اقتصاد کشاورزی و توسعه، سال ۳۰، شمارهٔ ۱۲۰، زمستان ۱۴۰۱

DOI: 10.30490/AEAD.2023.357440.1412

مقاله پژوهشی

به‌کارگیری مدل داده‌کاوی هیبریدی (الگوریتم ژنتیک–موجک– شبکهٔ عصبی عمیق– شبیه‌سازی مونت‌کارلو) برای پیش‌بینی قیمت محصولات کشاورزی: مطالعهٔ موردی قیمت آتی زعفران در بورس کالای کشاورزی*

رضا حیدری[۱]، سید محمدرضا حاج سیدجوادی[۲]

تاریخ دریافت: ۱۴۰۰/۱۰/۱۹ تاریخ پذیرش: ۱۴۰۱/۵/۲۳

چکیده

پیش‌بینی قیمت و روند تغییرات آن از مهم‌ترین عوامل در تصمیم‌گیری و تدوین راهبردهای مربوط به محصولات کشاورزی است. هدف مطالعهٔ حاضر ارائه یک مدل یا الگوی داده‌کاوی هیبریدی شامل مجموعه مدل‌های غیرخطی الگوریتم ژنتیک، تبدیل موجک، شبکه عصبی عمیق و روش مونت‌کارلو برای پیش‌بینی دقیق قیمت محصولات کشاورزی بود. این الگوی پیشنهادی از نوع هیبریدی دومرحله‌ای و مدل پایه هیبریدی

---

* مطالعه حاضر برگرفته از طرح پژوهشی داخلی مؤسسهٔ پژوهش‌های برنامه‌ریزی، اقتصاد کشاورزی و توسعهٔ روستایی با عنوان «طراحی مدل پیش‌بینی قیمت قرارداد آتی زعفران در بورس کالای کشاورزی» می‌باشد.

۱- نویسندهٔ مسئول و استادیار پژوهشی، مؤسسهٔ پژوهش‌های برنامه‌ریزی، اقتصاد کشاورزی و توسعهٔ روستایی، تهران، ایران.
(rezaheidari3631@gmail.com)

۲- پژوهشگر، مؤسسهٔ پژوهش‌های برنامه‌ریزی، اقتصاد کشاورزی و توسعهٔ روستایی، تهران، ایران.

غیرخطی- غیرخطی بود و در آن، از الگوریتم ژنتیک برای تعیین وقفه بهینه سری زمانی قیمت، از تابع موجک برای نوفه‌زدایی داده‌های قیمت، از شبکه عصبی عمیق برای پیش‌بینی قیمت، از روش مونت‌کارلو برای شبیه‌سازی محتمل‌ترین احتمال قیمت و در نهایت، از محاسبات پیچیده نرم برای انجام «پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید» برای دوره زمانی دوم تا دهم اردیبهشت ۱۳۹۹ استفاده شد. نتایج مقایسه الگوی پیشنهادی «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» با سه الگوی رقیب «الگوریتم ژنتیک-شبکه عصبی عمیق- مونت‌کارلو»، «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت‌کارلو» و «الگوریتم ژنتیک- شبکه عصبی ساده- مونت کارلو»، با استفاده از معیارهای ارزیابی، نشان داد که الگوی پیشنهادی نسبت به سه الگوی رقیب دارای عملکرد بهتری در پیش‌بینی قیمت زعفران آتی است؛ همچنین، استفاده از شبکه عصبی عمیق در مقایسه با شبکه عصبی ساده و نیز به‌کارگیری نظریهٔ موجک برای نوفه‌زدایی و استفاده از روش مونت‌کارلو برای شبیه‌سازی قیمت‌های پیش‌بینی‌شده دقت پیش‌بینی قیمت آتی زعفران را افزایش می‌دهد. علاوه بر این، استفاده از محاسبات نرم برای انجام «پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید» نشان داد که الگوی پیشنهادی از کارآیی لازم و دقت بالا برای پیش‌بینی کوتاه‌مدت قیمت آتی زعفران برخوردار بوده، به‌گونه‌ای که میزان خطای محاسباتی کمتر از یک درصد (۰/۶ درصد) است. بنابراین، مطالعه حاضر در دستیابی به شاخص میزان دقت حداکثری، سناریوسازی روند قیمت‌های آتی، تحلیل حساسیت مؤلفه‌های مؤثر بر قیمت و سرانجام، پیش‌بینی قیمت آینده از جایگاهی بسیار مناسب برخوردار است. با توجه به نتایج به‌دست‌آمده، استفاده از الگوی پیشنهادی برای پیش‌بینی قیمت محصولات کشاورزی توصیه می‌شود.

کلیدواژه‌ها: *پیش‌بینی، قیمت آتی زعفران، الگوریتم ژنتیک، نظریهٔ موجک، شبکهٔ عصبی عمیق، روش مونت‌کارلو.*

طبقه‌بندی JEL : C45, C53, Q11

مقدمه

قیمت کالاهای کشاورزی نقش مهمی در رفاه افراد جامعه دارد و تغییرات آن بر بخش بزرگی از زندگی مردم در کشورهای کمترتوسعه‌یافته و در حال توسعه تأثیر می‌گذارد. کشورهای سراسر جهان، به‌دلیل افزایش قیمت جهانی مواد غذایی، در معرض ناآرامی‌های اقتصادی و سیاسی قرار دارند (Shiferaw, 2019). طی دوره ۲۰۰۶ تا اواسط ۲۰۰۸، قیمت جهانی کالاهای کشاورزی به‌طور قابل توجهی افزایش یافت، به‌گونه‌ای که قیمت محصولات کشاورزی تقریباً دو برابر شد (Nazlioglu, 2011; Ajmera et al., 2012). طی سال‌های اخیر نیز با وقوع همه‌گیری ویروس کرونا و به‌دنبال آن، جنگ روسیه و اوکراین، قیمت کالاهای کشاورزی بسیار افزایش یافته، به‌گونه‌ای که با روند صعودی در قیمت کالاهای کشاورزی، نگرانی‌های کشورهای واردکننده مواد غذایی و نهاده‌های کشاورزی چند برابر شده است (Nazlioglu and Soytas, 2011).

قیمت محصولات کشاورزی به‌عنوان یک علامت اجتماعی بین عرضه و تقاضا، به‌دلایل متعدد از جمله افزایش تقاضای جهانی، تغییرات پیچیده مرتبط با عوامل چرخه‌ای، تغییرات اقلیمی، سیاست‌ها و تناقض بین تولید و بازاریابی، به‌شدت در نوسان است. افزایش مداوم قیمت مواد غذایی ناشی از افزایش سریع تقاضا برای مواد غذایی بیش از هشت‌صد میلیون نفر در سراسر جهان را به‌دلیل سوءتغذیه مزمن، در معرض تهدید مستقیم قرار داده و از این‌رو، قیمت کالاهای کشاورزی توجه سیاست‌گذاران، محققان دانشگاهی و شرکت‌ها را برای پیش‌بینی قیمت محصولات غذایی به خود جلب کرده است (Weng et al., 2019; Shao and Dai, 2018).

افزایش قیمت مواد غذایی و نوسان‌های آن با اثرات اقتصادی خرد و کلان مهم از جمله اثرات نامطلوب بر تولید ناخالص داخلی یک کشور، تورم، مشکلات خانوار، افزایش فقر، زندگی مالی مردم، کاهش تغذیه و نیز کاهش خدمات ضروری آموزش و پرورش و خدمات بهداشتی همراه است. علاوه بر این، قیمت کالاهای کشاورزی بر بازار زمین‌های کشاورزی، سیاست‌های دولت و کل صنایع وابسته به کشاورزی تأثیر می‌گذارد. نوسان در قیمت کالاهای کشاورزی بهره‌برداران بخش کشاورزی را تحت فشار عاطفی و مالی قرار می‌دهد، زیرا کار سخت آنها در طول سال‌ها به هدر می‌رود (Chuluunsaikhan et al., 2020; Sabu and Kumar, 2020). بنابراین، برای هدایت صحیح تولیدات کشاورزی، ایجاد تعادل صحیح بین عرضه و تقاضا، افزایش درآمد کشاورزان، کمک به کشاورزان برای برنامه‌ریزی کشت محصول بعدی خود، کمک به دولت و کشاورزان و تجار و مصرف‌کنندگان برای به‌دست آوردن آگاهی روشن از بازار، طراحی برنامه‌های تجاری، تنظیم دقیق امور مالی افراد و نیز کاهش خطرات و عدم قطعیت‌های قابل کنترل، دقت در پیش‌بینی قیمت و همچنین، اطلاع قبلی از قیمت‌ها ضروری است (Zhang et al., 2020). پیش‌بینی قیمت محصولات کشاورزی برای بسیاری از بازیگران اقتصادی مهم است و جایگاهی ویژه در تصمیم‌گیری آنها دارد (Ly et al., 2021).

مدل‌های پیش‌بینی قیمت کالاهای کشاورزی مؤلفه‌ای مهم و ضروری در مدیریت بازار و برنامه‌ریزی تصمیم‌گیری محسوب می‌شوند، زیرا پیش‌بینی قیمت کالاهای کشاورزی چشم‌انداز آینده را شبیه‌سازی کرده، به تولیدکنندگان این امکان را می‌دهد تا تصمیمات بهتری بگیرند و به بهینه‌سازی راهبرد فروش محصول خود کمک کنند. بنابراین، چارچوب پیش‌بینی دقیق یک جزء اساسی در پیش‌بینی قیمت محصولات کشاورزی است که با بهره‌گیری از آن، می‌توان از وقوع بسیاری از مسائل و

مشکلات مربوط به تقاضا و عرضه کالاهای کشاورزی جلوگیری کرد (Xiong et al., 2015; Yang et al., 2016).

پس از سال‌ها تحقیق و پژوهش، دستاوردهای پژوهشیِ قابل توجهی در روش‌های پیش‌بینی قیمت محصولات کشاورزی حاصل شده، که عبارت‌اند از: الف) روش‌های تحلیل کیفی شامل روش تحقیق بازار و روش احتمال ذهنی و ب) روش‌های تحلیل کمی شامل روش پیش‌بینی آمار ریاضی، روش تحلیل سری زمانی، روش‌های پیش‌بینی داده‌کاوی اعم از مدل‌های هوشمند ساده و هیبریدی.

سؤال اساسی برای پژوهشگران این است که «در مواجهه با روش‌های متعدد پیش‌بینی قیمت، کدام روش‌ها مناسب‌ترند؟». همین موضوع یک چالش اساسی برای اکثر پژوهشگران است، زیرا درک آنها از مفهوم، ویژگی‌ها و کاربرد مدل‌های پیش‌بینی قیمت، جامع و صحیح نیست. بنابراین، لازم است که مدل‌های پیش‌بینی قیمت محصولات کشاورزی توسعه یابد تا به نتایج دقیق‌تر بینجامد (Wang et al., 2020).

