش ترکیبی جدید پیشنهاد کرده‌اند، که براساس نسخه‌ای بهبودیافته از روش تجزیه‌ی تجربی (EMD) و شبکه‌ی LSTM بوده است. در روش اخیر، ابتدا داده‌های نویزی سهام با استفاده از نسخه‌ی جدید EMD، که از روش درون‌یابی Akima به جای درون‌یابی مکعبی استفاده می‌کند، به چندین مؤلفه‌ی جدا به نام توابع مد ذاتی (IMF) تقسیم می‌شوند. سپس زیرمؤلفه‌های با همبستگی بالا برای ساخت مدل LSTM استفاده می‌شوند. این مدل ترکیبی با مدل‌های دیگر مثل LSTM ساده، ماشین بردار پشتیبان (SVM)، جنگل تصادفی، و درخت تصمیم مقایسه شده و نتایج نشان داده است که مدل پیشنهادی ترکیبی-Akima-EMD LSTM نسبت به سایر مدل‌ها عملکرد بهتری داشته و به عنوان یک مدل مؤثر برای پیش‌بینی داده‌های مالی پیچیده و غیرخطی توصیه شده است.

چن و همکاران (۲۰۲۴)،^[۱۸] یک رویکرد انتخاب ویژگی بر پایه‌ی الگوریتم LASSO با شبکه‌ی LSTM آبخاری (Ca-LSTM) ترکیب شده پیشنهاد کرده‌اند. ایده‌ی مدل اخیر براساس بازسازی ویژگی‌های داده و تقویت تأثیر ویژگی‌های منفرد از طریق آبخاری‌سازی بوده است. از طریق مقایسه‌ی جامع نتایج به‌دست‌آمده با مدل‌های رایج دیگر مانند: LSTM، XGboost، SVR، و VMD-LSTM در چهار مجموعه‌ی داده با ویژگی‌های مختلف، برتری مدل پیشنهادی اثبات شده است. همچنین عنوان شده است که هر چه دوره‌ی زمانی داده‌های استفاده‌شده برای پیش‌بینی کوتاه‌تر و نوسان‌های سهام کمتر باشد، مدل عملکرد بهتری دارد. همچنین، زمانی که شبکه‌ی LSTM آبخاری در مؤلفه‌هایی که توسط شبکه‌ی VMD تولید شده‌اند، استفاده می‌شود، عملکرد بهتری نسبت به روش‌های مرجع دارد.

گولمز^{۱۷} (۲۰۲۳)،^[۱۹] یک شبکه‌ی LSTM جدید را که توسط الگوریتم بهینه‌سازی خرگوش مصنوعی (ARO) بهینه شده بود، پیشنهاد کرده است. در مدل اخیر با نام LSTM-ARO، از داده‌های قیمت سهام شاخص DJIA برای پیش‌بینی استفاده و در آن، داده‌ها به فرم جدیدی تبدیل شده‌اند، که از ۲۰ روز گذشته برای پیش‌بینی قیمت روز بعد استفاده می‌کنند. در مدل مذکور، یک شبکه‌ی LSTM جدید با پارامترهای مختلف که به متغیرهای الگوریتم ARO متصل شده‌اند، ایجاد شده است. تغییر متغیرها باعث تغییر در معماری LSTM شده و الگوریتم ARO بهترین معماری را برای دستیابی به نتایج

برای پیش‌بینی قیمت‌های چندگانه، از جمله بالاترین و پایین‌ترین قیمت سهام در روز معاملاتی به‌طور همزمان پیشنهاد کرده‌اند. در پژوهش لیو^{۱۱} (۲۰۱۹)،^[۱۱] از الگوریتم LSTM برای پیش‌بینی نوسان‌های شاخص S&P ۵۰۰ و سهم اپل استفاده شده است. در مطالعه‌ی اخیر، از داده‌های تاریخی برای آموزش و اعتبارسنجی مدل استفاده شده است.

با توجه به کاستی‌های مدل‌های تکی به‌کاررفته در پیش‌بینی داده‌های پیچیده، مدل‌های ترکیبی که دو یا چند مدل مجزا را ادغام می‌کنند، برای حل مشکلاتی مانند دقت پایین مدل و پیش‌بینی‌های عقب‌افتاده پیشنهاد شده‌اند. به این ترتیب، روش‌های پیش‌پردازش مختلفی برای کاهش نویز یا بهینه‌سازی داده‌ها و فرآیند مدل معرفی و سپس داده‌های پردازش‌شده برای به‌دست‌آوردن نتیجه‌ی نهایی وارد مدل پیش‌بینی می‌شوند. در ادامه، به بررسی مهم‌ترین پژوهش‌ها با رویکرد ترکیبی و مشابه با مدل پیشنهادی پرداخته شده است.

زنگ^{۱۲} و همکاران (۲۰۲۲)،^[۱۲] یک مدل ترکیبی AGA-LSTM جهت پیش‌بینی قیمت سهام ارائه کرده‌اند؛ که سهم اصلی آن بهینه‌سازی مدل LSTM با استفاده از الگوریتم ژنتیک تطبیقی (AGA) بوده است. مکانیزم بهینه‌سازی براساس رتبه‌بندی انفرادی با تنظیم خودکار ساختار شبکه‌ی مدل و ترکیبی از فراپارامترها، می‌تواند به‌طور قابل‌توجهی تعداد دفعه‌های تکرار تنظیم هایپرپارامتر را کاهش دهد. طبق نتایج تحلیل انجام‌شده بر روی ۶ شاخص مختلف بازارهای سهام، مدل پیشنهادی در هر فرم عملکرد بهتری نسبت به LSTM و سایر مدل‌های یادگیری ماشین داشته است.

التوفیسی^{۱۳} (۲۰۲۱)،^[۱۳] یک مدل جدید پیش‌بینی بازار سهام را معرفی کرده است، که شامل سه فاز اصلی بوده است: استخراج ویژگی، انتخاب ویژگی بهینه، و پیش‌بینی. در ابتدا، ویژگی‌های آماری، مانند: میانگین، انحراف معیار، واریانس، چولگی، و کشیدگی و مهم‌ترین اندیکاتورهای تکنیکال از داده‌های جمع‌آوری‌شده بازار سهام استخراج شده‌اند. برای به‌دست‌آوردن بهترین نتایج پیش‌بینی‌شده، انتخاب مرتبط‌ترین ویژگی‌ها توسط یک مدل ترکیبی جدید با نام (RDAWA) انتخاب شده است. در مرحله‌ی آخر پیش‌بینی، قیمت سهام با استفاده از ویژگی‌های انتخاب‌شده با یک مدل طبقه‌بندی‌کننده، شامل: ماشین بردار پشتیبان (SVM)، جنگل تصادفی ۱ (RF1)، جنگل تصادفی ۲ (RF2)، و شبکه‌ی عصبی بهینه‌شده (NN) صورت گرفته است. برای پیش‌بینی دقیق‌تر، آموزش شبکه‌ی NN توسط RDAWA پیشنهادی از طریق تنظیم دقیق وزن بهینه انجام شده است. در نهایت، عملکرد کار پیشنهادی نسبت به سایر مدل‌های مرسوم با توجه به معیارهای خاص مقایسه شده است.

