

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

DOI: 10.30490/AEAD.2023.357440.1412

مقاله پژوهشی

به کارگیری مدل داده کاوی هیبریدی (الگوریتم ژنتیک-موجک- شبکه عصبی عمیق - شبیه سازی مونت کارلو) برای پیش بینی قیمت محصولات کشاورزی: مطالعه موردی قیمت آتی زعفران در بورس کالای کشاورزی*

رضا حیدری^۱، سید محمدرضا حاج سیدجواد^۲

تاریخ پذیرش: ۱۴۰۱/۵/۲۳

تاریخ دریافت: ۱۴۰۰/۱۰/۱۹

چکیده

پیش بینی قیمت و روند تغییرات آن از مهم ترین عوامل در تصمیم گیری و تدوین راهبردهای مربوط به محصولات کشاورزی است. هدف مطالعه حاضر ارائه یک مدل یا الگوی داده کاوی هیبریدی شامل مجموعه مدل های غیرخطی الگوریتم ژنتیک، تبدیل موجک، شبکه عصبی عمیق و روش مونت کارلو برای پیش بینی دقیق قیمت محصولات کشاورزی بود. این الگوی پیشنهادی از نوع هیبریدی دومرحله ای و مدل پایه هیبریدی

* مطالعه حاضر برگرفته از طرح پژوهشی داخلی مؤسسه پژوهش های برنامه ریزی، اقتصاد کشاورزی و توسعه روستایی با عنوان «طراحی مدل پیش بینی قیمت قرارداد آتی زعفران در بورس کالای کشاورزی» می باشد.

۱- نویسنده مسئول و استادیار پژوهشی، مؤسسه پژوهش های برنامه ریزی، اقتصاد کشاورزی و توسعه روستایی، تهران، ایران. (rezaheidari3631@gmail.com)

۲- پژوهشگر، مؤسسه پژوهش های برنامه ریزی، اقتصاد کشاورزی و توسعه 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 مونت کارلو 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 - مونت کارلو» با سه الگوی رقیب «الگوریتم ژنتیک - شبکه عصبی عمیق - مونت کارلو»، «الگوریتم ژنتیک - تبدیل موجک - شبکه عصبی ساده - مونت کارلو» 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 نظریه موجک برای نوفه‌زدایی و استفاده از روش مونت کارلو برای شبیه‌سازی قیمت‌های پیش‌بینی شده 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 کارایی 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، نظریه موجک، شبکه عصبی عمیق، روش مونت کارلو.

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 کمتر توسعه‌یافته 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 به عنوان یک علامت اجتماعی بین عرضه و تقاضا، به دلایل متعدد از جمله افزایش تقاضای جهانی، تغییرات پیچیده مرتبط با عوامل چرخه‌ای، تغییرات اقلیمی، سیاست‌ها و تناقض بین تولید و بازاریابی، به شدت 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 به دلیل سوء تغذیه مزمن، در معرض تهدید مستقیم قرار داده و از این رو، 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 (Chuluunsai Khan 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 به دست 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 پژوهشی 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 عمده 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 مخاطره قیمتی بازار آتی این محصول باشد. سؤال اساسی این است که «در مواجهه با روش‌های متعدد پیش‌بینی قیمت، برای پیش‌بینی قیمت آتی

زعفران، کدام روش ها را باید انتخاب کرد؟». بنابراین، هدف مطالعه حاضر ارائه یک مدل داده کاوی هیبریدی به منظور پیش بینی دقیق قیمت آتی زعفران در بورس کالای کشاورزی در قالب ابزاری مناسب و کارآمد برای نیل به اهداف مطالعه (پیش بینی قیمت محصولات کشاورزی) است. از سوی دیگر، با توجه به ماهیت غیرخطی و نامانایی در رفتار سری زمانی قیمت، به کارگیری یک مدل هیبریدی متشکل از مدل های غیرخطی تبدیل موجک، الگوریتم ژنتیک، شبکه عصبی عمیق و روش مونت کارلو یک نوآوری در کسب نتایج دقیق محسوب می شود.