در ایران، کشاورزی از جمله فعالیت‌هایی است که در آن، همواره تصمیم‌گیری در شرایط مخاطره (ریسک) و عدم اطمینان صورت می‌گیرد. تصادفی بودن عرضه محصولات کشاورزی نیز همراه با کشش‌ناپذیری تقاضا برای اکثر محصولات کشاورزی باعث ایجاد نوسان‌های شدید در قیمت محصولات عرضه‌شده در بازار می‌شود. در این میان، بورس کالای کشاورزی ایران به‌عنوان یک بازار جدید با ایجاد بازاری فراگیر، یکپارچه، متشکل و سازمان‌یافته در صدد رفع مشکلات موجود در بازار محصولات کشاورزی برآمده است. زعفران، به‌عنوان گران‌ترین محصول کشاورزی و دارویی جهان، جایگاهی ویژه در بورس کالای کشاورزی ایران دارد. خرید و فروش زعفران در بورس کالای کشاورزی ایران یکی از پررونق‌ترین تجارت‌ها در حوزه ادویه و خشکبار است. بورس کالای زعفران نیز مانند هر بازار دیگری، همیشه با محدودیت‌ها و مسائل ساختاری روبه‌رو بوده، که بخش عمدهٔ این مشکلات مربوط به نوسان‌های قیمتی این محصول است. دولت‌های مختلف، به‌منظور تنظیم بازار و تعیین قیمت زعفران برای مصرف‌کنندگان و صادرکنندگان، تلاش‌های زیادی کرده‌اند، اما به‌دلیل عدم استفاده از ابزارهای مناسب (مانند مدل‌های پیش‌بینی قیمت)، کمتر شاهد بازدهی مناسب در این بخش بوده‌ایم. از طرف دیگر، در قرارداد آتی زعفران، خرید این محصول با قیمت معین در روزهای آتی «امروز» انجام می‌شود و از این‌رو، پیش‌بینی قیمت زعفران با استفاده از مدل‌های مناسب می‌تواند کمک زیادی به خریداران در راستای کاهش مخاطرهٔ قیمتی بازار آتی این محصول باشد. سؤال اساسی این است که «در مواجهه با روش‌های متعدد پیش‌بینی قیمت، برای پیش‌بینی قیمت آتی

زعفران، کدام روش‌ها را باید انتخاب کرد؟». بنابراین، هدف مطالعهٔ حاضر ارائه یک مدل داده‌کاوی هیبریدی به‌منظور پیش‌بینی دقیق قیمت آتی زعفران در بورس کالای کشاورزی در قالب ابزاری مناسب و کارآمد برای نیل به اهداف مطالعه (پیش‌بینی قیمت محصولات کشاورزی) است. از سوی دیگر، با توجه به ماهیت غیرخطی و نامانایی در رفتار سری زمانی قیمت، به‌کارگیری یک مدل هیبریدی متشکل از مدل‌های غیرخطی تبدیل موجک، الگوریتم ژنتیک، شبکه عصبی عمیق و روش مونت‌کارلو یک نوآوری در کسب نتایج دقیق محسوب می‌شود.

پس از مقدمه، در بخش مبانی نظری و ادبیات موضوع، به بررسی انواع مدل‌های داده‌کاوی برای پیش‌بینی قیمت پویای محصولات کشاورزی و برخی از مطالعات پیشین در این زمینه و در بخش مدل تحقیق و روش برآورد، به معرفی ویژگی‌های مدل پیشنهادی و اجزای آن و در بخش نتایج نیز به ارزیابی شیوهٔ استخراج داده‌ها و نتایج حاصل از مدل پیشنهادی و عملکرد آن در مقایسه با مدل‌های رقیب پرداخته می‌شود؛ و در پایان پژوهش نیز نتیجه‌گیری و پیشنهادهایی ارائه می‌شود.

پیشینه تحقیق

بر اساس ادبیات نظری، در سراسر جهان، به‌کارگیری الگوهای داده‌کاوی برای پیش‌بینی قیمت به‌عنوان یک اصل اساسی در عملکرد خوب اقتصادی مورد پذیرش قرار گرفته و البته، در زمینه پیش‌بینی قیمت، از الگوهای مرسوم، هوشمند و هیبریدی زیادی نیز در داخل و خارج کشور استفاده شده است که در ادامه، به برخی از آنها اشاره می‌شود.

نتایج مطالعه رشید و همکاران (Rasheed et al., 2021) نشان داد که مدل حافظهٔ بلند–کوتاه‌مدت[1]، در مقایسه با مدل‌های سری زمانی آماری و یادگیری ماشین معمولی، عملکرد قابل توجهی در پیش‌بینی قیمت گندم دارد. نصار و همکاران (Nassar et al., 2020)، در مقایسه مدل‌های پیش‌بینی قیمت یادگیری عمیق با هشت مدل روی مجموعه داده‌های سری زمانی سبزی‌ها، میوه‌ها و گل‌ها، نشان دادند که مدل حافظهٔ بلند–کوتاه‌مدت (LSTM) و مدل ترکیبی LSTM و شبکهٔ عصبی کانولوشنی[2] در پیش‌بینی دقیق قیمت محصولات تازه کارآمدترند. سابو و کومار (Sabu and Kumar, 2020) از مدل‌های سری زمانی و یادگیری ماشین برای پیش‌بینی قیمت ماهانه دانه درخت خرما[3] در ایالت کرالای هند استفاده کردند و دریافتند که شبکه عصبی LSTM دارای دقت پیش‌بینی

1. Long-Short Term Memory (LSTM)
2. Convolution Neural Networks (CNN)
3. arecanut

بالاتری نسبت به سایر مدل‌های رقیب است. در مطالعه دیگری، چن و همکاران (Chen et al., 2019) نشان دادند که با استفاده از تحلیل موجک، نوفه داده‌های سری زمانی قیمت کلم کاهش می‌یابد و پیش‌بینی با مدل LSTM روی داده‌های نرمال‌سازی‌شده نیز دقت پیش‌بینی در قیمت کلم را بهبود می‌بخشد. وانگ و همکاران (Wang et al., 2018) از یک مدل هیبریدی برای پیش‌بینی قیمت ماهانه سیر چینی طی سال‌های ۲۰۱۷–۲۰۱۰ استفاده کردند؛ مدل پیشنهادی آنها متشکل از مدل میانگین متحرک خودبازگشتی[1] به‌عنوان بخش خطی و مدل ماشین بردار پشتیان[2] به‌عنوان بخش غیرخطی بود. نتایج این مطالعه نشان داد که مدل هیبریدی ARIMA-SVM نسبت به دو مدل ARIMA و SVM عملکرد بهتری در پیش‌بینی قیمت سیر دارد و می‌توان از آن به‌عنوان روشی مؤثر برای پیش‌بینی قیمت کوتاه‌مدت سیر استفاده کرد. در مطالعه وانگ و همکاران (Wang et al., 2017)، از یک مدل هیبریدی برای پیش‌بینی قیمت آتی گندم، ذرت و سویا استفاده شد. در این روش، چهار مدل هیبریدی بر پایه شبکه عصبی انتشار برگشتی[3] و بهینه‌سازی ازدحام ذرات[4] پیشنهاد شده است. نتایج این مطالعه نشان داد که همه «مدل‌های هیبریدی همراه با شیوهٔ تجزیه» عملکرد بهتری نسبت به مدل‌های انفرادی دارند. پلیام و بونراد (Polyiam and Boonrawd, 2017)، برای پیش‌بینی قیمت کاساوا (نوعی سبزی ریشه‌ای[5]) در تایلند در دوره ۲۰۱۵–۲۰۰۵، از یک مدل هیبریدی متشکل از شبکه عصبی مصنوعی و ماشین بردار پشتیبان (SVM-ANN) استفاده کردند. نتایج این مطالعه نشان داد که مدل هیبریدی پیشنهادی دارای کمترین خطا نسبت به مدل‌های دیگر است. همچنین، نتایج مطالعات لی و همکاران (Li et al, 2014)، وانگ و گوپتا (Wang and Gupta, 2013)، هاون و همکاران (Haven et al., 2012) و جمازی و آلویی (Jammazi and Aloui, 2012) نشان داد که استفاده از تبدیل موجک برای نوفه‌زدایی داده‌ها موجب افزایش دقت و عملکرد الگوی مورد استفاده برای پیش‌بینی می‌شود.

افزون بر این، در مطالعه محمدی و همکاران (Mohammadi et al., 2017)، با استفاده از ترکیب تبدیل موجک و شبکه عصبی، مدلی به‌منظور پیش‌بینی روزانه قیمت گاز طبیعی در ایران ارائه شد. نتایج این مطالعه نشان داد که استفاده از تبدیل موجک نوفه‌های داده‌های قیمت گاز را کاهش داده و مدل

1. Autoregressive Moving-Average Model (ARIMA)
2. Support Vector Machine (SVM)
3. Back Propagation Neural Network (BPNN)
4. Particle Swarm Optimization (PSO)
5. cassava

هیبریدی تبدیل موجک- شبکه عصبی، عملکرد پیش‌بینی قیمت گاز را در مقایسه با مدل شبکه عصبی بهبود بخشیده است. صادقی و دهقانی فیروزآبادی (Sadeghi and Dehghani Firoozabadi, 2017) از تحلیل موجک برای نوفه‌زدایی از سری‌های زمانی بورس اوراق بهادار تهران استفاده کردند. در این مطالعه برای نوفه‌زدایی از دو روش دندروگرام و تابع موجک هار و دابشیز استفاده شد و نتایج مطالعه حکایت از عملکرد بهتر نوفه‌زدایی با استفاده از موجک دابشیز در سری‌های زمانی مورد نظر داشت. راعی و همکاران (Rai et al., 2015) به پیش‌بینی شاخص قیمت بورس سهام با استفاده از مدل ترکیبی شبکه عصبی و تبدیل موجک از ابتدای فروردین سال ۱۳۸۵ تا پایان خرداد ۱۳۹۲ پرداختند. نتایج این مطالعه حاکی از بهبود معنی‌دار در پیش‌بینی شبکه عصبی با استفاده از داده‌های نوفه‌زدایی‌شده است. در تحقیقی دیگر، راعی و محمودی آذر (Rai and Mahmoudi Azar, 2014)، با استفاده از مدل‌های آریما، شبکه عصبی و نوفه‌زدایی موجک، اقدام به پیش‌بینی بازده آتی بازار سهام کردند. نتایج این مطالعه نشان داد که کاهش نوفه داده‌ها عملکرد پیش‌بینی بازده شاخص را بهبود می‌بخشد؛ به بیان بهتر، مدل شبکه عصبی- موجکی (نوفه‌زدایی سیگنال) عملکردی بهتر از مدل‌های آریما و شبکه عصبی دارد و مدل‌های شبکه عصبی نیز قدرت پیش‌بینی بهتری را نسبت به مدل‌های آریما نشان می‌دهد؛ همچنین، مقادیر مربوط به آزمون دایبولد- ماریانو این نتایج را تأیید می‌کند.

بررسی ادبیات نظری پیرامون پیش‌بینی قیمت محصولات کشاورزی در مطالعهٔ حاضر نشان داد که: ۱- مطالعات زیادی با مدل‌های مختلف برای پیش‌بینی قیمت‌های بسیاری از کالاهای کشاورزی در خارج از ایران در مواردی به‌صورت تک‌مرحله‌ای استفاده شده است؛ ۲- تحقیقات کاربردی بسیار کمی با کاربرد مدل‌های هیبریدی بر پایه مدل‌های یادگیری عمیق در پیش‌بینی قیمت کالاهای کشاورزی وجود دارد؛ ۳- تا آنجا که تاکنون هیچ مدل هیبریدی از نوع مدل پیشنهادی مطالعهٔ حاضر با به‌کارگیری و معرفی اجزا (تبدیل موجک برای نوفه‌زدایی، الگوریتم ژنتیک برای تعیین وقفه بیهینه، شبکه عصبی عمیق برای پیش‌بینی و روش مونت‌کارلو برای شبیه‌سازی) به‌منظور پیش‌بینی قیمت محصولات کشاورزی ارائه نشده است.

مدل تحقیق و روش برآورد

استفاده از مدل‌های هیبریدی یک راه معمول در بهبود دقت پیش‌بینی‌ها و غلبه بر محدودیت‌های مدل‌های تکی است. ادبیات موضوع مربوط به مدل‌های هیبریدی بسیار گسترده بوده

و مطالعات فراوان در این زمینه صورت گرفته است. اصل اساسی و منطقی در هیبرید کردن مدل‌ها بر این پایه استوار است که به‌کارگیری یک روش از روش‌های موجود رویکردی جامع برای پیش‌بینی به‌طور حتمی نیست و قابلیت به‌کارگیری در هر شرایط و هر نوع داده را ندارد. از این‌رو، با هیبرید کردن مدل‌های مختلف، می‌توان نقاط ضعف یک مدل را با استفاده از نقاط قوت مدل‌های دیگر بهبود بخشید. یافته‌های تجربی و نظری نیز نشان می‌دهد که هیبرید مدل‌های متفاوت، یک راه مؤثر و کارآ برای بهبود دقت پیش‌بینی‌هاست (Moghaddasi and Jaleh Rajabi, 2011). با توجه به اهمیت مدل‌های هیبریدی و استفاده از آن در مطالعهٔ حاضر، در ادامه، خلاصه‌ای از معرفی این نوع مدل‌ها ارائه می‌شود.