چونگ و شین^{۱۴} (۲۰۱۸)،^[۱۴] یک مدل پیش‌بینی قیمت مبتنی بر مدل ترکیبی GA-LSTM پیشنهاد کرده‌اند؛ که در آن از الگوریتم ژنتیک جهت بهینه‌سازی پارامترهای الگوریتم LSTM استفاده شده است. به عبارت دیگر، الگوریتم ژنتیک به جستجوی مقدار بهینه یا نزدیک به بهینه برای اندازه‌ی پنجره‌ی زمانی و تعداد واحدهای LSTM در یک شبکه پرداخته است.

در پژوهش نیو^{۱۵} و همکاران (۲۰۲۰)،^[۱۵] یک مدل ترکیبی VMD-LSTM جهت پیش‌بینی قیمت سهام ارائه شده است. مدل پیشنهادی براساس الگوریتم

¹⁵ Niu¹⁶ Qiu¹⁷ Gülmez¹¹ Liu¹² Zeng¹³ Alotaibi¹⁴ Chung & Shin

معیار و مدل‌های متشکل از سایر الگوریتم‌های مشابه مقایسه شده‌اند. به‌طور مثال، مدل‌های مبتنی بر تجزیه با مدل‌های تجزیه‌ی دیگر مقایسه شده است؛ در واقع، عملکرد الگوریتم‌های تجزیه‌ی متفاوت بررسی و مقایسه شده‌اند و یا در مدل‌های مبتنی بر انتخاب ویژگی، عملکرد الگوریتم‌های فراابتکاری متفاوت در انتخاب ویژگی مقایسه شده‌اند. اما هدف اصلی پژوهش حاضر، مقایسه‌ی کلی‌تر پژوهش‌های ذکر شده و در واقع، مقایسه‌ی دو رویکرد متفاوت در پیش‌بینی قیمت بوده است، که پیشتر در هیچ پژوهشی انجام نشده است و می‌تواند فارغ از الگوریتم‌های مختلف تجزیه و یا انتخاب ویژگی، رویکرد برتر در پیش‌بینی قیمت سهام را معرفی کند.

۳. روش‌شناسی پژوهش

در ادامه، رویکردهای الگوریتم ژنتیک و شبکه‌ی LSTM و معیارهای خطا، معرفی و چارچوب اصلی مدل ترکیبی پیشنهادی ارائه شده است. سپس در بخش بعد، نتایج پیش‌بینی تجربی قیمت سهام، که با استفاده از مدل پیشنهادی به‌دست آمده است، ارائه و با نتایج مدل پایه مقایسه شده است.

۱.۳. شبکه‌ی عصبی کوتاه‌مدت طولانی (LSTM)

مدت کوتاهی پس از آموزش اولین شبکه‌های عصبی بازگشتی با استفاده از الگوریتم پس‌انتشار، مشکلات یادگیری وابستگی‌های بلندمدت به‌دلیل ناپدید شدن و یا رشد انفجاری گرادیان توجه شده است. یکی از اولین و موفق‌ترین روش‌ها برای حل مشکل گرادیان‌های ناپدیدشونده در قالب مدل حافظه‌ی کوتاه‌مدت طولانی (LSTM) توسط هوکرایتر و اشمید هابر^{۱۸} (۱۹۹۷) ارائه شده است. LSTM ها بسیار شبیه به شبکه‌های عصبی بازگشتی استاندارد هستند؛ اما در مدل اخیر، هر گره‌ی معمولی بازگشتی با یک سلول حافظه جایگزین می‌شود و هر سلول حافظه، حاوی یک حالت داخلی است.^[۲۱]

همان‌طور که اشاره شد، در شبکه‌های LSTM یک سلول وجود دارد، که وظیفه‌ی آن حفظ و انتقال اطلاعات مهمی است که باید در حافظه‌ی شبکه باقی بمانند و در گذر از هر مرحله ممکن است تغییر کنند و اطلاعاتی به آن‌ها اضافه یا از آن‌ها کاسته شود؛ که این عمل توسط عملگرهای ضرب و جمع انجام می‌شود و در نتیجه‌ی آن، خطای محاسباتی مدل در عمق بسیار زیاد شبکه نیز منتشر می‌شود و مشکل کاهش شدید گرادیان در مدل مذکور اتفاق نمی‌افتد.^[۲۲]

اما تفاوت اصلی بین شبکه‌های LSTM و RNN این است که در شبکه‌های LSTM، گیت‌هایی، شامل: فراموشی، ورودی، و خروجی وجود دارد. گیت‌ها همچون دروازه‌هایی هستند که اجازه‌ی عبور اطلاعات را در اختیار دارند و با توابعی همچون سیگموئید و عملیات ریاضی مختلف ساخته می‌شوند. تابع سیگموئید بر روی ترکیب خروجی لایه‌ی مخفی قبل و ورودی مرحله‌ی جدید اعمال می‌شود و در نهایت، مقداری بین ۰ و ۱ را بر می‌گرداند، که میزان حفظ یا حذف اطلاعات را تعیین می‌کند (رابطه‌ی ۱):

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \quad (1)$$

در گام بعدی، علاوه بر اعمال تابع سیگموئید، تابع $\tanh$ هم بر ترکیبی از خروجی لایه‌ی پنهان مرحله‌ی قبل و ورودی مرحله‌ی کنونی اعمال می‌شود تا دانش و اطلاعات جدیدی را بسازد؛ اما این نکته که چند درصد از اطلاعات مذکور برای مرحله‌ی بعد لازم است حفظ شود، توسط تابع سیگموئید تعیین

بهبهینه پیدا کرده است. برای ارزیابی کیفیت مدل پیشنهادی، مدل اخیر با شبکه‌های LSTM1D, LSTM2D, LSTM3D, ANN و LSTM-GA مقایسه شده و نتایج نشان داده‌اند که مدل LSTM-ARO، بهترین عملکرد را در میان آن‌ها داشته و پیش‌بینی‌های بسیار دقیقی ارائه داده است.

نصیری و عبادزاده (۲۰۲۳)،^[۲۰] دو روش جدید برای پیش‌بینی چندمرحله‌ای قیمت سهام براساس دو روش مختلف تجزیه، یعنی تبدیل خطی کسینوسی گسسته (DCT) و تبدیل غیرخطی تجزیه‌ی حالت‌های متغیر (VMD) ارائه کرده‌اند. روش DCT-MFRFNN براساس DCT و شبکه‌ی عصبی فازی بازگشتی چندمنظوره (MFRFNN) طراحی و از DCT برای کاهش نوسان‌های سری‌های زمانی و ساده‌سازی ساختار آن و از MFRFNN برای پیش‌بینی قیمت سهام استفاده شده است. روش VMD-MFRFNN، که مبتنی بر VMD و MFRFNN است، مزایای هر دو را با هم ترکیب می‌کند. VMD-MFRFNN از دو مرحله تشکیل شده است: در مرحله‌ی تجزیه، سیگنال ورودی به چند تابع حالت ذاتی (IMF) توسط VMD تجزیه می‌شود. در مرحله‌ی پیش‌بینی، هر IMF به یک MFRFNN جداگانه داده می‌شود و سیگنال‌های پیش‌بینی شده برای بازسازی خروجی نهایی جمع می‌شوند.