پس از مقدمه، در بخش مبانی نظری و ادبیات موضوع، به بررسی انواع مدل های داده کاوی برای پیش بینی قیمت پویای محصولات کشاورزی و برخی از مطالعات پیشین در این زمینه و در بخش مدل تحقیق و روش برآورد، به معرفی ویژگی های مدل پیشنهادی و اجزای آن و در بخش نتایج نیز به ارزیابی شیوه استخراج داده ها و نتایج حاصل از مدل پیشنهادی و عملکرد آن در مقایسه با مدل های رقیب پرداخته می شود؛ و در پایان پژوهش نیز نتیجه گیری و پیشنهادهایی ارائه می شود.

پیشینه تحقیق

بر اساس ادبیات نظری، در سراسر جهان، به کارگیری الگوهای داده کاوی برای پیش بینی قیمت به عنوان یک اصل اساسی در عملکرد خوب اقتصادی مورد پذیرش قرار گرفته و البته، در زمینه پیش بینی 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 می شود.

نتایج مطالعه رشید و همکاران (Rasheed et al., 2021) نشان داد که مدل حافظه بلند-کوتاه مدت^۱، در مقایسه با مدل های سری زمانی آماری و یادگیری ماشین معمولی، عملکرد قابل توجهی در پیش بینی قیمت گندم دارد. نصار و همکاران (Nassar et al., 2020)، در مقایسه مدل های پیش بینی قیمت یادگیری عمیق با هشت مدل روی مجموعه داده های سری زمانی سبزی ها، میوه ها و گل ها، نشان دادند که مدل حافظه بلند-کوتاه مدت (LSTM) و مدل ترکیبی LSTM و شبکه عصبی کانولوشنی^۲ در پیش بینی دقیق قیمت محصولات تازه کارآمدترند. سابو و کومار (Sabu and Kumar, 2020) از مدل های سری زمانی و یادگیری ماشین برای پیش بینی قیمت ماهانه دانه درخت خرما^۳ در ایالت کرالای هند استفاده کردند و دریافتند که شبکه عصبی LSTM دارای دقت پیش بینی

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

بالا تری 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 نرمال‌سازی شده 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 ۲۰۱۰-۲۰۱۷ استفاده کردند؛ مدل پیشنهادی آنها متشکل از مدل میانگین متحرک خودبازگشتی^۱ به عنوان بخش خطی و مدل ماشین بردار پشتیبان^۲ به عنوان 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 به عنوان 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 آبی گندم، ذرت و سویا استفاده شد. در این روش، چهار مدل هیبریدی بر پایه شبکه عصبی انتشار برگشتی^۳ و بهینه‌سازی ازدحام ذرات^۴ پیشنهاد شده است. نتایج این مطالعه نشان داد که همه «مدل‌های هیبریدی همراه با شیوه تجزیه» عملکرد بهتری نسبت به مدل‌های انفرادی دارند. پلیام و بونراد (Polyiam and Boonrawd, 2017)، برای پیش‌بینی قیمت کاساوا (نوعی سبزی ریشه‌ای^۵) 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)، هاوون 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 به منظور 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 نوفه‌زدایی شده است. در تحقیقی دیگر، راعی و محمودی آذر (Rai and Mahmoudi Azar, 2014)، با استفاده از مدل‌های آریمای 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 آریمای شبکه عصبی دارد و مدل‌های شبکه عصبی نیز قدرت پیش‌بینی بهتری را نسبت به مدل‌های آریمای نشان می‌دهد؛ همچنین، مقادیر مربوط به آزمون دایبولد- ماریانو این نتایج را تأیید می‌کند.

بررسی ادبیات نظری پیرامون پیش‌بینی قیمت محصولات کشاورزی در مطالعه حاضر نشان داد که: ۱- مطالعات زیادی با مدل‌های مختلف برای پیش‌بینی قیمت‌های بسیاری از کالاهای کشاورزی در خارج از ایران در مواردی به صورت 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 مطالعه حاضر با به کارگیری و معرفی اجزا (تبدیل موجک برای نوفه‌زدایی، الگوریتم ژنتیک برای تعیین وقفه بیهینه، شبکه عصبی عمیق برای پیش‌بینی و روش مونت کارلو برای شبیه‌سازی) به منظور 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 مطالعه 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^۱ و ۲- مدل‌های هیبریدی مبتنی بر چارچوب «تجزیه و جمع‌بندی»^۲. 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^۳ (Reboredo and Rivera-Castro, 2013)، تجزیه تجربی^۴ (Zhang et al., 2008) و تجزیه و تحلیل طیف منفرد^۵ (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، به گونه ای که در این نوع مدل هیبریدی، ابتدا قیمت محصولات کشاورزی در چندین مرحله مورد تجزیه و تحلیل قرار می گیرد و سپس، با استفاده از روش های منتخب، طی یک مسیر مشخص، این اجزا به طور جداگانه تکمیل می شوند؛ و در نهایت، با دریافت یک درک کلی از نتایج قیمت، نتایج پیش بینی حاصل می شود.