به‌طور کلی، مفهوم مدل‌های هیبریدی شامل روش‌هایی است که به ترکیب نتایج پیش‌بینی مدل‌های مختلف با هم می‌پردازند تا نتیجه پیش‌بینی جدید و دقیق‌تری به‌دست آید. متداول‌ترین راهبرد هیبریدی تلاش برای یافتن وزن‌ها با استفاده از رگرسیون شناخته شده و بیشترین کاربرد را تاکنون داشته است. در این راهبرد، مدل‌های منفرد تکمیل می‌شود، به‌گونه‌ای که قوانین بالقوه در یک دنباله سری زمانی را می‌توان با دقت بیشتری دنبال کرد. مدل‌های هیبریدی را می‌توان در دو دسته خلاصه کرد: ۱- مدل‌های هیبریدی عمومی[1] و ۲- مدل‌های هیبریدی مبتنی بر چارچوب «تجزیه و جمع‌بندی»[2]. در مدل‌های هیبریدی عمومی، اصل مدل مبتنی بر تقسیم‌بندی قیمت محصولات کشاورزی به‌صورت افقی (موازی) به دو قسمت و سپس، پیش‌بینی این دو قسمت به‌طور جداگانه با روش‌های مختلف است که سرانجام، با روی هم قرار دادن آنها، نتیجه پیش‌بینی به‌دست می‌آید. گرچه مدل‌های هیبریدی عمومی تا حدی عملکرد پیش‌بینی را بهبود می‌بخشند، اما اغلب نمی‌توانند به‌طور کامل، با نامانایی سری‌های زمانی تصادفی و نامنظم مقابله کنند. بنابراین، بر اساس روش‌های تجزیه رایج مانند تجزیه و تحلیل موجک[3] (Reboredo and Rivera-Castro, 2013)، تجزیه تجربی[4] (Zhang et al., 2008) و تجزیه و تحلیل طیف منفرد[5] (Vautard et al., 1992)، یک مفهوم امیدوارکننده از «تجزیه و جمع‌بندی» برای افزایش توانایی پیش‌بینی مدل‌های موجود در قالب مدل‌های هیبریدی مبتنی بر چارچوب «تجزیه و جمع‌بندی» پدید آمده، که با مدل هیبریدی

1. general hybrid model
2. hybrid model based on the “decomposition and ensemble” framework
3. Wavelet Analysis (WA)
4. Empirical Mode Decomposition (EMD)
5. Singular Spectrum Analysis (SSA)

عمومی در نحوه اجرا (محاسبه نتایج) متفاوت است، به‌گونه‌ای که در این نوع مدل هیبریدی، ابتدا قیمت محصولات کشاورزی در چندین مرحله مورد تجزیه و تحلیل قرار می‌گیرد و سپس، با استفاده از روش‌های منتخب، طی یک مسیر مشخص، این اجزا به‌طور جداگانه تکمیل می‌شوند؛ و در نهایت، با دریافت یک درک کلی از نتایج قیمت، نتایج پیش‌بینی حاصل می‌شود.

به‌طور کلی، در مدل‌های هیبریدی، ممکن است تجزیه و جمع‌بندی فرآیند تبدیل اجرا از طریق مدل‌های منتخب به یکدیگر تک‌مرحله‌ای[1] یا دومرحله‌ای[2] باشد. در روش تک‌مرحله‌ای، مدل پایه در بخش ورودی داده‌ها یا داده‌های خروخی (نتایج مدل پایه) هیبرید می‌شود. اما، در روش دومرحله‌ای، هر دو بخش فرآیند هیبریدی در هر دو بخش ورودی داده‌ها و خروجی نتایج اتفاق می‌افتد. همچنین، بیان این نکته ضروری است که ساختار مدل‌های هیبریدی نیز متفاوت است و با توجه به میزان اهمیت دقت پیش‌بینی در مطالعه مورد نظر، به سه حالت کلی تقسیم می‌شود، که عبارت‌اند از: ۱- مدل پایه هیبریدی خطی- خطی (مدل هیبریدی ابتدایی)، ۲- مدل پایه هیبریدی خطی- غیرخطی (مدل هیبریدی متوسط) و ۳- مدل پایه هیبریدی غیرخطی- غیرخطی (مدل هیبریدی پیشرفته). مدل هیبریدی مورد استفاده در مطالعظ حاضر از نوع هیبریدی دومرحله‌ای و مدل پایه هیبریدی غیرخطی- غیرخطی بوده و از این‌رو، کامل‌ترین نوع مدل هیبریدی است.

در شکل ۱، شمای کلی مدل داده‌کاوی هیبریدی پیشنهادی برای پیش‌بینی قیمت محصولات کشاورزی (در مطالعهٔ حاضر، قیمت آتی زعفران نگین) نشان داده شده است. ویژگی‌های مدل پیشنهادی پژوهش حاضر یعنی، «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» (GA-WA-LSTM-MCM) عبارت‌اند از:

۱- مطابق با مبانی نظری، این مدل پیشنهادی از نوع هیبریدی دومرحله‌ای و مدل پایه غیرخطی- غیرخطی است. توضیح آنکه در مدل پیشنهادی، داده‌های ورودی به مدل شبکه عصبی عمیق (LSTM) توسط مدل ناپارامتریک تبدیل موجک گسسته (WA) نوفه‌زدایی می‌شود. همچنین، خروجی مدل LSTM توسط مدل غیرخطی مونت‌کارلو شبیه‌سازی می‌شود تا بالاترین احتمال رخداد قیمت در آینده به‌دست آید.

1. One-step integrated
2. two-step integrated

۲- از الگوریتم ژنتیک (GA) برای تعیین وقفه بهینه سری زمانی قیمت، از تابع موجک برای نوفه‌زدایی سری زمانی قیمت و از شبکه عصبی عمیق برای پیش‌بینی سری زمانی قیمت استفاده شده است.

۳- از شیوهٔ مونت‌کارلو (MCM) برای شبیه‌سازی محتمل‌ترین احتمال قیمت به‌همراه آزمون قیمت‌های جدید[1] در تکمیل فرآیند پیش‌بینی استفاده شده است.

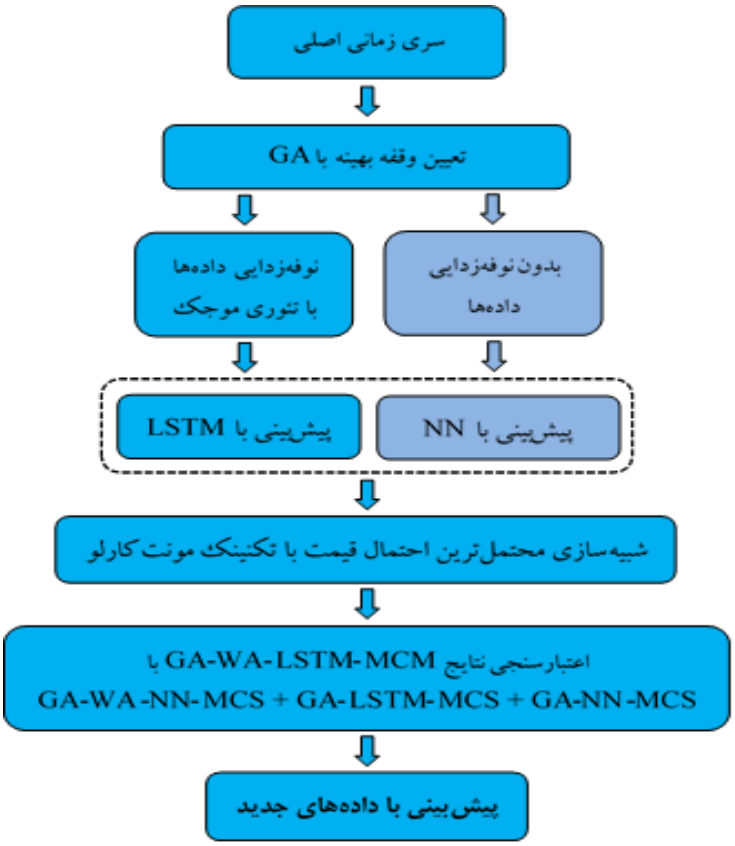
۴- استفاده از روابط و معادلات محاسبات پیچیده نرم[2] برای انجام «پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید»[3] یکی دیگر از نوآوری‌های مطالعهٔ حاضر است.

برای نشان دادن دقت پیش‌بینی مدل پیشنهادی «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» با استفاده از معیارهای ارزیابی، خروخی حاصل از پیش‌بینی مدل پیشنهادی مطالعهٔ حاضر با سه مدل دیگر یعنی، مدل‌های «الگوریتم ژنتیک- شبکه عصبی عمیق- مونت‌کارلو»، «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت‌کارلو» و «الگوریتم ژنتیک- شبکه عصبی ساده- مونت‌کارلو» مقایسه شده است. علت انتخاب مدل‌های رقیب نزدیک به مدل پیشنهادی این است که اثر هر کدام از اجزای الگوی پیشنهادی در افزایش دقت پیش‌بینی مشخص شود. برای نمونه، در الگوی رقیب «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت‌کارلو»، به‌جای شبکه عصبی عمیق در الگوی پیشنهادی، از شبکه عصبی ساده استفاده شده است تا کارایی افزایش دقت پیش‌بینی توسط شبکه عصبی عمیق در مدل پیشنهادی مشخص شود. همچنین، در دو مدل رقیب «الگوریتم ژنتیک- شبکه عصبی عمیق- مونت‌کارلو» و «الگوریتم ژنتیک- شبکه عصبی ساده-مونت‌کارلو»، تابع موجک حذف شده است تا اثر نوفه‌زدایی داده‌ها در افزایش دقت پیش‌بینی قیمت محصولات کشاورزی مشخص شود. در ادامه، هر کدام از اجزای مدل پیشنهادی معرفی می‌شوند.

1. new credit data

۲- محاسبات پیچیده نرم (soft computing)؛ به مجموعه‌ای از شیوه‌های جدید محاسباتی در علوم رایانه، هوش مصنوعی، یادگیری ماشینی و بسیاری از زمینه‌های کاربردی دیگر اطلاق می‌شود. در تمامی این زمینه‌ها به مطالعه، مدل‌سازی و تحلیل پدیده‌هایی بسیار پیچیده نیاز است که در گذشته، شیوه‌های علمی دقیق در حل آسان، تحلیلی و کامل آنها موفق نبوده‌اند.

3. deployment score new data



مأخذ: یافته‌های پژوهش

شکل ۱- شمای کلی مدل پیشنهادی برای پیش‌بینی قیمت محصولات کشاورزی

تبدیل موجک

پیشینه پژوهش نشان داد که استفاده از تبدیل موجک در نوفه‌زدایی داده‌های سری زمانی قیمت سبب افزایش دقت پیش‌بینی مدل می‌شود. در پژوهش حاضر، نخست، نوفه‌زدایی داده‌ها با استفاده از تبدیل موجک صورت می‌گیرد و سپس، داده‌های نوفه‌زدایی‌شده به‌عنوان داده‌های ورودی برای پیش‌بینی به شبکه عصبی عمیق منتقل می‌شود. به دیگر سخن، تبدیل موجک به‌عنوان یک روش پیش‌پردازش به‌کار می‌رود. تبدیل موجک یک ابزار قوی ریاضی است که یک سری زمانی مانا یا نامانا را به دو بخش تقریبات و جزئیات تقسیم می‌کند. تبدیل موجک با استفاده از توابع پایه‌ای، یک

سری زمانی را به فضای فرکانس می‌برد و سپس، سری زمانی را در زمان و مقیاس‌های مختلف نشان می‌دهد. تبدیل موجک قابلیت استفاده برای هر دو سری‌های زمانی پیوسته و گسسته را دارد. زمانی که سری زمانی گسسته باشد، باید از موجک گسسته مادر استفاده کرد. از جمله توابع مهم تبدیل موجک گسسته می‌توان به توابع هار، دابیشز، سیملت، کویفلت و میر اشاره کرد. به‌طور کلی، موجک گسسته مادر به صورت رابطه (۱) تعریف می‌شود:

$$\Psi_{m,n}\left(\frac{t-u}{s}\right)=\frac{1}{\sqrt{s_o^{\frac{m}{2}}}}\Psi\left(\frac{t-nu_o s_o^m}{s_o^m}\right) \qquad (۱)$$

که در آن، n مکان موجک و m اندازه موجک است. معمولاً برای ساده‌تر شدن مقادیر s_o =۲ و u =۱ در نظر گرفته می‌شود. در این حالت، تابع موجک گسسته مادر به‌صورت رابطه (۲) درمی‌آید:

$$w_{m,n}=2^{-\frac{m}{2}}\sum_{t=0}^{N-1}\Psi(2^{-m}t-n)x(t) \qquad (۲)$$

که در آن، $w_{m,n}$ ضرایب حاصل از استفاده از تابع موجک برای سری زمانی گسسته در حالت $s=۲^m$ و $u=۲^m \times n$ است. بر اساس نظریه مالات[1] در سال ۱۹۹۸، یک سری زمانی x(t) می‌تواند به دو بخش تقریبات و جزئیات تجزیه شود. بنابراین، سری زمانی اصلی یک تابع گسسته با استفاده از معکوس تابع موجک به‌صورت رابطه (۳) به‌دست می‌آید:

$$x(t)=T+\sum_{m=1}^{M}\sum_{t=0}^{M-m-1}w_{m,n}2^{-\frac{m}{2}}\Psi(2^{-m}t-n) \qquad (۳)$$

شکل ساده رابطه (۳) به‌صورت رابطه (۴) است:

$$x(t)=A_M(t)+\sum_{m=1}^{M}D_M(t) \qquad (۴)$$

که در آن، $A_m(t)$ بخش تقریبات یا به بیان دیگر، جزء هموارشدهٔ سری زمانی و $D_m(t)$ (m=1,2,…,M) جزئیات و یا نوفه موجود در سری زمانی است (Shabri and Samsudin, 2014).