نتایج تجربی نشان داده است که VMD-MFRFNN عملکرد بهتری نسبت به سایر روش‌های پیشرفته داشته و همچنین، DCT-MFRFNN در تمامی آزمایش‌ها عملکرد بهتری نسبت به MFRFNN و DCT-LSTM داشته است، که تأثیر مثبت DCT در عملکرد MFRFNN را نشان می‌دهد.

با بررسی پژوهش‌های انجام‌شده مشاهده می‌شود که در مطالعات خارجی و داخلی با گذر زمان، روش‌های یادگیری عمیق جایگزین روش‌های سنتی یادگیری ماشین در مدل‌های پیشنهادی شده‌اند. همان‌طور که پیشتر گفته شده است، مطالعات بررسی‌شده در پژوهش حاضر در دو دسته‌ی کلی مدل‌های پیش‌بینی مبتنی بر رویکرد تجزیه و همچنین مدل‌های پیش‌بینی مبتنی بر رویکرد انتخاب ویژگی قرار می‌گیرند. در مدل‌های مبتنی بر رویکرد تجزیه، از یک الگوریتم تجزیه در ترکیب با مدل پیش‌بینی اصلی استفاده شده است. پژوهش‌های انجام‌شده در این حوزه به مرور زمان و با ظهور الگوریتم‌های تجزیه‌ی جدید به روز شده و نتایج مدل‌های اخیر نیز بهبود یافته‌اند. به‌طور کلی بیشتر پژوهش‌های این حوزه از الگوریتم‌های مشهور EMD و WT در مدل‌های خود استفاده کرده‌اند؛ در حالی که در سال‌های اخیر، الگوریتم‌های تجزیه‌ی جدیدتری مانند EEMD, CEEMD, و VMD معرفی و کمتر در مدل‌های پیش‌بینی استفاده شده‌اند. بنابراین، یکی از اهداف پژوهش حاضر، مقایسه‌ی عملکردی الگوریتم‌های تجزیه‌ی جدید با الگوریتم‌های تجزیه‌ی EMD و WT است. اما با بررسی پیشینه‌ی پژوهش‌های دسته‌ی دوم، یعنی پژوهش‌های مرتبط یا رویکرد انتخاب ویژگی، که از ترکیب یک الگوریتم فراابتکاری در ترکیب با یک مدل پیش‌بینی توسعه می‌یابند، می‌توان گفت که پژوهش‌های مذکور نسبت به حوزه‌ی مبتنی بر تجزیه‌ی سری زمانی جدیدتر است. به‌طوری که با مرور ادبیات می‌توان دریافت که از الگوریتم‌های فراابتکاری بیشتر و اکثراً در بهینه‌سازی هابیر پارامترهای مدل استفاده شده است. همچنین، از الگوریتم‌های مذکور جهت انتخاب ویژگی در پژوهش‌های جدیدتر معرفی و استفاده شده است. بررسی پژوهش‌های دو حوزه‌ی اخیر نشان می‌دهد که در آن‌ها جهت بررسی کارایی مدل‌ها، نتایج پیش‌بینی با نتایج مدل‌های

توسط دراگومیرتسکی و زوسو^{۱۹} (۲۰۱۳) ارائه شده است.^[۲۴] روش VMD، در واقع یک سیگنال نامانا را به توابع حالت ذاتی چندگانه تجزیه می‌کند. بنابراین، روش VMD برای پرداختن به داده‌های پیچیده‌ای، مانند سری‌های زمانی مالی مناسب‌تر بوده و به‌طورگسترده‌ای برای پیاده‌سازی تجزیه‌ی سیگنال‌های چندبسامدی به‌صورت تطبیقی استفاده شده است، که در آن باند بسامد را می‌توان به‌طورهمزمان تخمین زد، تا به‌طورمؤثر خطاها را بین باندهای بسامدی مختلف متعادل سازد. روش VMD با الهام از روش EMD فرض می‌کند که سیگنال اصلی f از تعدادی توابع حالت ذاتی (IMF)، که در رابطه‌ی ۷ با u_k نشان داده شده است، تشکیل شده است:

$$f(t) = \sum_k u_k(t) \sum_k A_k(t) \cos[\phi_k(t)] \quad (7)$$

که در آن، A_k دامنه‌ی u_k و $\phi_k(t)$ فاز u_k هستند و بسامد مرکزی هر IMF نیز با بسامد لحظه‌ای آن طبق رابطه‌ی ۲ در نظر گرفته می‌شود.

با شروع یک مسئله‌ی حذف نویز رایج، سیگنال مشاهده‌شده‌ی f به‌صورت رابطه‌ی ۸ فرموله می‌شود، که در آن f ، سیگنال اصلی و η یک افزونه‌ی نوفه‌ی سفید است:

$$f_0 = f + \eta \quad (8)$$

برای بازیابی سیگنال اصلی f از تنظیم Tikhonov برای کمینه‌سازی f مطابق رابطه‌ی ۹ استفاده می‌شود:

$$\min_f \|f - f_0\|_2^2 + \alpha \|\partial_t f\|_2^2 \quad (9)$$

سپس معادلات اوایلر- لاگرانژ به‌راحتی در حوزه‌ی بسامد به‌دست می‌آیند و مطابق رابطه‌ی ۱۰ حل می‌شوند:

$$\hat{f}(\omega) = \frac{\hat{f}_0}{1 + \alpha \omega^2} \quad (10)$$

که در آن، $\hat{f}(\omega)$ تبدیل فوریه‌ی f و α نشان‌دهنده‌ی واریانس نوفه‌ی سفید هستند. برای ارزیابی پهنای باند یک حالت، مسئله‌ی تغییرات محدود به‌صورت رابطه‌ی ۱۱ فرموله می‌شود؛ که در آن، $\{u_k\} = \{u_1, \dots, u_k\}$ مجموعه‌ی IMF ها، $\{\omega_k\} = \{\omega_1, \dots, \omega_k\}$ مجموعه‌ی بسامدهای

مرکزی، و عبارت $(\sigma(t) + \frac{j}{\pi t}) * u_k(t)$ نشانگر تبدیل هیلبرت u_k هستند.

$$\text{s.t. } \sum_k u_k = f \quad (11)$$

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\sigma(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}$$