به طور کلی، در مدل های هیبریدی، ممکن است تجزیه و جمع بندی فرآیند تبدیل اجرا از طریق مدل های منتخب به یکدیگر تک مرحله ای^۱ یا دو مرحله ای^۲ باشد. در روش تک مرحله ای، مدل پایه در بخش ورودی داده ها یا داده های خروجی (نتایج مدل پایه) هیبرید می شود. اما، در روش دو مرحله ای، هر دو بخش فرآیند هیبریدی در هر دو بخش ورودی داده ها و خروجی نتایج اتفاق می افتد. همچنین، بیان این نکته ضروری است که ساختار مدل های هیبریدی نیز متفاوت است و با توجه به میزان اهمیت دقت پیش بینی در مطالعه مورد نظر، به سه حالت کلی تقسیم می شود، که عبارت اند 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 مطالعه حاضر از نوع هیبریدی دو مرحله ای و مدل پایه هیبریدی غیرخطی - غیرخطی بوده و از این رو، کامل ترین نوع مدل هیبریدی است.

در شکل ۱، شمای کلی مدل داده کاوی هیبریدی پیشنهادی برای پیش بینی قیمت محصولات کشاورزی (در مطالعه حاضر، قیمت آتی زعفران نگین) نشان داده شده است. ویژگی های مدل پیشنهادی پژوهش حاضر یعنی، «الگوریتم ژنتیک - تبدیل موجک - شبکه عصبی عمیق - مونت کارلو» (GA-WA-LSTM-MCM) عبارت اند از:

۱- مطابق با مبانی نظری، این مدل پیشنهادی از نوع هیبریدی دو مرحله ای و مدل پایه غیرخطی - غیرخطی است. توضیح آنکه در مدل پیشنهادی، داده های ورودی به مدل شبکه عصبی عمیق (LSTM) توسط مدل ناپارامتریک تبدیل موجک گسسته (WA) نوفه زدایی می شود. همچنین، خروجی مدل LSTM توسط مدل غیرخطی مونت کارلو شبیه سازی می شود تا بالاترین احتمال رخداد قیمت در آینده به دست 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 شیوه مونت کارلو (MCM) برای شبیه‌سازی محتمل‌ترین احتمال قیمت به همراه آزمون قیمت‌های جدید^۱ در تکمیل فرآیند پیش‌بینی استفاده شده است.

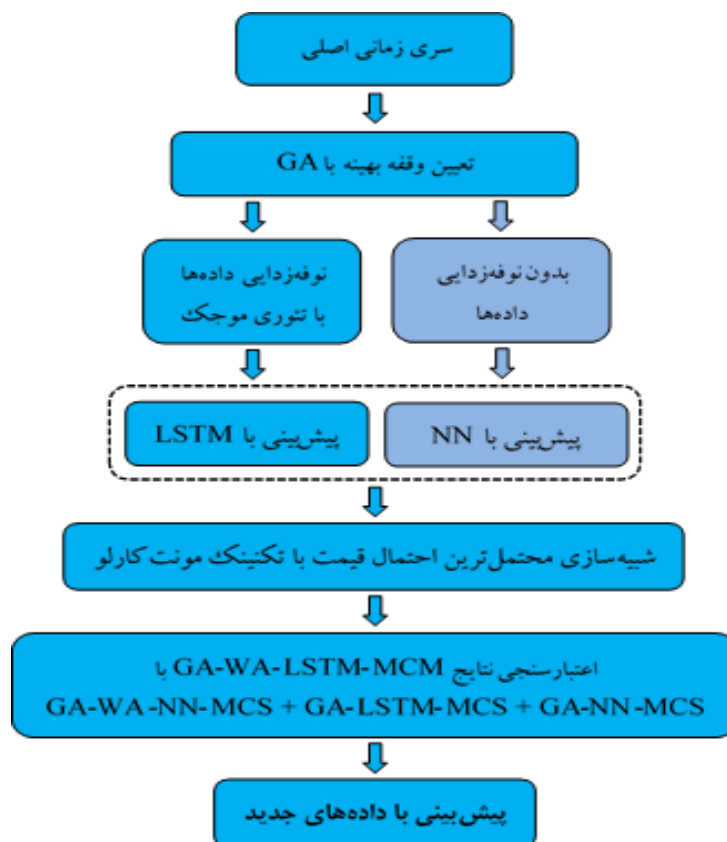
۴- استفاده از روابط و معادلات محاسبات پیچیده نرم^۲ برای انجام «پیش‌بینی خارج از نمونه با مجموعه داده‌های جدید»^۳ یکی دیگر از نوآوری‌های مطالعه حاضر است.