1. Malat theory

الگوریتم ژنتیک

الگوریتم ژنتیک یکی از روش‌های جدید بهینه‌سازی است که بیشتر برای بهینه‌سازی مسائل بسیار پیچیده و غیرخطی به‌کار می‌رود. الگوریتم ژنتیک الهامی از علم ژنتیک و نظریهٔ تکامل داروین بوده و بر اساس بقای برترین‌ها یا انتخاب طبیعی استوار است. ایده اساسی این الگوریتم انتقال خصوصیات موروثی توسط کروموزوم‌هاست. این الگوریتم با الهام از طبیعت بر پایهٔ اصل تکاملی «پایداری بهترین‌ها» استوار است. یکی از برتری‌های این الگوریتم نسبت به سایر مدل‌های مشابه عدم وابستگی آن به فرضیه‌های آماری محدودکننده و نرمال بودن توزیع نسبت‌ها یا برابری واریانس یا کوواریانس ماتریس نسبت‌هاست (Yang, 2007). الگوریتم ژنتیک، با تعریف یک کروموزوم که باید بهینه باشد، شروع می‌شود. هر درایه در هر کروموزوم «ژن» نامیده می‌شود. مجموعه‌ای از این کروموزوم‌ها جمعیت الگوریتم را شکل می‌دهند. هر تکرار برای بهینه‌سازی فرآیند مسئله داده‌شده به‌عنوان یک نسل شناخته می‌شود. جمعیت افراد با استفاده از نسل‌های مداوم در هر نسل جدید توسعه می‌یابد. در هر مرحله، جمعیت جدید تولیدشده از نظر ژنتیکی بهتر از جمعیت قبلی است و شانس بیشتری برای ادامه حیات دارد. این روند تا حصول یک راه حل بهینه یا شرایط رضایت‌بخش توسط کاربران ادامه می‌یابد. تبدیل جمعیت افراد به نسل جدید توسط سه عامل اصلی ژنتیکی انتخاب[1]، آمیزش[2] و جهش[3] انجام می‌شود. بعد از ایجاد نسل‌های متعدد، الگوریتم ژنتیک روی راه حل بهینه همگرا می‌شود. اجرای موفقیت‌آمیز الگوریتم ژنتیک بستگی به ترکیب تعادلی پارامترهای کنترلی دارد که شامل احتمال آمیزش، احتمال جهش و اندازه جمعیت است (Sangwana et al., 2016; Raikar et al., 2016).

شبکه عصبی عمیق

در حال حاضر، مدل‌های شبکه عمیق از دقیق‌ترین روش‌های در دسترس برای پیش‌بینی به‌شمار می‌روند. به‌طور کلی، شبکه‌های عمیق از پشت سر هم قرار دادن شبکه‌های کم‌عمق (مانند شبکه عصبی تک‌لایه) تشکیل می‌شوند. در نتیجه، شبکه‌های عمیق قادر به یادگیری و استخراج ویژگی‌های سلسه‌مراتبی غیرخطی است و می‌تواند الگوهای آماری پیچیده را ذخیره کند. مدل‌های عمیق کارآیی خود را برای کاربردهای مختلف به اثبات رسانده‌اند. شبکه‌های عصبی

1. selection
2. crossover
3. mutation

عمیق در ساختار خود از نوع یادگیری بدون نظارت[1] استفاده می‌کنند، زیرا داده‌های امروزی که برای پردازش آنها از هوش مصنوعی استفاده می‌شود، بسیار زیاد و شامل میلیون‌ها داده است و تهیه این داده‌ها به‌صورت دستی عملاً غیرممکن است. از این‌رو، شبکه‌های عصبی عمیق خود داده‌های ورودی را پردازش کرده، شروع به تعلیم خود می‌کنند و هرچه داده‌های بیشتری در اختیار آن قرار گیرد، قدرتمندتر می‌شود و به‌راحتی بسیار هوشمندتر از انسان عمل می‌کند. در این نوع شبکه عصبی، ابتدا ساختارهای ابتدایی داده‌های ورودی را دریافت می‌کنند و سپس، با ترکیب توابع خطی و غیرخطی بسیار، ساختارهایی پیچیده و سطح بالا از داده‌های ورودی را تولید و پردازش می‌کنند.

پس از مشاهده کاستی‌هایی در کاربرد شبکه عصبی بازگشتی[2]، مفهومی به‌عنوان حافظه بلندمدت و کوتاه‌مدت در شکبه عصبی پدید آمد که به‌اختصار، مدل حافظهٔ بلند–کوتاه‌مدت (LSTM) نامیده شد. مدل LSTM، برخلاف شبکه عصبی بازگشتی (RNN)، در مواجهه با دنباله‌های طولانی مشکلی ندارد و با سازوکار طراحی‌شده در این‌گونه الگوها به‌خوبی اجازه کار با دنباله‌های طولانی‌تر را می‌دهند (Goodfellow et al., 2016). به دیگر سخن، این مدل برخلاف شبکه عصبی بازگشتی سنتی که صرفاً جمع متوازن سیگنال‌های ورودی را محاسبه می‌کند و سپس، از یک تابع فعال‌سازی عبور می‌دهد، مدل LSTM از یک حافظه C_t در زمان t بهره می‌برد. خروجی h_t و یا فعال‌سازی واحد LSTM به‌صورت رابطه (۵) تعریف می‌شود و در آن، Γ_o دروازه خروجی کنترل‌کنندهٔ میزان محتوایی است که از طریق حافظه ارائه می‌شود. دروازه خروجی نیز از طریق رابطه (۶) محاسبه می‌شود که در آن، σ تابع فعال‌سازی softmax است. عبارت W_o نیز یک ماتریس اریب است. سلول حافظه C_t نیز با فراموشی نسبی حافظه فعلی و اضافه کردن محتوای حافظه جدید به‌صورت $\widehat{C_t}$ در قالب رابطه (۷) به‌روزرسانی می‌شود که در آن، محتوای حافظه جدید از طریق رابطه (۸) به‌دست می‌آید. این میزان از حافظه فعلی که باید فراموش شود، توسط دروازه فراموشی F_t کنترل می‌شود و آن میزانی که از محتوای حافظه جدید که باید به سلول حافظه جدید اضافه شود، توسط دروازه به‌روزرسانی انجام می‌شود. این عمل نیز با محاسبه توسط روابط (۹) و (۱۰) انجام می‌گیرد:

$$h_t = \Gamma_o \tanh(C_t) \quad (۵)$$

$$\Gamma_o = \sigma(W_0.[h_{t-1}, X_t] + b_o) \quad (۶)$$

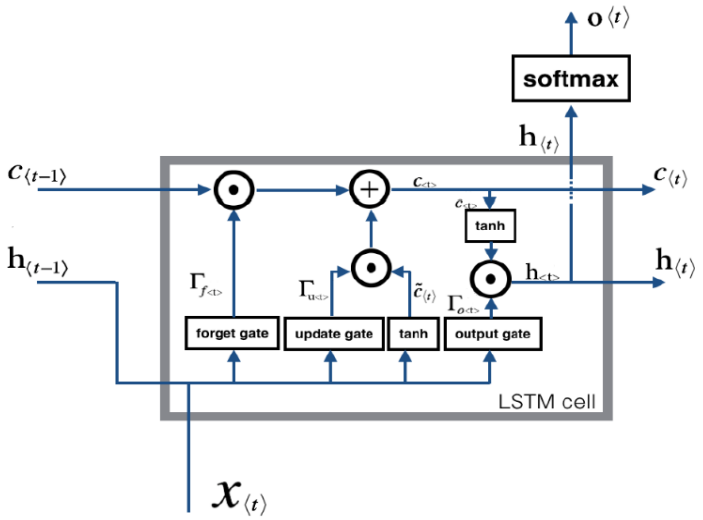
1. unsupervised learning
2. Recurrent Neural Network (RNN)

$$C_t = \Gamma_f . C_{t-1} + \Gamma_u . \widehat{C_t} \quad (۷)$$

$$\widehat{C_t} = tanh(W_c . [h_{t-1}, X_t] + b_c) \quad (۸)$$

$$\Gamma_f = \sigma(W_f . [h_{t-1}, X_t] + b_f) \quad (۹)$$

$$\Gamma_u = \sigma(W_u . [h_{t-1}, X_t] + b_u) \quad (۱۰)$$



شکل ۲- ساختار شبکه عصبی عمیق (RNN-LSTM)

در مدل شبکه عصبی LSTM، مفاهیمی جدید وجود دارد که در شبکه عصبی بازگشتی سنتی وجود ندارند. در این شبکه، به‌اصطلاح، سه دروازه وجود دارد که از طریق آنها، شبکه نسبت به کنترل جریان داده درون خود اقدام می‌کند. این سه دروازه عبارت‌اند از: ۱- دروازه نسیان یا فراموشی[1]، ۲- دروازه به‌روزرسانی[2] که به دروازه ورودی معروف است و ۳- دروازه خروجی[3]. علاوه بر این سه دروازه، یک سلول حافظه[4] نیز وجود دارد که از آن به‌اختصار با عنوان C یاد می‌شود. اینها مفاهیم تازه در این شبکه به‌شمار می‌روند و شبکه، علاوه بر این چهار مفهوم جدید دارای یک ورودی، از حافظه

1. forget gate
2. update gate
3. input gate or output gate
4. memory cell

پنهان یا همان h و ورودی یا همان x نیز بهره می‌برد و دو خروجی تولید می‌کند (یک خروجی C_t و خروجی دیگر، h_t است که خود به دو بخش تقسیم می‌شود: بخشی به گام زمانی بعد منتقل می‌شود و بخشی نیز در صورت نیاز به تولید خروجی، در گام زمانی فعلی مورد استفاده قرار می‌گیرد).

دروازه فراموشی که در عبارات بالا به‌صورت Γ_f نمایش داده شده است، وظیفه کنترل جریان اطلاعات از گام زمانی قبلی را دارد. این دروازه مشخص می‌کند که «آیا اطلاعات حافظه از گام زمانی قبل مورد استفاده قرار گیرد یا خیر؟» و اگر باید از گام زمانی قبل چیزی وارد شود، «به چه میزان باشد؟». دروازه به‌روزرسانی که در عبارات بالا به‌صورت Γ_u نمایش داده شده است، وظیفه کنترل جریان اطلاعات جدید را بر عهده دارد. این دروازه مشخص می‌کند که «آیا در گام زمانی فعلی باید از اطلاعات جدید استفاده شود یا خیر؟» و اگر بلی، «به چه میزان؟». همچنین، دروازه خروجی که در عبارات بالا به‌صورت Γ_o نمایش داده شده است، مشخص می‌کند که «چه میزان از اطلاعات گام زمانی قبل با اطلاعات گام زمانی فعلی به گام زمانی بعد منتقل شود؟» (Zolfaghari et al., 2020).