برای کمینه‌سازی معادله‌ی مذکور، از تابع لاگرانژ تقویت‌شده با معرفی یک جمله‌ی جریمه‌ی درجه‌ی دوم و ضریب لاگرانژی λ مطابق رابطه‌ی ۱۲ استفاده می‌شود:

$$L(\{u_k\}, \{\omega_k\}, \lambda) := \alpha \sum_k \left\| \partial_t \left[\left(\sigma(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \frac{1}{2} \|f(t) - \sum_k u_k(t)\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \quad (12)$$

می‌شود و سپس آن‌ها به سلول حالت اضافه می‌شوند. روابط ۲ و ۳، نشان‌دهنده‌ی فرآیند مذکور هستند:

$$i_t = \sigma(w_i * [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\dot{C}_t = \tanh(W_c * [h_{t-1}, x_t] + b_c) \quad (3)$$

سلول حالت طبق رابطه‌ی ۴ به روزرسانی می‌شود:

$$C_t = f_t * C_{t-1} + i_t * \dot{C}_t \quad (4)$$

در مرحله‌ی آخر، برای تولید خروجی، یک تابع $\tanh$ بر روی اطلاعات سلول وضعیت اعمال شده است، تا مقادیر بین ۰ تا ۱ ایجاد شود؛ چرا که جمع‌کردن اطلاعات در هر مرحله ممکن است بازه‌ی مقادیر سلول را تغییر دهد. همچنین یک تابع سیگموئید بر روی ترکیبی از خروجی لایه‌ی پنهان قبلی و ورودی این مرحله اعمال شود و درنهایت، خروجی سیگموئید در خروجی $\tanh$ ضرب می‌شود و نتیجه‌ی نهایی مطابق روابط ۵ و ۶ به‌دست خواهد آمد:

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

۲.۳. الگوریتم ژنتیک (GA)

الگوریتم ژنتیک (GA)، در واقع از نظریه‌ی تکامل داروین الهام گرفته شده است و کاربردهای بسیار زیادی در یافتن جواب مسائلی از قبیل: بهینه‌سازی، مکان‌یابی، و ... دارد. یکی از نقاط قوت الگوریتم GA، سرعت بالای آن در حرکت به سمت جواب مسئله است، که باعث می‌شود در مسائلی که فضای جواب بسیار وسیعی دارند و بررسی تمام جواب‌ها عملاً غیرممکن است، بسیار استفاده شود. از دیگر ویژگی‌های مثبت الگوریتم GA می‌توان به جستجوی هم‌زمان در چند جهت و همچنین عدم نیاز به ویژگی‌های تحدب و پیوستگی تابع هدف مسئله اشاره کرد.

شیوه‌ی کار الگوریتم GA به بیان ساده به این صورت است که الگوریتم با مجموعه‌ای از نقاط نامزد کار خود را آغاز می‌کند و با استفاده از عملگرهای ژنتیکی در جهت یافتن پاسخ بهتر پیش می‌رود.

هر کروموزوم در الگوریتم، یک نقطه در فضای جستجوی جواب یا به‌عبارت دیگر، یک راه حل مسئله را نشان می‌دهد. کروموزوم‌ها از مجموعه‌ای از ژن‌ها تشکیل می‌شوند و مجموعه‌ای از کروموزوم‌ها، که با کدهای باینری نمایش داده می‌شوند، یک جمعیت را تشکیل می‌دهند. شیوه‌ی کار الگوریتم ژنتیکی معمولاً به این صورت است که با تولید جمعیتی از کروموزوم‌ها آغاز می‌شود. در گام بعد، کروموزوم‌های ساخته‌شده بررسی و ارزیابی می‌شوند و کروموزوم‌هایی که با توجه به تابع هدف مسئله، پاسخ بهینه‌تری تولید می‌کنند، شانس بیشتری برای تولید مثل پیدا می‌کنند.^[۲۳]

الگوریتم ژنتیک تا جایی ادامه می‌یابد که جمعیت همگرا شده باشد؛ به این معنی که الگوریتم فرزندان که تفاوت قابل‌توجهی با نسل قبلی داشته باشند، تولید نکند. در این مرحله عنوان می‌شود که الگوریتم مجموعه‌ای از پاسخ‌ها را برای مسئله ارائه کرده است.

۳.۳. روش تجزیه‌ی حالت متغیر

روش تجزیه‌ی حالت متغیر (VMD)، یکی از جدیدترین روش‌های تجزیه‌ی سیگنال و مشابه روش تجزیه‌ی حالت تجربی (EMD) است، که برای اولین بار

- **گام دوم:** از روش رولت برای عملیات انتخاب استفاده می‌شود و تناسب هر کروموزوم در جمعیت محاسبه می‌شود. احتمال انتخاب هر فرد به نسبت تناسب کروموزوم‌هاست و مجموع احتمال انتخاب کروموزوم ۱ است. در طول الگوریتم، جمعیت یک بار در هر دور با توجه به احتمال به روزرسانی می‌شود.

- **گام سوم:** برای انجام عملیات متقاطع از روش تقاطع چندنقطه‌ای^{۲۰} استفاده می‌شود و کروموزوم‌ها بین دو فرد با احتمال متقاطع ۰/۸ رد و بدل خواهند شد. در فرآیند الگوریتم، یک عملیات متقاطع بر روی هر کروموزوم در هر چرخه انجام می‌شود و روش تولید یک احتمال تصادفی است، که اگر مقدار آن کمتر از احتمال متقاطع باشد، تعویض انجام می‌شود و در غیر این صورت، تعویض صورت نمی‌گیرد.

- **گام چهارم:** این الگوریتم از روش جهش بیت پایه^{۲۱} برای انجام عملیات جهش استفاده شده است. در فردهای معاصر، با احتمال کمی یک ژن تغییر می‌کند. احتمال تغییر برابر با ۰/۲ تنظیم شده است. الگوریتم در هر دور یک احتمال تصادفی تولید می‌کند. اگر احتمال تصادفی کمتر از احتمال تلاقی باشد، جهش انجام می‌شود؛ در غیر این صورت، جهش صورت نمی‌گیرد.

در مرحله‌های ۲ تا ۴، الگوریتم ۱۰۰ بار تکرار می‌شود و در پایان الگوریتم، جمعیت بهینه‌ی نزدیک به راه‌حل بهینه تولید می‌شود. پارامترهای الگوریتم ژنتیک در مدل پیشنهادی در جدول ۱ ارائه شده است:

- مرحله‌ی دوم

ویژگی‌های انتخاب‌شده توسط الگوریتم ژنتیک در مرحله‌ی قبل به‌عنوان ویژگی‌های ورودی مدل LSTM در نظر گرفته می‌شوند و پیش‌بینی قیمت بسته‌شدن سهم در روز آتی با استفاده از ویژگی‌های ورودی مذکور انجام می‌شود. مقادیر هاپرپارامترهای شبکه‌ی LSTM در مدل پیشنهادی، در جدول ۲ ارائه شده است.