برای نشان دادن دقت پیش‌بینی مدل پیشنهادی «الگوریتم ژنتیک-تبدیل موجک-شبکه عصبی عمیق-مونت کارلو» با استفاده از معیارهای ارزیابی، خروجی حاصل از پیش‌بینی مدل پیشنهادی مطالعه حاضر با سه مدل دیگر یعنی، مدل‌های «الگوریتم ژنتیک-شبکه عصبی عمیق-مونت کارلو»، «الگوریتم ژنتیک-تبدیل موجک-شبکه عصبی ساده-مونت کارلو» و «الگوریتم ژنتیک-شبکه عصبی ساده-مونت کارلو» 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-مونت کارلو»، به جای 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-مونت کارلو» و «الگوریتم ژنتیک-شبکه عصبی ساده-مونت کارلو»، 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 آنان، 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 نوفه‌زدایی شده به عنوان داده‌های ورودی برای پیش‌بینی به شبکه عصبی عمیق منتقل می‌شود. به دیگر سخن، تبدیل موجک به عنوان یک روش پیش‌پردازش به کار 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، کوپفلت و میر اشاره کرد. به‌طور کلی، موجک گسسته مادر به صورت رابطه (۱) تعریف می‌شود:

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

که در آن، n مکان موجک و m اندازه موجک است. معمولاً برای ساده‌تر شدن مقادیر $s_o = 2$ و $u = 1$ در نظر گرفته می‌شود. در این حالت، تابع موجک گسسته مادر به صورت رابطه (۲) درمی‌آید:

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

که در آن، $w_{m,n}$ ضرایب حاصل از استفاده از تابع موجک برای سری زمانی گسسته در حالت $s = 2^m$ و $u = 2^m \times n$ است. بر اساس نظریه مالات^۱ 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 به صورت رابطه (۳) به دست می‌آید:

$$x(t) = T + \sum_{m=1}^M \sum_{t=0}^{M-m-1} w_{m,n} 2^{-\frac{m}{2}} \Psi(2^{-m}t - n) \quad (3)$$

شکل ساده رابطه (۳) به صورت رابطه (۴) است:

$$x(t) = A_M(t) + \sum_{m=1}^M D_M(t) \quad (4)$$

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

الگوریتم ژنتیک

الگوریتم ژنتیک یکی از روش های جدید بهینه سازی است که بیشتر برای بهینه سازی مسائل بسیار پیچیده و غیرخطی به کار می رود. الگوریتم ژنتیک الهامی از علم ژنتیک و نظریه تکامل داروین بوده و بر اساس بقای برترین ها یا انتخاب طبیعی استوار است. ایده اساسی این الگوریتم انتقال خصوصیات موروثی توسط کروموزوم هاست. این الگوریتم با الهام از طبیعت بر پایه اصل تکاملی «پایداری بهترین ها» استوار است. یکی از برتری های این الگوریتم نسبت به سایر مدل های مشابه عدم وابستگی آن به فرضیه های آماری محدودکننده و نرمال بودن توزیع نسبت ها یا برابری واریانس یا کوواریانس ماتریس نسبت هاست (Yang, 2007). الگوریتم ژنتیک، با تعریف یک کروموزوم که باید بهینه باشد، شروع می شود. هر درایه در هر کروموزوم «ژن» نامیده می شود. مجموعه ای از این کروموزوم ها جمعیت الگوریتم را شکل می دهند. هر تکرار برای بهینه سازی فرآیند مسئله داده شده به عنوان یک نسل شناخته می شود. جمعیت افراد با استفاده از نسل های مداوم در هر نسل جدید توسعه می یابد. در هر مرحله، جمعیت جدید تولید شده 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 رضایت بخش توسط کاربران ادامه می یابد. تبدیل جمعیت افراد به نسل جدید توسط سه عامل اصلی ژنتیکی انتخاب^۱، آمیزش^۲ و جهش^۳ انجام می شود. بعد از ایجاد نسل های متعدد، الگوریتم ژنتیک روی راه حل بهینه همگرا می شود. اجرای موفقیت آمیز 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، مدل های شبکه عمیق از دقیق ترین روش های در دسترس برای پیش بینی به شمار می روند. به طور کلی، شبکه های عمیق از پشت سر هم قرار دادن شبکه های کم عمق (مانند شبکه عصبی تک لایه) تشکیل می شوند. در نتیجه، شبکه های عمیق قادر به یادگیری و استخراج ویژگی های سلسه مراتبی غیرخطی است و می تواند الگوهای آماری پیچیده را ذخیره کند. مدل های عمیق کارایی خود را برای کاربردهای مختلف به اثبات رسانده اند. شبکه های عصبی