روش مونت‌کارلو

روش مونت‌کارلو به هر شیوه‌ای گفته می‌شود که از طریق نمونه‌سازی آماری، پاسخ‌های تقریبی برای مسائل کمی فراهم می‌کند. شبیه‌سازی مونت‌کارلو بیشتر به‌منظور توصیف روشی برای انتشار عدم قطعیت‌های موجود در ورودی مدل به عدم قطعیت‌ها در خروجی مدل به‌کار می‌رود. بنابراین، مونت‌کارلو نوعی شبیه‌سازی است که صریحاً و به‌صورت کمی، عدم قطعیت را نمایش می‌دهد. این روش متکی به فرآیند نمایش صریح عدم قطعیت با تعیین ورودی‌ها به‌عنوان توزیع‌های احتمال است. اگر ورودی‌های توصیف‌کننده یک نظام غیرقطعی باشند، آنگاه پیش‌بینی عملکرد پیشرو الزاماً غیرقطعی است. این بدان معنی است که نتیجه هر گونه تحلیل مبتنی بر ورودی‌های نمایش داده‌شده با توزیع‌های احتمال، خود یک توزیع احتمال است.

هر الگوریتم محاسباتی و مدل‌سازی مونت‌کارلو در شکل ساده یا پیچیده شامل چند جزء اصلی است. این اجزا عبارت‌اند از اعداد تصادفی، تابع توزیع احتماال، قواعد نمونه‌برداری، تخمین خطا و کاهش انحراف معیار پراکندگی، موازی‌سازی- بردارسازی محاسباتی و الگوریتم متروپلیس. هسته اصلی هر روش مدل‌سازی تصادفی یا روش مونت‌کارلو بر مبنای استفاده مداوم از اعداد تصادفی است. روش مونت‌کارلو یک الگوریتم محاسباتی است که از نمونه‌گیری تصادفی برای محاسابه نتایج استفاده می‌کند. از آنجا که نتیجه شبیه‌سازی یک نظام غیرقطعی یک گزارش مشروط است، برای نمونه، «اگر سد بسازیم، ماهی‌های سالمون منقرض می‌شوند»، نتیجه یک شبیه‌سازی احتمالی (مونت‌کارلو) نیز

یک احتمال مشروط است: «اگر سد بسازیم، به احتمال بیست درصد، ماهی‌های سالمون منقرض می‌شوند». این نتیجه (در این مورد، بیان کمی احتمال منقرض شدن) اغلب برای تصمیم‌گیرندگانی که از نتایج شبیه‌سازی استفاده می‌کنند، بسیار مفیدتر است.

به‌منظور محاسبه توزیع احتمال کارآیی پیش‌بینی‌شده، لازم است که عدم قطعیت‌های ورودی به عدم قطعیت‌های خروجی منتقل شود که برای این کار، روش‌های گوناگون وجود دارد. شبیه‌سازی مونت‌کارلو احتمالاً رایج‌ترین شیوهٔ انتشار عدم قطعیت موجود در جنبه‌های مختلف یک نظام به کارآیی پیش‌بینی‌شده است. در شبیه‌سازی مونت‌کارلو، کل نظام به تعداد دفعات زیادی اجرا می‌شود (برای نمونه، صد هزار بار). به هر بار شبیه‌سازی، تحقق[1] نظام گفته می‌شود. برای هر تحقق، تمام پارامترهای غیرقطعی نمونه‌برداری می‌شود (یعنی، یک مقدار تصادفی از توزیع اختصاصی مربوط به هر پارامتر انتخاب می‌شود)؛ سپس، این نظام در طول زمان (با معین بودن مجموعه پارامترهای ورودی) شبیه‌سازی می‌شود، به‌گونه‌ای که کارآیی نظام بتواند محاسبه شود. این کار منتج به ایجاد تعداد زیادی نتیجه مستقل و جداگانه می‌شود، که هر کدام نمایشگر یک «آینده» احتمالی برای نظام است (یعنی، یک مسیر احتمالی که نظام احتمالاً با گذشت زمان دنبال خواهد کرد). نتایج تحقق‌های مستقل نظام به شکل توزیع‌های احتمالی خروجی‌های ممکن درخواهند آمد. در نتیجه، خروجی‌ها به‌صورت مقادیر تک نیستند، بلکه توزیع احتمال هستند. در حقیقت، روش مونت‌کارلو یک روش شبیه‌سازی است که با در نظر گرفتن پویایی نظام به پیش‌بینی تغییرات آینده آن پرداخته، این تغییرات را شبیه‌سازی می‌کند (Alinejad et al., 2017; Ghias, 2014).

معیارهای ارزیابی پیش‌بینی مدل‌ها

معمولاً برای ارزیابی مدل‌های پیاده‌سازی‌شده، از برخی شاخص‌های آماری استفاده می‌شود که به ارزیابی تفاوت و انحراف بین داده‌های واقعی و پیش‌بینی‌شده می‌پردازند. در مطالعهٔ حاضر، از شاخص‌های مجذور میانگین مربعات خطا[2] و ضریب تعیین (R^2) استفاده می‌شود. هر کدام از این دو شاخص می‌تواند مدل پیاده‌سازی‌شده در مطالعهٔ حاضر را اعتبارسنجی کند، به‌گونه‌ای که مدلی دارای بهترین عملکرد است که کمترین مقدار مجذور میانگین مربعات (RMSE) یا بیشترین مقدار R^2 را داشته باشد. شاخص‌های یادشده به‌صورت روابط زیر است:

1. realization
2. Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \quad (۱۱)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2} \quad (۱۲)$$

که در این روابط، Y_i داده‌های واقعی (مشاهداتی)، $\hat{Y}_i$ داده‌های پیش‌بینی‌شده (شبیه‌سازی‌شده)، $\overline{Y}_i$ میانگین داده‌های واقعی و n دوره زمانی ارزیابی است (Yang et al., 2016). برای انجام پژوهش حاضر، از نرم‌افزارهای متلب[1]، رپید ماینر[2]، کلمنتین[3] و پایتون[4] استفاده شد.

داده‌ها و نتایج تجربی

در تحقیق حاضر، برای یافتن بهترین مدل برای پیش‌بینی قیمت پویای آتی زعفران نگین، از یک مدل داده‌کاوی هیبریدی پیشنهادی استفاده شد. بر این اساس، نتایج حاصل از بررسی ویژگی داده‌ها، انتخاب وقفه بهینه با استفاده از الگوریتم ژنتیک، نوفه‌زادیی داده‌ها با استفاده از نظریهٔ موجک، تعیین الگوی بهینه در شبکه عصبی عمیق، مقایسه نتایج الگوی پیشنهادی با مدل رقیب و در نهایت، پیش‌بینی خارج از نمونه با داده‌های جدید برای قیمت آتی زعفران نگین ارائه شده است.

بررسی ویژگی داده‌ها

داده‌های مورد استفاده در مطالعهٔ حاضر شامل ۳۸۴ داده از قیمت روزانه سررسید قراردادهای آتی زعفران نگین در بورس کالای ایران در دوره زمانی ۱۳۹۸/۰۲/۰۱ تا ۱۳۹۹/۰۲/۱۵ است. مشخصات داده‌های مورد بررسی در جدول ۱ آمده است. بررسی ویژگی داده‌ها نشان داد که دامنه واریانس قیمت‌های آتی زعفران نگین از میانگین آنها کمتر است، بنابراین، میزان پراکندگی داده‌ها در وضعیت نسبتاً مناسب قرار دارد. نتایج حاصل از دو آزمون کولموگروف و شاپیرو[5] نیز نشان داد که داده‌های قیمت آتی زعفران نگین از توزیع نرمال برخوردار نیست. همچنین، بررسی آزمون مقادیر بحرانی[6] نیز نشان داد که سری زمانی داده‌های قیمت آتی زعفران نگین در ابتدای دوره دارای کمترین مقدار و در اواخر دوره، دارای بیشترین مقدار است؛ از این‌رو، قیمت آتی زعفران نگین در دوره

1. Matlab
2. Rapid Miner
3. Clementine
4. Python
5. Kolmogorov and Shapiro
6. extreme values

منتخب دارای نوسان بالاست و پیش‌بینی قیمت برای نمونه‌های خارج از نمونه از روابط غیرخطی تبعیت کرده و از ساختاری پیچیده برخوردار است.

جدول ۱- ویژگی‌های داده‌های روزانه سررسید قراردادهای آتی زعفران نگین در دوره مورد بررسی

تعداد داده‌ها	میانگین	انحراف معیار	حداقل	حداکثر	آزمون نرمال بودن کولموگروف		آزمون نرمال بودن شاپیرو	
					آماره	سطح معنی‌داری	آماره	سطح معنی‌داری
۳۵۴	۱۰۵۲۵۳,۹	۲۰۰۵۴,۹	۷۰۲۹۴	۱۷۲۱۴۹	۰/۱۸۳	۰/۰۰۰	۰/۹۱۱	۰/۰۰۰

مأخذ: یافته‌های پژوهش

انتخاب وقفه بهینه با استفاده از الگوریتم ژنتیک

یکی از مراحل مهم قبل از پیش‌بینی قیمت آتی زعفران نگین تعیین تعداد وقفه‌های بهینه مؤثر بر قیمت است. در مطالعهٔ حاضر، برای یافتن تعداد وقفه بهینه، از الگوریتم ژنتیک استفاده شد، چراکه این الگوریتم از قانون پارسیمونی اقتصادسنجی یا قانون حداقل صرفه‌جویی استفاده می‌کند. در واقع، قانون پارسیمونی پیشنهاد می‌دهد که برای تعیین تعداد وقفه بهینه و جلوگیری از حذف داده‌ها در مدل‌های غیرخطی، از یک الگوریتم متعالی استفاده شود. الگوریتم ژنتیک نیز یکی از بهترین الگوریتم‌های متعالی برای بهینه‌یابی است. از این‌رو، در پژوهش حاضر، برای تعیین وقفه بهینه در مدل‌های غیرخطی، از الگوریتم ژنتیک استفاده شد. تعریف الگوریتم ژنتیک نیز در نرم‌افزار داده‌کاوی Rapid Miner7.2 انجام شد. در واقع، این نرم‌افزار یکی از قوی‌ترین نرم‌افزارهای محاسبات الگوریتم‌های داده‌کاوی است و امکان ارتباط با سایر محیط‌های نرم‌افزاری از جمله Matlab، R و Python را نیز دارد. پس از چندین مرتبه آزمون و خطا در اجرای الگوریتم ژنتیک، در مدل نهایی، برای دستیابی به نقطه بهینه، روش Sequential انتخاب شد و جمعیت اولیه پنجاه نفر، نرخ جهش شش درصد، نرخ تلاقی ۲۵ درصد و چرخ رولت برای انتخاب نسل بعد به‌عنوان پارامترهای نهایی تعیین شد.

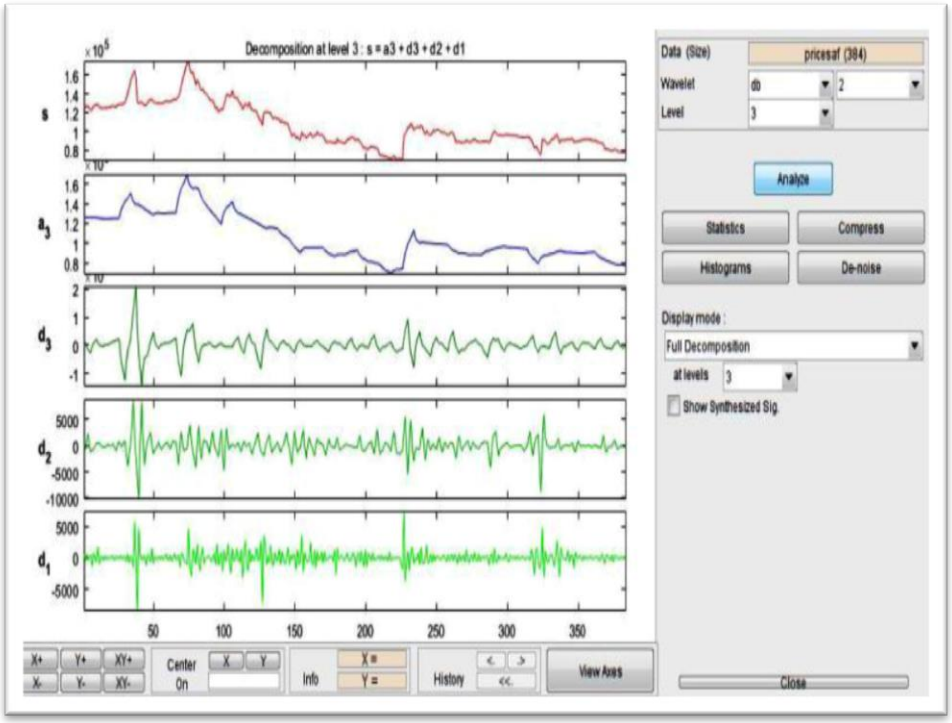
نتایج حاصل از انجام چرخه بهینه‌یابی برای انتخاب تعداد وقفه‌های بهینه با استفاده از مدل نهایی تعیین‌شده در الگوریتم ژنتیک نشان داد که بین یک تا شش وقفه برای قیمت آتی زعفران نگین، وقفه بهینه برای سری زمانی قیمت آتی این محصول «دو» است. بنابراین، قیمت آتی زعفران

نگین با دو وقفه می‌تواند بر روند پیش‌بینی آن مؤثر باشد. همچنین، می‌توان از متغیر قیمت دووقفه‌ای زعفران نگین برای فرآیند نوفه‌زدایی استفاده کرد.