جدول ۱. مقادیر پارامترهای الگوریتم ژنتیک.

Item	Parameter	Value
1	Chromosome Size	25
2	Population Size	100
3	Crossover Rate	0.8
4	Mutation Rate	0.02
5	Number of Generations	100
6	Fitness Function	MSE

جدول ۲. مقادیر پارامترهای شبکه‌ی عصبی LSTM.

Item	Parameter	Value
1	Model Step	4
2	Number of Hidden Layer Neurons	300
3	Dropout Rate	0.2
4	Activation Function	Relu
5	Optimizer Function	Adam
6	Batch Size	256
7	Number of Epochs	100

نتیجه‌ی نهایی با به روزرسانی بسامد مرکزی و پهنای باند هر IMF در یک دنباله‌ی بهینه‌سازی به نام ADMM به‌دست می‌آید. روش به روزرسانی بسامد را می‌توان به‌صورت رابطه‌ی ۱۳ بیان کرد:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2} \quad (13)$$

که در آن، $\hat{u}_k^{n+1}(\omega)$ ، $\hat{f}(\omega)$ و $\hat{\lambda}(\omega)$ عبارت‌اند از تبدیل‌های فوریه‌ی عبارت‌های u_k^{n+1} ، $f(t)$ و $\lambda(\omega)$. معادله‌ی اخیر را می‌توان به‌عنوان یک فیلتر وینر شناسایی کرد، که در آن $\hat{u}_k^{n+1}(\omega)$ خروجی فیلتر $\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2$ با سیگنال قبل از $1 + 2\alpha(\omega - \omega_k)^2$ است. در روش‌های تکراری، بسامد مرکزی هر u_k با روش گرانش طیف توان مجدداً تخمین زده می‌شود. معادله‌های به روزشده‌ی ω_k^{n+1} مطابق رابطه‌ی ۱۴ به‌دست می‌آیند:

$$\omega_k^{n+1} = \frac{\int_{-\infty}^{\infty} \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |\hat{u}_k(\omega)|^2 d\omega} \quad (14)$$

فرآیند VMD زمانی متوقف می‌شود که خطای نسبی e کمتر از حد همگرایی ε باشد (مطابق رابطه‌ی ۱۵):

$$e = \frac{\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon \quad (15)$$

۳.۳. تشریح مدل‌های پیشنهادی

در پژوهش حاضر، یک مدل پیش‌بینی قیمت سهام دو مرحله‌ای با ترکیب شبکه‌ی عصبی حافظه‌ی کوتاه‌مدت طولانی و الگوریتم ژنتیک و با رویکرد انتخاب ویژگی پیشنهاد می‌شود، که در ادامه به شرح مراحل مدل مذکور پرداخته شده است:

- **مرحله‌ی اول:** در ابتدا، با استفاده از الگوریتم ژنتیک به انتخاب ویژگی‌های با اهمیت پرداخته شده است. جهت دسترسی به مهم‌ترین ویژگی‌ها برای پیش‌بینی هر سهم، این گام‌ها به ترتیب انجام می‌شود:

- **گام اول:** مقداردهی اولیه‌ی تصادفی جمعیت. جمعیت الگوریتم ژنتیک با pop مطابق رابطه‌ی ۱۶ بیان می‌شود:

$$pop = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mk} \end{pmatrix} \quad (16)$$

که در آن، هر ردیف یک کروموزوم (یا مجموعه‌ی ویژگی) را نشان می‌دهد. طول کروموزوم (k) ، تعداد کل ویژگی‌ها را نشان می‌دهد، که در پژوهش حاضر مساوی ۲۵ در نظر گرفته شده است. مقدار m ، اندازه‌ی جمعیت را نشان می‌دهد، که در پژوهش حاضر مساوی ۱۰۰ در نظر گرفته شده است. همچنین a_{ij} ها یکی از دو مقدار ۰ و ۱ را نشان می‌دهند، که مقدار ۱ نشانگر انتخاب و مقدار ۰ نشانگر عدم انتخاب هر ویژگی است.

²⁰ Multi point Crossover

²¹ Bit Flip Mutation

- مرحله‌ی سوم

که در آن‌ها، y_t نشانگر مقدار واقعی، $\hat{y}_t$ نشان‌دهنده‌ی مقدار پیش‌بینی‌شده‌ی مدل و n برابر با تعداد نمونه است. به‌طورکلی، هر چه مقادیر MAE و MSE کوچک‌تر باشند، فاصله‌ی بین مقدار پیش‌بینی‌شده و مقدار واقعی کمتر خواهد بود و به‌عبارت دیگر، دقت پیش‌بینی مدل بالاتر است.

۴.۳. ویژگی‌های استفاده‌شده در مدل پیشنهادی

در مدل پیشنهادی برای ۵ سهم موردنظر، مقادیر ۲۵ ویژگی که شامل اطلاعات قیمتی و تعدادی از مهم‌ترین اندیکاتورهای تحلیل تکنیکال است، محاسبه شده است. همچنین برای انتخاب مهم‌ترین ویژگی‌ها در پیش‌بینی، قیمت هر سهم به‌عنوان ورودی به الگوریتم ژنتیک وارد شده و از فهرست ویژگی‌های منتخب الگوریتم ژنتیک در مرحله‌ی دوم برای پیش‌بینی قیمت سهام استفاده شده است. در ادامه و در جدول ۳، فهرست ۲۵ ویژگی مذکور ارائه شده است.

برای ارزیابی کیفیت پیش‌بینی یک مدل توسعه‌یافته معمولاً چند معیار عملکرد استفاده می‌شود. در پژوهش حاضر، از معیارهای میانگین خطای مطلق (MAE)، ریشه‌ی میانگین مربعات خطا ($RMSE$)، و ضریب تعیین (R^2) برای ارزیابی دقت پیش‌بینی استفاده می‌شود. معیارهای گفته‌شده به ترتیب به‌صورت روابط ۱۷ و ۱۸ محاسبه می‌شوند:

$$MAE = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n} \quad (17)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}} \quad (18)$$

جدول ۳. ویژگی‌های استفاده‌شده در مدل GA-LSTM.

Item	parameter	value
1	High	Highest price of the stock during the trading day
2	Low	Lowest price of the stock during the trading day
3	Open	Opening price of the stock on the trading day
4	TP	Typical price of the stock
5	Adj Close	Adjusted closing price of the stock on the trading day
6	Volume	Trading volume of the stock on the trading day
7	SMA(5)	5-day Simple Moving Average
8	SMA(20)	20-day Simple Moving Average
9	SMA(50)	50-day Simple Moving Average
10	EMA(10)	10-day Exponential Moving Average
11	EMA(20)	20-day Exponential Moving Average
12	EMA(100)	100-day Exponential Moving Average
13	MACD	MACD line (Moving Average Convergence Divergence)
14	Signal_line	Signal line of the MACD indicator
15	ATR	ATR (Average True Range) indicator
16	ROC	ROC (Rate of Change) indicator
17	RSI	RSI (Relative Strength Index) indicator
18	SO%K	Stochastic Oscillator %K indicator
19	William%R	Williams %R indicator
20	Lower_Band	Lower band of the Bollinger Bands indicator
21	Upper_Band	Upper band of the Bollinger Bands indicator
22	SD20	20-day Standard Deviation of price
23	MOM	Momentum indicator
24	SnP_Close	Closing value of the S&P 500 index on the trading day
25	DJIA_Close	Closing value of the Dow Jones Industrial Average (DJIA) index on the trading day

خطاهای پیش‌بینی IMF های با بسامد پایین، نسبت به خطای پیش‌بینی IMF های با بسامد بالا، تأثیر مهم‌تری در نتیجه‌ی پیش‌بینی نهایی خواهد داشت.