1. selection
2. crossover
3. mutation

عمیق در ساختار خود از نوع یادگیری بدون نظارت^۱ 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^۲، مفهومی به‌عنوان حافظه بلندمدت و کوتاه‌مدت در شبکه عصبی پدید آمد که به‌اختصار، مدل حافظه 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 کنترل‌کننده میزان محتوای 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 $\tilde{C}_t$ 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:

$$h_t = \Gamma_o \tanh(C_t) \quad (5)$$

$$\Gamma_o = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (6)$$

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

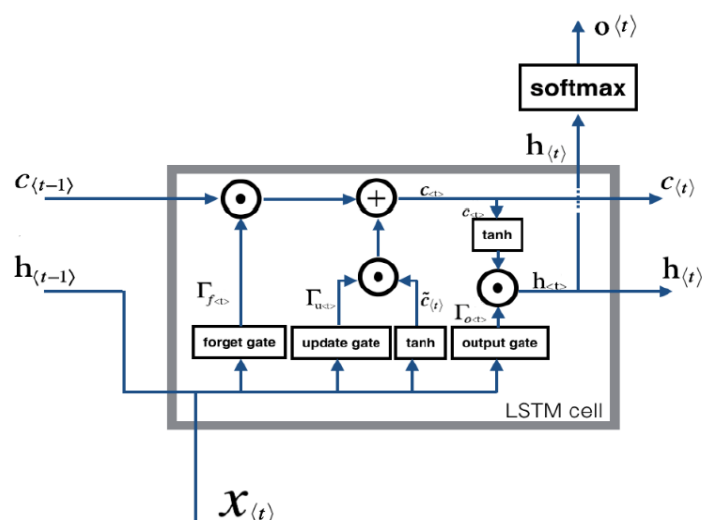
به کارگیری مدل داده کاوی هیبریدی.....

$$C_t = \Gamma_f \cdot C_{t-1} + \Gamma_u \cdot \widehat{C}_t \quad (7)$$

$$\widehat{C}_t = \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \quad (8)$$

$$\Gamma_f = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f) \quad (9)$$

$$\Gamma_u = \sigma(W_u \cdot [h_{t-1}, X_t] + b_u) \quad (10)$$



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، به اصطلاح، 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^۱، ۲- دروازه به‌روزرسانی^۲ که به دروازه ورودی معروف است و ۳- دروازه خروجی^۳. علاوه بر این سه دروازه، یک سلول حافظه^۴ نیز وجود دارد که از آن به اختصار 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 به صورت 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 به صورت 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 به صورت 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 مونت کارلو

روش مونت کارلو به هر شیوه‌ای گفته می‌شود که از طریق نمونه‌سازی آماری، پاسخ‌های تقریبی برای مسائل کمی فراهم می‌کند. شبیه‌سازی مونت کارلو بیشتر به‌منظور توصیف روشی برای انتشار عدم قطعیت‌های موجود در ورودی مدل به عدم قطعیت‌ها در خروجی مدل به کار می‌رود. بنابراین، مونت کارلو 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 مونت کارلو در شکل ساده یا پیچیده شامل چند جزء اصلی است. این اجزا عبارت‌اند از اعداد تصادفی، تابع توزیع احتمال، قواعد نمونه‌برداری، تخمین خطا و کاهش انحراف معیار پراکندگی، موازی‌سازی - بردارسازی محاسباتی و الگوریتم متروپلیس. هسته اصلی هر روش مدل‌سازی تصادفی یا روش مونت کارلو بر مبنای استفاده مداوم از اعداد تصادفی است. روش مونت کارلو یک الگوریتم محاسباتی است که از نمونه‌گیری تصادفی برای محاسبه 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 (مونت کارلو) 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.