نوفه‌زادیی داده‌های قیمت آتی زعفران نگین با استفاده از نظریهٔ موجک

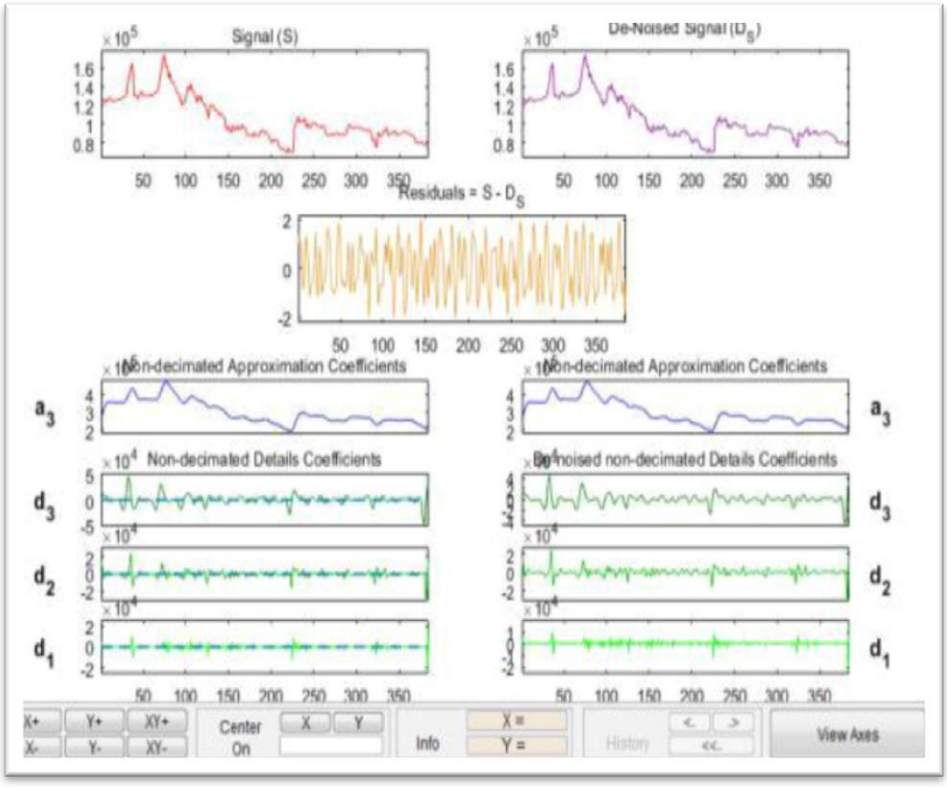
یکی از روش‌های بالا بردن دقت پیش‌بینی سری‌های زمانی دارای ریشه واحد نوفه‌زدایی (حذف نوفه) با استفاده از روش تبدیل موجک است. در پژوهش حاضر، برای نوفه‌زدایی داده‌ها، از مدل موجک گسسته در جعبه ابزار موجک نرم‌افزار متلب استفاده شد. حذف نوفه با استفاده از تبدیل موجک شامل سه مرحله اساسی تجزیه، حذف نوفه و بازسازی است. در مطالعهٔ حاضر، به‌منظور پیدا کردن نوع و طول موجک مناسب به‌منظور تجزیه سری زمانی، برای طول‌های مختلف توابع موجک دابشیز، کویفلت و سیملت[1]، سری زمانی شاخص تجزیه شده و سپس، با انجام عمل معکوس با استفاده از تبدیل موجک گسسته حداکثر همپوشانی، سری زمانی شاخص بازسازی شده است. آنگاه خانواده موجکی که کمترین مقدار خطا را به خود اختصاص داده، به‌عنوان خانواده موجک مناسب انتخاب شده است. بنابراین، در مرحله اول، نوع و طول موجک مناسب از میان انواع متنوع از خانواده موجک انتخاب شده و پس از آن، عملیات تجزیه و بازسازی آن در راستای دستیابی به حداقل خطا چندین مرتبه تکرار می‌شود. در مطالعهٔ حاضر، برای تجزیه داده‌های قیمت از توابع موجک سیملت، هار، دابشیز و کویفلت استفاده شد. برای انتخاب سطح بهینه تجزیه داده‌ها، ابتدا داده‌ها به شش سطح تجزیه شد و سپس، به کمک توابع و انتخاب‌گرهای موجود در منوی فرمان نرم‌افزار متلب و با استفاده از کمترین مقدار معیار RMSE، سطح بهینه تجزیه داده‌ها برابر با دو انتخاب شد. نتایج حاصل از این تجزیه در شکل ۳ نشان داده شده است. در این شکل، سری S سیگنال اصلی قیمت آتی زعفران نگین، a1 و d1 سری تقریب‌زده‌شده جزئیات فیلترشده هستند. نتایج حاصل از مراحل نوفه‌زدایی قیمت آتی زعفران نگین نیز در شکل ۴ و در نهایت، نتایج حاصل از مرحله پایانی نوفه‌زدایی (بازسازی قیمت آتی زعفران نگین) در شکل ۵ آمده است.. در شکل ۵، نمودار پیوسته (رنگ قهوه‌ای) و نقطه‌چین (سیاه رنگ)، به‌ترتیب، نشان‌دهنده قیمت آتی زعفران نگین بدون نوفه‌زدایی و با نوفه‌زدایی است. اختلاف این دو نمودار نیز به‌عنوان نوفه محسوب می‌شود.

1. Dabshyse, Coiflet and Symlet



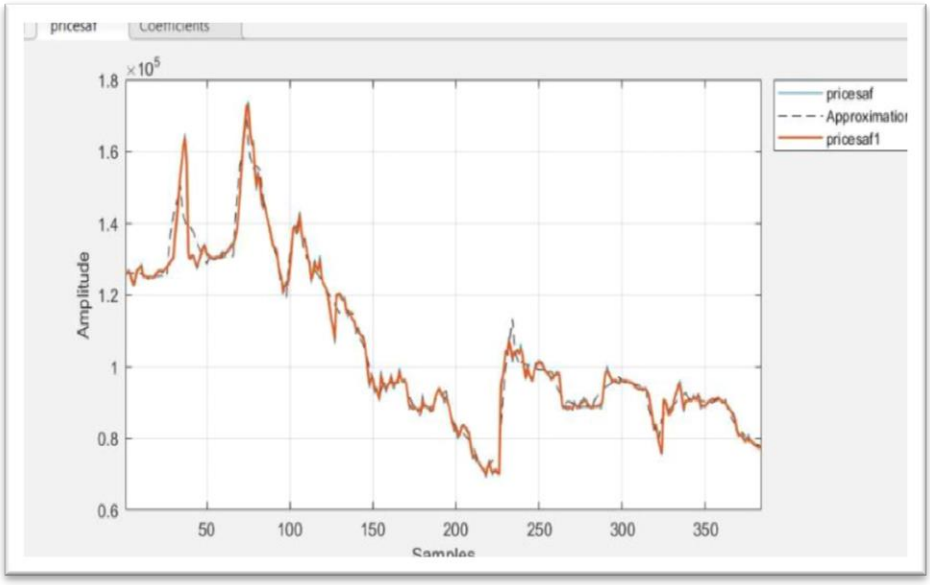
مأخذ: یافته‌های پژوهش

شکل ۳- سری زمانی قیمت آتی زعفران نگین و مؤلفه‌های تجزیه‌شده توسط موجک دابشیز ۳ با سطح بهینه ۲



مأخذ: یافته‌های پژوهش

شکل ۴- مراحل نوفه‌زدایی قیمت آتی زعفران نگین به‌همراه باقی‌مانده و سیگنال نوفه‌زدایی‌شده



مأخذ: یافته‌های پژوهش

شکل ۵- سری زمانی نهایی نوفه‌زدایی قیمت آتی زعفران نگین

تعیین الگوی بهینه در مدل شبکه عصبی عمیق

برای راه‌اندازی مدل شبکه عصبی عمیق، از الگوریتم آموزشی H2O استفاده شد. در روش آموزش H2O، تعداد گره‌ها و تعداد لایه‌های پنهان نقش مؤثری در دستیابی به دقت بالای پیش‌بینی مدل دارند. تعداد لایه‌های پنهان و گره‌های مدل شبکه عصبی از طریق کاهش گرادیان و کاهش میزان خطا محاسبه می‌شود. در پژوهش حاضر، برای افزایش دقت پیش‌بینی، دخالتی از سوی محقق برای تعیین تعداد گره‌ها در مدل لحاظ نشد و این مهم از طریق الگوریتم‌های هوشمند تدوین‌شده در توابع موجود انجام شده است. در این راستا، تعیین تعداد گره‌ها در لایه‌های پنهان توسط روش Epoch انجام شد. روش Epoch به‌صورت خودکار چند مرتبه داده‌های ورودی را وارد می‌کند تا حالت بهینه به‌دست آید. حالت بهینه نیز از طریق logloss مشخص می‌شود. در واقع، زمانی که logloss حالت نزولی پیدا کند، حالت بهینه به‌دست آمده است .در مطالعه حاضر، برای اطمینان از حالت بهینه، آموزش مدل تا یک مرحله بعد از حالت نزولی شدن چهل مرتبه با تعداد گره‌های مختلف انجام شد. پس از تعیین تعداد گره‌های لایه پنهان، خروجی‌های لایه پنهان که به‌صورت عدد پیوسته

است، وارد تابع فعال‌ساز می‌شود. در پژوهش حاضر، از تابع فعالساز بسیار کارآی Maxout استفاده شده است. در واقع، تابع فعال‌ساز Maxout مشخص‌کنندهٔ حد آستانه بهینه‌شده از طریق تغییر تعداد خطاست. این حد آستانه در دامنهٔ یک تا منفی یک تغییر می‌کند. در پژوهش حاضر، حد آستانه برای ورودی‌های مختلف متفاوت بود و سرانجام، پس از حل ۲۱ بار آزمایش، به نقطه بهینه رسید.

مقایسه نتایج حاصل از مقادیر پیش‌بینی‌شده توسط مدل داده‌کاوی هیبریدی پیشنهادی

در پژوهش حاضر، برای افزایش دقت پیش‌بینی، از شبیه‌سازی مونت‌کارلو استفاده شده است. در واقع، در روش مونت‌کارلو، بر اساس توزیع احتمال وقفه‌ها، برای مقادیر پیش‌بینی‌شده، به‌دفعات زیاد (مثلاً ده هزار بار) شبیه‌سازی تکرار می‌شود و میانگین آنها به‌عنوان مقدار پیش‌بینی‌شده مدل در نظر گرفته می‌شود، که دارای بالاترین احتمال است. به بیان دیگر، نقش روش مونت‌کارلو در مدل پیشنهادی این است که با محتمل‌ترین حالت، مقدار پیش‌بینی‌شده قیمت را مشخص می‌کند. نتایج حاصل از مقادیر شبیه‌سازی‌شده توسط مونت‌کارلو برای مدل پیشنهادی (GA-WA-LSTM-MCM) و مدل رقیب (GA-WA-NN-MCM) در جدول ۲ آمده است. شایان یادآوری است که برای نشان دادن اثر نوفه‌زدایی روی دقت پیش‌بینی قیمت زعفران نگین، پیش‌بینی با استفاده از دو مدل منتخب در دو مرحله یعنی، با استفاده از داده‌های اصلی (نوفه‌زدایی‌نشده) و داده‌های قیمت نوفه‌زدایی‌شده انجام شد. بر اساس نتایج جدول ۲، مقایسه وزن‌های وقفه‌های اول و دوم در همه مدل‌های منتخب نشان می‌دهد که قیمت روز قبل (وقفه اول) بیشترین اهمیت را در بررسی و پیش‌بینی قیمت آتی زعفران نگین دارد.