مرحله‌ی ۴: در مرحله‌ی آخر، عملکرد این مدل نیز با معیارهای میانگین خطای مطلق (MAE)، ریشه‌ی میانگین مربعات خطا (RMSE)، و ضریب تعیین (R^2) ارزیابی می‌شود.

۴. یافته‌های پژوهش

همان‌طور که پیشتر گفته شد، در پژوهش حاضر از داده‌های قیمتی بازه‌های روزانه، ۵ سهم بزرگ بازار بورس نیویورک، شامل: گلدمن ساکس (GS)، یونایتد هلث (UNH)، جی پی مورگان (JPM)، آی بی ام (IBM)، و جانسون اند جانسون (JNJ) استفاده شده است. بازه‌ی تاریخی بررسی‌شده‌ی هر سهم از تاریخ ۲۰۰۱/۰۱/۰۱ تا ۲۰۲۰/۰۱/۰۱ و به استثناء روزهای تعطیل بوده است. از ۸۰٪ داده‌های ابتدایی نمونه به عنوان داده‌ی آموزشی و از ۲۰٪ باقیمانده به‌عنوان داده‌ی آزمایشی استفاده شده است.

در جدول ۵، ویژگی‌های منتخب هر سهم و در جدول ۶، نتایج مدل GA-LSTM که با استفاده از رویکرد انتخاب ویژگی و با ۲۵ ویژگی ذکرشده، توسعه داده شده است؛ با نتایج مدل پیشنهادی دوم یعنی مدل VMD-LSTM، که با استفاده از تجزیه‌ی سری زمانی ویژگی قیمت بسته‌شدن سهام توسعه داده‌شده مقایسه شده است. علاوه بر این، نتایج دو مدل پیشنهادی، با نتایج سایر مدل‌های معیار نیز مقایسه شده است (جدول ۷).

مطابق جدول ۶، علی‌رغم عملکرد مناسب و خطای کم هر دو مدل پیشنهادی، مدل GA-LSTM که مبتنی بر انتخاب ویژگی توسعه داده شده است، در ۴ سهم از ۵ سهم بررسی‌شده، با توجه به تمام معیارهای خطای گزارش‌شده، نتایج دقیق‌تری ارائه کرده و مدل VMD-LSTM نیز در پیش‌بینی قیمت سهم جانسون اند جانسون (JNJ) نسبت به مدل دیگر عملکرد بهتری داشته است.

همچنین نتایج جدول ۷ نشان می‌دهد که مدل مبتنی بر الگوریتم ژنتیک نسبت به دو مدل EMD-LSTM و WT-LSTM نیز عملکرد بهتری داشته و توانایی پیش‌بینی را در مقایسه با مدل منفرد بسیار افزایش داده است. از نتایج می‌توان استنباط کرد که رویکرد انتخاب ویژگی می‌تواند علاوه بر کاهش عملیات مدل با استفاده از ویژگی‌های کمتر، دقت مدل را نیز به شکل چشمگیری افزایش دهد. به‌عنوان مثال، در شکل ۱، نمودار سمت چپ نتیجه‌ی پیش‌بینی مدل GA-LSTM و نمودار سمت راست نتیجه‌ی مدل VMD-LSTM در پیش‌بینی قیمت سهم یونایتد هلث مشاهده می‌شود.

در ادامه، به‌منظور بررسی معناداری تفاوت نتایج دو مدل پیشنهادی، از معیار MAE در آزمون دیبولد-ماریانو^{۲۲} استفاده شده است. آزمون اخیر نشان می‌دهد که قدرت پیش‌بینی یکسان تحت فرضیه‌ی $DM \sim N(0, 1)$ است. بنابراین، فرضیه‌ی قدرت پیش‌بینی یکسان در سطح ۹۵٪، زمانی که $|DM| > 1/96$ باشد، رد می‌شود.

همچنین، آزمون دیبولد-ماریانو برای خطای پیش‌بینی مدل‌های VMD-LSTM و GA-LSTM استفاده شده و $DM = 2/217$ به‌دست آمده است. بنابراین، مطابق نتایج به‌دست‌آمده، در سطح اطمینان ۹۵٪، قدرت

با توجه به عملکرد خوب الگوریتم تجزیه‌ی VMD و شبکه‌ی عصبی LSTM، در پژوهش حاضر، مدل ترکیبی نوآورانه‌ی دوم با عنوان VMD-LSTM توسعه داده شده است؛ که در ادامه، به شرح مراحل مدل مذکور پرداخته شده است:

مرحله‌ی ۱: روش VMD برای تجزیه‌ی سری زمانی اصلی قیمت سهام به K زیرمجموعه‌ی مستقل متقابل استفاده می‌شود؛ که با $IMF_1, IMF_2, \dots, IMF_k$ نشان داده می‌شوند، که نشانگر نوسان‌های جزئی مختلف قیمت از بسامد بالا تا بسامد پایین است، یعنی سری زمانی قیمت بر حسب IMF ها به صورت رابطه‌ی ۱۹ بازسازی می‌شود:

$$x(t) = \sum_{k=1}^k IMF_k(t) \quad (19)$$

جهت کاهش تأثیر نوسان‌ها و تسهیل بهینه‌سازی فرآیند حل، هر IMF به‌دست‌آمده از الگوریتم VMD در محدوده‌ی $[0, 1]$ نرمالایز می‌شود. سپس هر سری IMF نرمالایز شده به‌عنوان ویژگی جدید ورودی مدل LSTM در نظر گرفته می‌شود.

مرحله‌ی ۲: هر جزء IMF به‌دست‌آمده توسط VMD به مجموعه‌ی داده‌های آموزشی و آزمایشی با یک نسبت ثابت تقسیم می‌شود و مجموعه‌های ورودی و خروجی براساس اندازه‌ی گام تقسیم می‌شوند. شبکه‌ی LSTM برای آموزش و ساخت مدل پیش‌بینی براساس مجموعه‌ی داده آموزشی استفاده می‌شود، که در آن پارامترهایی از جمله: نرخ یادگیری، اندازه‌ی دسته، تعداد تکرارها، و تعداد لایه‌های پنهان وجود دارد، که برای اطمینان از دقت پیش‌بینی مدل حیاتی هستند و باید از قبل تنظیم شوند. فهرست پارامترهای مذکور و مقادیر آن‌ها در جدول ۴ ارائه شده است. اندازه‌ی گام در هر دو مدل با استفاده از نتایج آزمایش‌های تکراری با گام‌های متنوع به‌دست آمده و روی مقدار ۴ تنظیم شده است؛ به این معنا که هم در مجموعه‌ی آموزشی و هم در مجموعه‌ی آزمون، مقدار پیش‌بینی را می‌توان براساس ورودی نقاط داده‌ی چهار مرتبه‌ی قبل به‌دست آورد.