به منظور 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 شیوه 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^۱ 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، به گونه‌ای 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 به صورت 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 مطالعه حاضر، از شاخص‌های مجذور میانگین مربعات خطا^۲ و ضریب تعیین (R^2) استفاده می‌شود. هر کدام از این دو شاخص می‌تواند مدل پیاده‌سازی‌شده در مطالعه حاضر را اعتبارسنجی کند، به گونه‌ای که مدلی دارای بهترین عملکرد است که کمترین مقدار مجذور میانگین مربعات (RMSE) یا بیشترین مقدار R^2 را داشته باشد. شاخص‌های یادشده به صورت روابط زیر است:

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

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (11)$$

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

که در این روابط، Y_i داده‌های واقعی (مشاهداتی)، $\hat{Y}_i$ داده‌های پیش‌بینی شده (شبیه‌سازی شده)، $\bar{Y}_i$ میانگین داده‌های واقعی و n دوره زمانی ارزیابی است (Yang et al., 2016). برای انجام پژوهش حاضر، از نرم‌افزارهای متلب^۱، ریپید ماینر^۲، کلمنتین^۳ و پایتون^۴ 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، نوفه‌زایی داده‌ها با استفاده از نظریه 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 مطالعه حاضر شامل ۳۸۴ داده از قیمت روزانه سرسید 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 کولموگوروف و شاپیرو^۵ نیز نشان داد که داده‌های قیمت آتی زعفران نگین از توزیع نرمال برخوردار نیست. همچنین، بررسی آزمون مقادیر بحرانی^۶ 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 این رو، قیمت آتی زعفران نگین در دوره

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

به کارگیری مدل داده کاوی هیبریدی.....

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	
					سطح معنی‌داری	آماره	سطح معنی‌داری	آماره
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 مطالعه 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 Miner 7.2 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

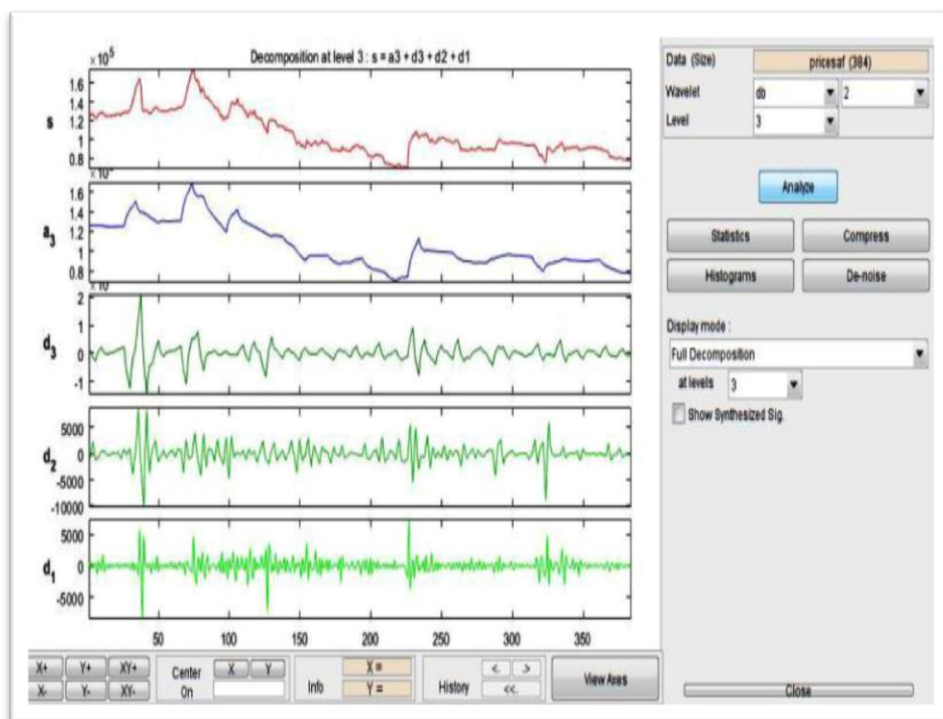
نگین با دو وقفه می‌تواند بر روند پیش‌بینی آن مؤثر باشد. همچنین، می‌توان از متغیر قیمت دووقفه‌ای زعفران نگین برای فرآیند نوفه‌زدایی استفاده کرد.