جدول ۲- نتایج حاصل از شبیه‌سازی مونت‌کارلو و وزن‌های وقفه‌های اول و دوم قیمت آتی زعفران نگین

مدل منتخب	وزن‌های هر کدام از وقفه‌ها		مقادیر واقعی و پیش‌بینی‌شده قیمت آتی زعفران نگین (ریال)		
	وزن وقفه اول	وزن وقفه دوم	قیمت در وقفه دوم	قیمت در وقفه اول	مقدار پیش‌بینی شده
GA-WA-LSTM-MCM	۰/۷۰۴	۰/۶۴۲	۱۰۵۰۹۹	۱۰۴۸۰۶	**۱۰۴۴۲۱**
GA- LSTM-MCM	۰/۷۰۴	۰/۶۳۰	۱۰۵۱۰۹	۱۰۴۷۸۷	**۱۰۴۵۷۷**
GA-WA-NN-MCM	۰/۹۸۸	۰/۲۲۱	۱۰۴۷۸۷	۱۰۵۱۰۹	**۱۰۴۲۰۹**
GA-NN-MCM	۰/۹۷۲	۰/۲۵۲	۱۰۵۰۹۹	۱۰۴۸۰۶	**۱۰۴۷۲۸**

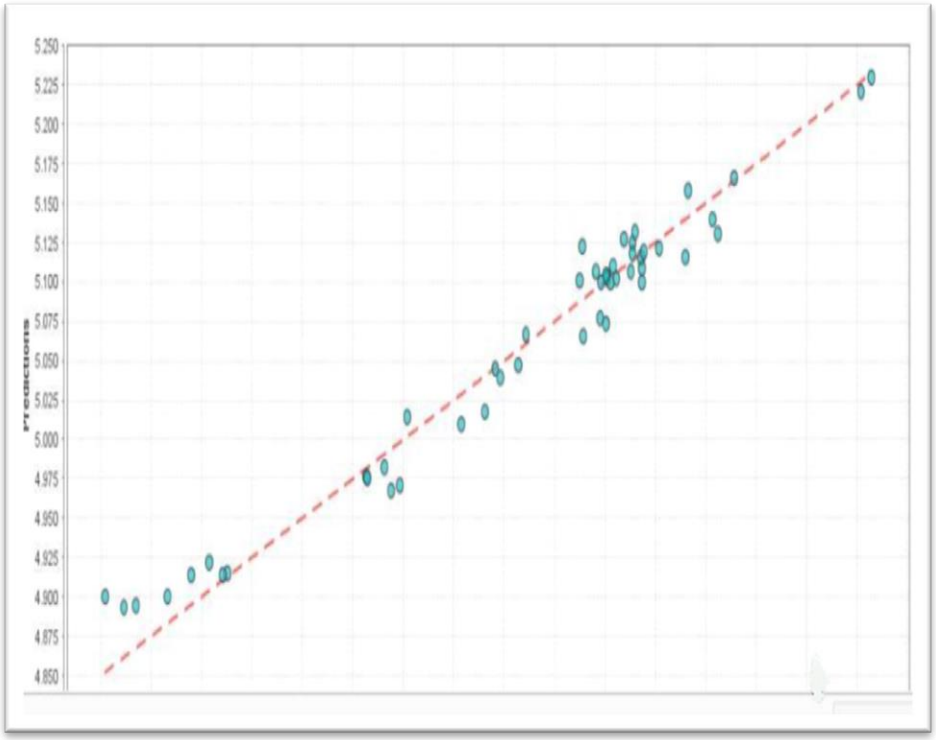
مأخذ: یافته‌های پژوهش

در جدول ۳، دقت پیش‌بینی مدل پیشنهادی با سه مدل رقیب با استفاده از معیارهای ارزیابی میانگین مجذور مربعات خطا (RMSE) و ضریب تعیین (R^2) مورد بررسی قرار گرفت. بر اساس نتایج جدول ۳، همه مدل‌های منتخب از تقریب مناسب برای پیش‌بینی قیمت آتی زعفران نگین در استفاده از داده‌های اصلی (نوفه‌دار) و داده‌های نوفه‌زدایی‌شده برخوردار بودند. مقایسه دقت پیش‌بینی مدل‌های نوفه‌زدایی‌شده توسط تابع موجک با مدل‌های نوفه‌زدایی‌نشده (یعنی، مقایسه مدل «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» با مدل «الگوریتم ژنتیک- شبکه عصبی عمیق- مونت‌کارلو» و مقایسه مدل «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت‌کارلو» با مدل «الگوریتم ژنتیک- شبکه عصبی ساده- مونت‌کارلو») نشان می‌دهد که به‌کارگیری نظریهٔ موجک برای نوفه‌زدایی داده‌ها سبب افزایش دقت پیش‌بینی قیمت آتی زعفران نگین شده است. همچنین، مقایسه دقت پیش‌بینی مدل پیشنهادی «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» با مدل «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت‌کارلو» نشان می‌دهد که استفاده از شبکه عصبی عمیق، در مقایسه با شبکه عصبی ساده، قدرت پیش‌بینی قیمت محصولات کشاورزی (قیمت آتی زعفران) را افزایش می‌دهد. در نهایت، مقایسه مدل پیشنهادی مطالعهٔ حاضر «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» با سه الگوی رقیب نشان می‌دهد که این الگو دارای عملکرد بهتری در پیش‌بینی قیمت است. نتیجه نهایی گویای آن است که استفاده از تابع موجک و مونت کارلو بر پایه شبکه عصبی عمیق سبب افزایش دقت پیش‌بینی قیمت محصولات کشاورزی می‌شود، که برازش عملکرد آن در شکل ۶ نشان داده شده است.

جدول ۳- مقایسه نتایج حاصل از پیش‌بینی مدل پیشنهادی و مدل‌های رقیب منتخب در داده‌های اصلی و نوفه‌زدایی‌شده

مدل منتخب	میانگین مربع مجذور خطا (RMSE)	ضریب تعیین (R^2)
GA-WA-LSTM-MCM	۱۲۸۸/۴	۰/۹۴۳
GA-LSTM-MCM	۱۷۱۵/۳	۰/۹۳۷
GA-WA-NN-MCM	۱۴۹۹/۶	۰/۹۱۹
GA-NN-MCM	۱۸۱۳/۴	۰/۸۹۱

مأخذ: یافته‌های پژوهش



مأخذ: یافته‌های پژوهش

شکل ۶- نمودار عملکرد مدل «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» برای قیمت آتی زعفران

پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید

همان‌گونه که گفته شد، استفاده از روابط و معادلات محاسبات پیچیده نرم برای انجام «پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید» یکی دیگر از نوآوری‌های مطالعهٔ حاضر بوده است. در واقع، در مطالعات مربوط به مقوله پیش‌بینی، تنها «مدل مناسب برای پیش‌بینی» معرفی شده و اما کاربرد عملی آن یعنی، ارائه مقادیر پیش‌بینی برای آینده مطرح نشده است. پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید مبتنی بر تعیین «سناریوهای داده‌های قیمت» و «روند تغییرات قیمت در سناریوهای تعیین شده» است. برای نمونه، در بازار بورس آتی زعفران، برخلاف بازار نقدی معاملات، در یک روز چندین قرارداد با سررسیدهای مختلف برای خرید زعفران وجود دارد. برای

انجام سناریوهای داده‌های قیمت از بین قراردادهای موجود، قراردادهایی که سررسید آنها به هم نزدیکترند، انتخاب می‌شوند و از داده‌های انتهایی قیمت قراردادها (مثلاً برای ده روز آخر) میانگین گرفته می‌شود. این اعداد میانگین‌گیری‌شده به‌عنوان داده‌های سناریوی قیمت (ستون دوم جدول ۳) انتخاب می‌شوند. رسم روند این داده‌ها نیز روند افزایشی یا کاهشی پیش‌بینیی خارج از نمونه را تعیین خواهد کرد. نتایج حاصل از پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید در جدول ۴ آمده است. دوره مورد نظر برای پیش‌بینی خارج از نمونه از دوم تا دهم اردیبهشت ۱۳۹۹ است. سناریوهای داده‌های جدید قیمت بر اساس توزیع نرمال داده‌های اولیه کاهشی بوده و از ۸۷ تا ۸۰ هزار ریال در نظر گرفته شده است. مقایسه مقادیر پیش‌بینی‌شده خارج از نمونه با داده‌های واقعی بیانگر کارآیی و دقت بالای مدل پیشنهادی برای پیش‌بینی کوتاه‌مدت است.

جدول ۴- پیش‌بینی خارج از نمونه با استفاده از مدل «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو» برای قیمت آتی زعفران (ریال)

تاریخ آینده	سناریوهای داده‌های جدید قیمت (ریال)	مقادیر پیش‌بینی قیمت خارج از نمونه (ریال)	مقادیر واقعی قیمت (ریال)	اختلاف مقادیر پیش‌بینی‌شده با مقادیر واقعی (ریال)
۱۳۹۹/۲/۲	۸۷۰۰۰	۸۷۰۱۴	۸۷۲۴۱	۲۲۷
۱۳۹۹/۲/۳	۸۶۰۰۰	۸۵۵۲۱	۸۵۸۷۳	۳۵۲
۱۳۹۹/۲/۴	۸۵۰۰۰	۸۳۱۱۶	۸۱۷۷۷	-۱۳۳۹
۱۳۹۹/۲/۵	۸۴۰۰۰	۸۱۲۱۰	۸۰۵۷۵	-۶۳۵
۱۳۹۹/۲/۶	۸۳۰۰۰	۸۰۴۱۷	۸۱۰۹۴	۶۷۷
۱۳۹۹/۲/۷	۸۲۰۰۰	۷۹۸۶۵	۸۰۳۲۱	۴۵۶
۱۳۹۹/۲/۹	۸۱۰۰۰	۷۹۶۸۱	۷۹۱۴۴	-۵۳۷
۱۳۹۹/۲/۱۰	۸۰۰۰۰	۷۹۲۷۹	۷۹۳۴۱	۶۲

مأخذ: یافته‌های پژوهش

نتیجه‌گیری و پیشنهادها

هدف پژوهش حاضر طراحی یک مدل داده‌کاوی هیبریدی برای پیش‌بینی دقیق قیمت محصولات کشاورزی با بهره‌گیری از مجموعه مدل‌های غیرخطی الگوریتم ژنتیک، تبدیل موجک گسسته، شبکه عصبی عمیق و روش مونت‌کارلو بوده و دارای این ویژگی‌هاست: ۱- مدل پیشنهادی از نوع هیبریدی دومرحله‌ای و مدل پایه هیبریدی غیرخطی- غیرخطی است؛ ۲- در این مدل

پیشنهادی، از الگوریتم ژنتیک برای تعیین وقفه بهینه سری زمانی قیمت، از تابع موجک برای نوفه‌زدایی سری زمانی قیمت و از شبکه عصبی عمیق برای پیش‌بینی سری زمانی قیمت استفاده شده است؛ ۳- در این مدل، از روش مونت‌کارلو برای شبیه‌سازی محتمل‌ترین احتمال قیمت در تکمیل فرآیند پیش‌بینی استفاده شده است؛ و ۴- در الگوی پیشنهادی پژوهش، از روابط و معادلات محاسبات پیچیده نرم برای انجام «پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید» استفاده شده است. نتایج حاصل از مقایسه مدل پیشنهادی «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت‌کارلو»، با استفاده از معیارهای ارزیابی، با سه مدل رقیب «الگوریتم ژنتیک- شبکه عصبی عمیق- مونت‌کارلو»، «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت‌کارلو» و «الگوریتم ژنتیک- شبکه عصبی ساده- مونت کارلو» نشان داد که الگوی پیشنهادی نسبت به سه مدل رقیب دارای عملکرد بهتری در پیش‌بینی قیمت آتی زعفران است. همچنین، استفاده از شبکه عصبی عمیق در مقایسه با شبکه عصبی ساده، به‌کارگیری نظریهٔ موجک برای نوفه‌زدایی و نیز استفاده از روش مونت‌کارلو برای شبیه‌سازی قیمت‌های پیش‌بینی‌شده دقت پیش‌بینی قیمت آتی زعفران را افزایش می‌دهد. علاوه بر این، استفاده از محاسبات نرم برای انجام «پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید» نشان داد که مدل پیشنهادی از کارآیی لازم و دقت بالا برای پیش‌بینی کوتاه‌مدت قیمت آتی زعفران برخوردار است. قیمت آتی پیش‌بینی‌شده برای زعفران نگین با استفاده از مدل پیشنهادی، ۱۰۴۴۲۱ ریال برای هر گرم بوده و در واقع، این قیمت یک مرجع قیمتی در اصلاح قیمت بازاری زعفران نگین است، که از درجه اطمینان بالا و میزان حداقل مخاطره برای فعالان بازار بورس در یک دوره زمانی مشخص برخوردار بوده و نزدیک شدن به قیمت یادشده یا دور شدن از آن به‌لحاظ مدیریت مخاطره در تحلیل روند آتی بازار مهم است. با توجه به نتایج به‌دست‌آمده، پیشنهادهایی به‌شرح زیر ارائه می‌شود:

۱- مطالعه حاضر در دستیابی به شاخص میزان دقت حداکثری، سناریوسازی روند قیمت‌های آتی، تحلیل حساسیت مؤلفه‌های مؤثر بر قیمت و سرانجام، پیش‌بینی قیمت آینده با تکیه بر مدل‌های داده‌کاوی هوشمند از جایگاهی بسیار مناسب برخوردار است، از این‌رو، شایسته است که مدیریت بورس کالای ایران برای تعیین راهبردهای اتاق پایاپای معاملات، مدیریت مخاطره و کشف قیمت، به بهره‌گیری از مزایای این مدل پیشنهادی در پیش‌بینی قیمت آتی بپردازد. همچنین، پژوهشگران و کنشگران بازار بورس و علاقه‌مندان فن پیش‌بینی می‌توانند از این مدل در پیش‌بینی کوتاه‌مدت و بلندمدت سایر قراردادهای بورس آتی استفاده کنند.

۲- از آنجا که داده‌های سری زمانی قیمت محصولات کشاورزی دارای ماهیت غیرخطی و نامانایی زیادی است، پیشنهاد می‌شود که مانند مطالعهٔ حاضر، از روش تبدیل موجک برای نوفه‌زدایی داده‌ها و از الگوریتم‌های فراابتکاری به‌منظور پیدا کردن بیشترین تأثیر وقفه‌های متغیرهای مؤثر برای اجرای مدیریت مخاطره در بورس استفاده شود.

۳- استفاده از روش مونت‌کارلو برای تعیین یک سطح قیمت پیش‌بینی مرجع با «درجه اطمینان بالا و میزان حداقل مخاطره» از دیگر پیشنهادهای پژوهش به‌شمار می‌رود.

۴- برای مطالعات آتی، بهتر است که از سایر الگوهای داده‌کاوری مانند مدل ماشین بردار پشتیبان، مدل تصادفی جنگلی و ... نیز برای مقایسه با نتایج استفاده شود.

منابع

1. Ajmera, R., Kook, N. and Crilley, J. (2012). Impact of commodity price movements on CPI inflation. *Monthly Labor Review*, 135: 29-43.
2. Alinejad, M., Bakhtiari, B. and Ghaderi, K. (2017). The comparison of Monte Carlo method and the combined method of fuzzy logic-PSO. *Journal of Marine Engineering*, 13(52): 105-112. (Persian)
3. Chen, Q., Lin, X., Zhong, Y. and Xie, Z. (2019). Price prediction of agricultural products based on wavelet analysis-LSTM. *2019 IEEE International Conference on Parallel and Distributed Processing with Applications, Big Data and Cloud Computing, Sustainable Computing and Communications, Social Computing and Networking (ISPA/BDCloud/SocialCom/SustainCom)*, Xiamen, China, pp. 984-990. DOI: 10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00142.
4. Chuluunsaikhan, T., Ryu, G., Yoo, K.H., Rah, H. and Nasridinov, A. (2020). Incorporating deep learning and news topic modeling for forecasting pork prices: the case of South Korea. *Journal of Agriculture*, 10(11): 513. Available at file:///C:/Users/m.noushmand.APERI/Downloads/agriculture-10-00513-v2.pdf.
5. Ghias, M. (2014). Introduction on the method of Monte Carlo simulation. *Scientific Journal of BASPARESH*, 4(1): 67-77. (Persian)

6. Goodfellow, I., Bengio, Y. and Courville, A. (2016). Deep learning. NIT Press Book. Available at https://www. deeplearningbook.org.
7. Haven, E., Liu, X. and Shen, L. (2012). De-noising option prices with the wavelet method, *European Journal of Operational Research*, Elsevier, 222(1): 104-112. DOI: 10.1016/j.ejor.2012.04.020.
8. Jammazi, R. and Aloui, Ch. (2012). Crude oil price forecasting: experimental evidence from wavelet decomposition and neural network modeling. *Energy Economics*, 34: 828-841.
9. Li, X., He, K., Lai, K K. and Zou, Y. (2014). Forecasting crude oil price with multiscale de-noising ensemble model. *Journal of Mathematical Problems in Engineering*, (Special Issue): 1-19. Available at https://downloads.hindawi.com/journals/mpe/2014/716571.pdf.
10. Ly, R., Traore, F. and Dia, K. (2021). Forecasting commodity prices using long-short-term memory neural networks. IFPRI Discussion Paper 2000. Washington, DC: International Food Policy Research Institute (IFPRI). DOI: 10.2499/p15738coll2.134265.
11. Moghaddasi, R. and Jaleh Rajabi, M. (2011). Integrated modeling approach for the price forecasting of agricultural products. *Journal of Economics and Development of Agriculture (Sciences and Industries of Agriculture)*, 5(3): 355-364. (Persian)
12. Mohammadi, T., Taklif, A. and Zamani, S. (2017). Natural gas price forecasting using the combination of wavelet transform and artificial neural network (case study: US market). *Quarterly Journal of Iranian Economic Research*, 22(71): 1-26. (Persian)
13. Nassar, L., Okwuchi, I.E., Saad, M., Karray, F. and Ponnambalam, K. (2020). Deep learning based approach for fresh produce market price prediction. Proceedings of IEEE 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1-7.
14. Nazlioglu, S. and Soytas, U. (2011). World oil prices and agricultural commodity prices: evidence from an emerging market. *Journal of Energy Economics*, 33: 488-496.

15. Nazlioglu, S. (2011). World oil and agricultural commodity prices: evidence from nonlinear causality. *Journal of Energy Policy*, 39(5): 2935-2943.
16. Polyiam, K. and Boonrawd, P. (2017). A hybrid forecasting model of cassava price based on artificial neural network with support vector machine technique, Third International Conference on Information Management (ICIM), Chengdu, China, 2017, pp. 123-127. DOI: 10.1109/INFOMAN.2017.7950359.
17. Rai, R. and Mahmoudi Azar, M. (2014). The forecasting of the future return of the stock market using ARIMA, neural network and wavelet de-noising models. *Quarterly Journal of Asset Management and Financing*, 2(2:5): 1-16. (Persian)
18. Rai, R., Mohammadi, Sh. and Fenderski, H. (2015). The forecasting of the stock price index using neural network and wavelet transform. *Quarterly Journal of Asset Management and Financing*, 3(1:8): 55-74. (Persian)
19. Raikar, R.V., Wang, Ch-Yi., Shih, H.P. and Hong, J.H. (2016). Prediction of contraction scour using ANN and GA. *Flow Measurement and Instrumentation*, 50: 26-34.
20. Rasheed, A., Younis, M.S., Ahmad, F., Qadir, J. and Kashif, M. (2021). District wise price forecasting of wheat in Pakistan using deep learning. arXiv preprint arXiv:2103.04781. Available at https://arxiv.org/pdf/2103.04781.pdf.
21. Reboredo, J.C. and Rivera-Castro, M.A. (2013). A wavelet decomposition approach to crude oil price and exchange rate dependence. *Economic Modelling*, 32: 42-57. DOI: 10.1016/j.econmod.2012.12.028.
22. Sabu, K.M. and Kumar, T.M. (2020). Predictive analytics in agriculture: forecasting prices of Arecanuts in Kerala. *Journal of Procedia Computer Science*, 171: 699-708.
23. Sadeghi, H. and Dehghani Firoozabadi, Z. (2017). The de-noising of financial time series using wavelet analysis. *Journal of Financial Engineering and Securities Management*, 33: 299-315. (Persian)

24. Sangwana, K.S., Saxenaa, S. and Kanta, G. (2016). Optimization of machining parameters to minimize surface roughness using integrated ANN-GA approach, *The 22nd CIRP Conference on Life Cycle Engineering*, *Procedia CIRP*, 29: 305-310.
25. Shabri, A. and Samsudin, R. (2014). Daily crude oil price forecasting using hybridizing wavelet and artificial neural network model. Hindawi Publishing Corporation, Mathematical Problems Engineering, Vol. 2014, ID: 201402, pp. 10.
26. Shao, Y.E. and Dai, J.T. (2018). Integrated feature selection of ARIMA with computational intelligence approaches for food crop price prediction. *Complexity*, Vol. 2018, pp. 17. DOI: 10.1155/2018/1910520.
27. Shiferaw, Y.A. (2019). Time-varying correlation between agricultural commodity and energy price dynamics with Bayesian multivariate DCC-GARCH models. *Physica A: Statistical Mechanics and its Applications*, 526: 120807.
28. Vautard, R., Yiou, P. and Ghil, M. (1992). Singular-spectrum analysis: a toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenomena*, 58(1-4): 95-126.
29. Wang, L. and Gupta, S. (2013). Neural networks and wavelet de-noising for stock trading and prediction. In: W. Pedrycz and S.M. Chen (Eds) Time series analysis, modeling and applications. Intelligent Systems Reference Library, vol. 47, pp. 229-247. Springer, Berlin, Heidelberg. Available at https://doi.org/10.1007/978-3-642-33439-9_11.
30. Wang, B., Liu, P., Chao, Z., Junmei, W., Chen, W., Cao, N., O'Hare, G.M.P. and Wen, F. (2018). Research on hybrid model of garlic short-term price forecasting based on big data. *Journal of Computers, Materials and Continua (CMC)*, 57(2): 283-296.
31. Wang, D., Yue, Ch., Wei, Sh. and Lv, J. (2017). Performance analysis of four decomposition-ensemble models for one-day-ahead agricultural commodity futures price forecasting. *Journal of Algorithms*, 10(3): 108. DOI: 10.3390/a10030108.

32. Wang, L., Feng, J., Sui, X., Chu, X. and Mu, W. (2020). Agricultural product price forecasting methods: research advances and trend. *British Food Journal*, 122(7): 2121-2138.
33. Weng, Y., Wang, X., Hua, J., Wang, H., Kang, M. and Wang, F.Y. (2019). Forecasting horticultural products price using ARIMA model and neural network based on a large-scale data set collected by web crawler. *Journal of IEEE Transactions on Computational Social Systems*, 6(3): 547-553.
34. Xiong, T., Li, Ch., Bao, Y., Hu, Zh. and Zhang, L. (2015). A combination method for interval forecasting of agricultural commodity futures prices. *Journal of Knowledge-Based Systems*, 77: 92-102.
35. Yang, X. (2007). A modified particle swarm optimizer with dynamic adaptation Baoding: applied mathematics and Computation 51, June 2007, *Elsevier*, 189: 1205-1213.
36. Yang, Y., Chen, Y., Wang, Y., Li, C. and LiSchool, L. (2016). Modelling a combined method based on ANFIS and neural network improved by DE algorithm: a case study for short-term electricity demand forecasting. *Journal of Applied Soft Computing*, 49: 663-675.
37. Zhang, X., Lai, K.K. and Wang, Sh-Y. (2008). A new approach for crude oil price analysis based on Empirical Mode Decomposition. *Energy Economics*, Elsevier, 30(3): 905-918.
38. Zhang, D., Chen, S., Liwen, L. and Xia, Q. (2020). Forecasting agricultural commodity prices using model selection framework with time series features and forecast horizons. *Journal of IEEE Access*, 8: 28197-28209.
39. Zolfaghari, M., Sahabi, B. and Bakhtiaran, M.J. (2020). Designing a model for forecasting of return of the total stock market index (with emphasis on combined models of deep learning network and GARCH models). *Quarterly Journal of Asset Management and Financing*, 42: 138-171. (Persian)