مرحله‌ی ۳: نتیجه‌ی پیش‌بینی نهایی سری زمانی قیمت سهام اصلی با تجمیع خروجی‌های پیش‌بینی‌شده‌ی جداگانه‌ی هر IMF به‌دست می‌آید. در میان، زیرمجموعه‌ها، IMF های با بسامد پایین معمولاً از مقادیر بزرگی تشکیل شده- اند که نشانگر روند اصلی سری زمانی هستند. از منظر سطح دقت پیش‌بینی،

جدول ۴. تنظیم‌های شبکه‌ی عصبی LSTM.

Item	parameter	Value
1	Model Step	4
2	Number of Network Layers	3
3	Number of Hidden Layer Neurons	300
4	Dropout Rate	0.2
5	Activation Function	Relu
6	Optimizer Function	Adam
7	Batch Size	512
8	Number of Epochs	80

جدول ۵. فهرست ویژگی‌های منتخب هر سهم.

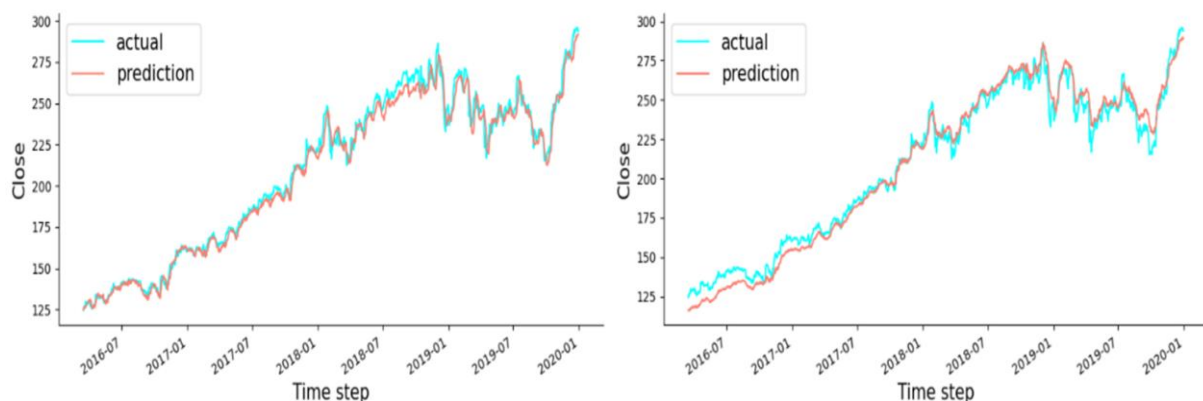
Item	Stock	Selected features
1	(UNH)	['High', 'Low', 'Volume', 'MACD', 'Adj Close', 'WT', 'ROC', 'EMA(10)', 'mom', 'RSI', 'atr', 'SnP_Close']
2	(GS)	['Adj Close', 'MA(50)', 'High', 'Low', 'Signal Line', 'MA(5)', 'ROC', 'sd(20)', 'atr', 'MACD', 'EMA(20)', 'RSI']
3	(IBM)	['High', 'Low', 'Signal Line', 'WT', 'MA(5)', 'SMA', 'sd(20)', 'atr', 'MACD', 'EMA', 'RSI']
4	(JPM)	['upper_band', 'lower_band', 'Low', 'Signal Line', 'WT', 'MA(5)', 'EMA(10)', 'mom', 'RSI', 'atr']
5	(JNJ)	['Open', 'High', 'Low', 'Signal Line', 'Adj Close', 'STOK', 'ROC', 'EMA(10)', 'RSI',]

جدول ۶. مقایسه‌ی نتایج دو مدل پیشنهادی.

Item	Stock	VMD-LSTM			GA-LSTM		
		MAE	RMSE	(%) R ²	MAE	RMSE	(%) R ²
1	(UNH)	5.292	6.445	98.15	3.77	5.093	98.84
2	(GS)	3.54	4.409	97.98	2.967	3.996	98.34
3	(IBM)	2.073	2.738	95.80	1.746	2.456	96.62
4	(JPM)	1.705	2.097	98.81	1.44	1.887	99.03
5	(JNJ)	1.284	1.733	96.42	1.295	1.886	96.09

جدول ۷. مقایسه‌ی نتایج مدل‌های پیشنهادی با سایر مدل‌ها.

Item	Model	(UNH)			(GS)			(JPM)		
		MAE	RMSE	(%) R ²	MAE	RMSE	(%) R ²	MAE	RMSE	(%) R ²
1	VMD-LSTM	5.292	6.445	98.15	3.54	4.409	97.98	1.705	2.097	98.81
2	GA-LSTM	3.77	5.093	98.84	2.967	3.996	98.34	1.44	1.887	99.03
3	EMD-LSTM	5.509	7.567	97.45	5.433	6.688	95.37	3.6	4.392	94.78
4	WT-LSTM	10.785	12.99	92.53	4.174	5.389	96.94	2.158	2.827	97.83
5	LSTM	7.508	9.559	95.91	3.615	4.738	97.67	2.718	3.396	96.87



شکل ۱. نتایج پیش‌بینی سهام unnh با دو مدل پیشنهادی.

پیش‌بینی مدل GA-LSTM بیشتر بوده و آزمون اخیر درستی آن را تأیید کرده است.

استنباط کرد که به‌طور کلی مدل‌های مبتنی بر رویکرد انتخاب ویژگی، که پیش‌بینی را با استفاده از ویژگی‌های متنوع انجام می‌دهند، عملکرد مناسب‌تری نسبت به مدل‌های مبتنی بر تجزیه‌ی سری زمانی ارائه می‌کنند، که فقط از ویژگی قیمتی جهت پیش‌بینی استفاده می‌کنند. همچنین نتایج پژوهش حاضر، حاکی از آن است که مدل پیشنهادی که در آن از الگوریتم ژنتیک جهت انتخاب ویژگی استفاده شده است، نسبت به مدل پایه که از تمام ویژگی‌ها در پیش‌بینی قیمت سهام بهره برده است، علاوه بر کاهش ابعاد و محاسبات مدل در نتیجه‌ی کاهش ویژگی‌های استفاده‌شده در مدل، دقت پیش‌بینی را نیز به شکل چشمگیری افزایش داده است. همچنین می‌توان استنباط کرد که استفاده از الگوریتم‌های فراابتکاری در کنار مدل‌های یادگیری عمیق می‌تواند عملکرد مدل‌های مذکور را در پیش‌بینی قیمت سهام بهبود بخشد.