نوفه‌زدایی داده‌های قیمت آتی زعفران نگین با استفاده از نظریه 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 مطالعه حاضر، به‌منظور پیدا کردن نوع و طول موجک مناسب به‌منظور تجزیه سری زمانی، برای طول‌های مختلف توابع موجک دابشیز، کوئفلت و سیملت^۱، 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 مطالعه حاضر، برای تجزیه داده‌های قیمت از توابع موجک سیملت، هار، دابشیز و کوئفلت 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، a_1 و d_1 سری تقریب‌زده شده 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)، به ترتیب، نشان‌دهنده قیمت آتی زعفران نگین بدون نوفه‌زدایی و با نوفه‌زدایی است. اختلاف این دو نمودار نیز به‌عنوان نوفه محسوب می‌شود.

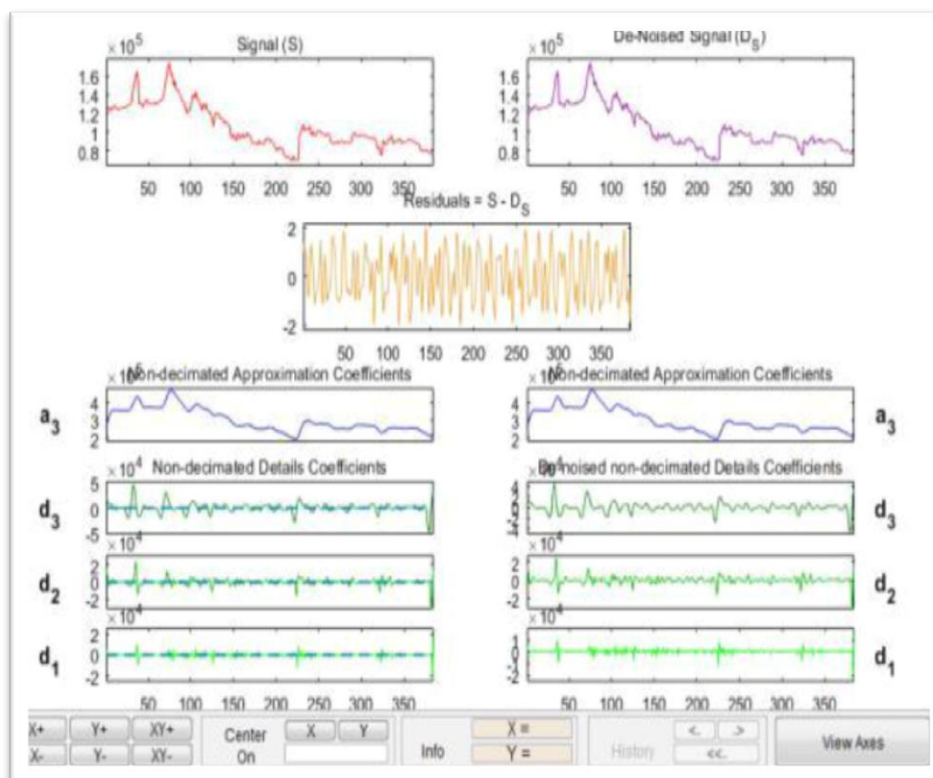
1. Dabshyse, Coiflet and Symlet

به کارگیری مدل داده کاوی هیبریدی.....