۵. نتیجه‌گیری و بحث

در پژوهش حاضر، دو مدل با دو رویکرد متفاوت انتخاب ویژگی و تجزیه‌ی سری زمانی ارائه شده و نتیجه گرفته شده است که هر دو مدل پیشنهادی، عملکرد مناسب و قابل اتکایی ارائه می‌کنند؛ اما مدل VMD-LSTM نسبت به سایر مدل‌های مشابه، مانند الگوریتم تجزیه‌ی حالت تجربی و الگوریتم تبدیل موجک عملکرد دقیق‌تری را ارائه کرده و همچنین در مقایسه، دو رویکرد تجزیه‌ی سری زمانی قیمت و انتخاب ویژگی، با توجه به نتایج به‌دست‌آمده می‌توان گفت که عملکرد مدل GA-LSTM با رویکرد انتخاب ویژگی دقیق‌تر بوده و می‌توان

References- منابع

- Berradi, Z., Lazaar, M., Mahboub, O. and Omara, H., 2021, June. A comprehensive review of artificial intelligence Techniques in Financial Market. In *2020 6th IEEE congress on information science and technology (CiSt)*, pp. 367-371. IEEE. <https://doi.org/10.1109/CiSt49399.2021.9357175>
- Le, D.Y.N., Maag, A. and Senthilananthan, S., 2020, November. Analysing stock market trend prediction using machine & deep learning models: A comprehensive review. In *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)*, pp. 1-10. IEEE. <https://doi.org/10.1109/CITISIA50690.2020.9371852>
- Arel, I., Rose, D.C. and Karnowski, T.P., 2010. Deep machine learning-a new frontier in artificial intelligence research [research frontier]. *IEEE computational intelligence magazine*, 5(4), pp.13-18. <https://doi.org/10.1109/MCI.2010.938364>
- Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. and Lew, M.S., 2016. Deep learning for visual understanding: A review. *Neurocomputing*, 187, pp.27-48. <https://doi.org/10.1016/j.neucom.2015.09.116>
- LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning. *nature*, 521(7553), pp.436-444. <https://doi.org/10.1038/nature14539>
- Adhikari, R. and Agrawal, R.K., 2014. A combination of artificial neural network and random walk models for financial time series forecasting. *Neural Computing and Applications*, 24, pp.1441-1449. <https://doi.org/10.1007/s00521-013-1386-y>
- Niu, H. and Wang, J., 2014. Financial time series prediction by a random data-time effective RBF neural network. *Soft computing*, 18, pp.497-508. <https://doi.org/10.1007/s00500-013-1070-2>
- Das, S.P., Achary, N.S. and Padhy, S., 2016. Novel hybrid SVM-TLBO forecasting model incorporating dimensionality reduction techniques. *Applied Intelligence*, 45, pp.1148-1165. <https://doi.org/10.1007/s10489-016-0801-3>
- Guo, T., Xu, Z., Yao, X., Chen, H., Aberer, K. and Funaya, K., 2016, October. Robust online time series prediction with recurrent neural networks. In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 816-825. Ieee. <https://doi.org/10.1109/DSAA.2016.92>
- Ding, G. and Qin, L., 2020. Study on the prediction of stock price based on the associated network model of LSTM. *International Journal of Machine Learning and Cybernetics*, 11(6), pp.1307-1317. <https://doi.org/10.1007/s13042-019-01041-1>
- Liu, Y., 2019. Novel volatility forecasting using deep learning-long short term memory recurrent neural networks. *Expert Systems with Applications*, 132, pp.99-109. <https://doi.org/10.1016/j.eswa.2019.04.038>
- Zeng, X., Cai, J., Liang, C. and Yuan, C., 2022. A hybrid model integrating long short-term memory with adaptive genetic algorithm based on individual ranking for stock index prediction. *Plos one*, 17(8), p.e0272637. <https://doi.org/10.1371/journal.pone.0272637>
- Alotaibi, S.S., 2021. Ensemble technique with optimal feature selection for Saudi stock market prediction: a novel hybrid red deer-grey algorithm. *IEEE Access*, 9, pp.64929-64944. <https://doi.org/10.1109/ACCESS.2021.3073507>
- Chung, H. and Shin, K.S., 2018. Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability*, 10(10), p.3765. <https://doi.org/10.3390/su10103765>

15. Niu, H., Xu, K. and Wang, W., 2020. A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network. *Applied Intelligence*, 50, pp.4296-4309. <https://doi.org/10.1007/s10489-020-01814-0>
16. Qiu, J., Wang, B. and Zhou, C., 2020. Forecasting stock prices with long-short term memory neural network based on attention mechanism. *PloS one*, 15(1), p.e0227222. <https://doi.org/10.1371/journal.pone.0227222>
17. Ali, M., Khan, D.M., Alshanbari, H.M. and El-Bagoury, A.A.A.H., 2023. Prediction of complex stock market data using an improved hybrid emd-lstm model. *Applied Sciences*, 13(3), p.1429. <https://doi.org/10.3390/app13031429>
18. Chen, X., Cao, L., Cao, Z. and Zhang, H., 2024. A multi-feature stock price prediction model based on multi-feature calculation, LASSO feature selection, and Ca-LSTM network. *Connection Science*, 36(1), p.2286188. <https://doi.org/10.1080/09540091.2023.2286188>
19. Gülmez, B., 2023. Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. *Expert Systems with Applications*, 227, p.120346. <https://doi.org/10.1016/j.eswa.2023.120346>
20. Nasiri, H. and Ebadzadeh, M.M., 2023. Multi-step-ahead stock price prediction using recurrent fuzzy neural network and variational mode decomposition. *Applied Soft Computing*, 148, p.110867. <https://doi.org/10.1016/j.asoc.2023.110867>
21. Mandic, D.P. and Chambers, J.A., 1999. Exploiting inherent relationships in RNN architectures. *Neural networks*, 12(10), pp.1341-1345. [https://doi.org/10.1016/S0893-6080\(99\)00076-3](https://doi.org/10.1016/S0893-6080(99)00076-3)
22. Wang, D., Fan, J., Fu, H. and Zhang, B., 2018. Research on optimization of big data construction engineering quality management based on RNN-LSTM. *complexity*, 2018(1), p.9691868. <https://doi.org/10.1155/2018/9691868>
23. Katoch, S., Chauhan, S.S. and Kumar, V., 2021. A review on genetic algorithm: past, present, and future. *Multimedia tools and applications*, 80, pp.8091-8126. <https://doi.org/10.1007/s11042-020-10139-6>
24. Dragomiretskiy, K. and Zosso, D., 2013. Variational mode decomposition. *IEEE transactions on signal processing*, 62(3), pp.531-544. <https://doi.org/10.1109/TSP.2013.2288675>