مأخذ: یافته‌های پژوهش

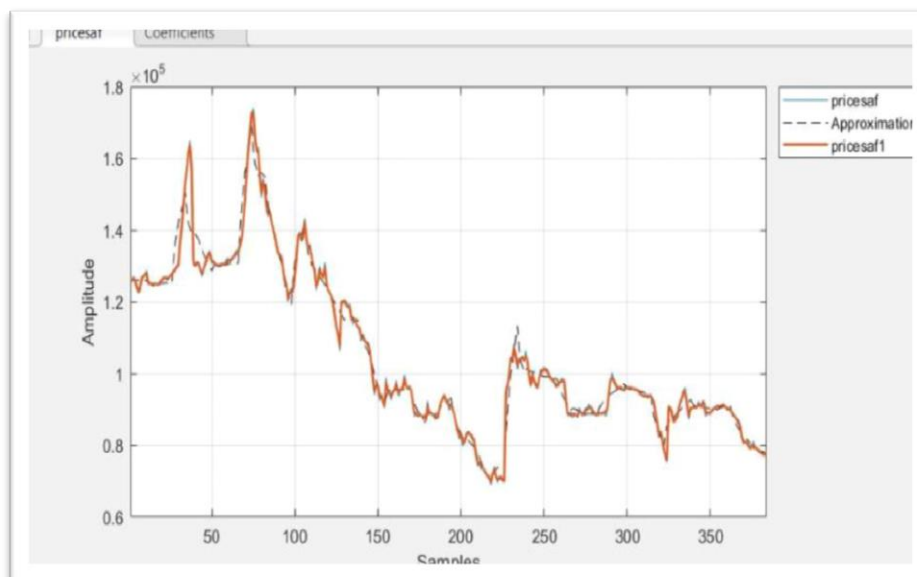
شکل ۳- سری زمانی قیمت آتی زعفران نگین و مؤلفه‌های تجزیه شده توسط موجک دابشیز ۳ با سطح بهینه ۲



مأخذ: یافته‌های پژوهش

شکل ۴- مراحل نوفه‌زدایی قیمت آتی زعفران نگین به همراه باقی مانده و سیگنال نوفه‌زدایی شده

به کارگیری مدل داده کاوی هیبریدی.....



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 گرادین 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 فعال ساز می شود. در پژوهش حاضر، از تابع فعال ساز بسیار کارآی Maxout استفاده شده است. در واقع، تابع فعال ساز Maxout مشخص کننده حد آستانه بهینه شده از طریق تغییر تعداد خطاست. این حد آستانه در دامنه یک تا منفی یک تغییر می کند. در پژوهش حاضر، حد آستانه برای ورودی های مختلف متفاوت بود و سرانجام، پس از حل ۲۱ بار آزمایش، به نقطه بهینه رسید.

مقایسه نتایج حاصل از مقادیر پیش بینی شده توسط مدل داده کاوی هیبریدی پیشنهادی

در پژوهش حاضر، برای افزایش دقت پیش بینی، از شبیه سازی مونت کارلو استفاده شده است. در واقع، در روش مونت کارلو، بر اساس توزیع احتمال وقفه ها، برای مقادیر پیش بینی شده، به دفعات زیاد (مثلاً ده هزار بار) شبیه سازی تکرار می شود و میانگین آنها به عنوان مقدار پیش بینی شده مدل در نظر گرفته می شود، که دارای بالاترین احتمال است. به بیان دیگر، نقش روش مونت کارلو در مدل پیشنهادی این است که با محتمل ترین حالت، مقدار پیش بینی شده قیمت را مشخص می کند. نتایج حاصل از مقادیر شبیه سازی شده توسط مونت کارلو 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 نوفه زدایی روی دقت پیش بینی قیمت زعفران نگین، پیش بینی با استفاده از دو مدل منتخب در دو مرحله یعنی، با استفاده از داده های اصلی (نوفه زدایی نشده) و داده های قیمت نوفه زدایی شده انجام شد. بر اساس نتایج جدول ۲، مقایسه وزن های وقفه های اول و دوم در همه مدل های منتخب نشان می دهد که قیمت روز قبل (وقفه اول) بیشترین اهمیت را در بررسی و پیش بینی قیمت آتی زعفران نگین دارد.

جدول ۲- نتایج حاصل از شبیه سازی مونت کارلو و وزن های وقفه های اول و دوم قیمت آتی زعفران نگین

مدل منتخب	وزن های هر کدام از وقفه ها		مقادیر واقعی و پیش بینی شده قیمت آتی زعفران نگین (ریال)		
	وزن وقفه اول	وزن وقفه دوم	قیمت در وقفه دوم	قیمت در وقفه اول	مقدار پیش بینی 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

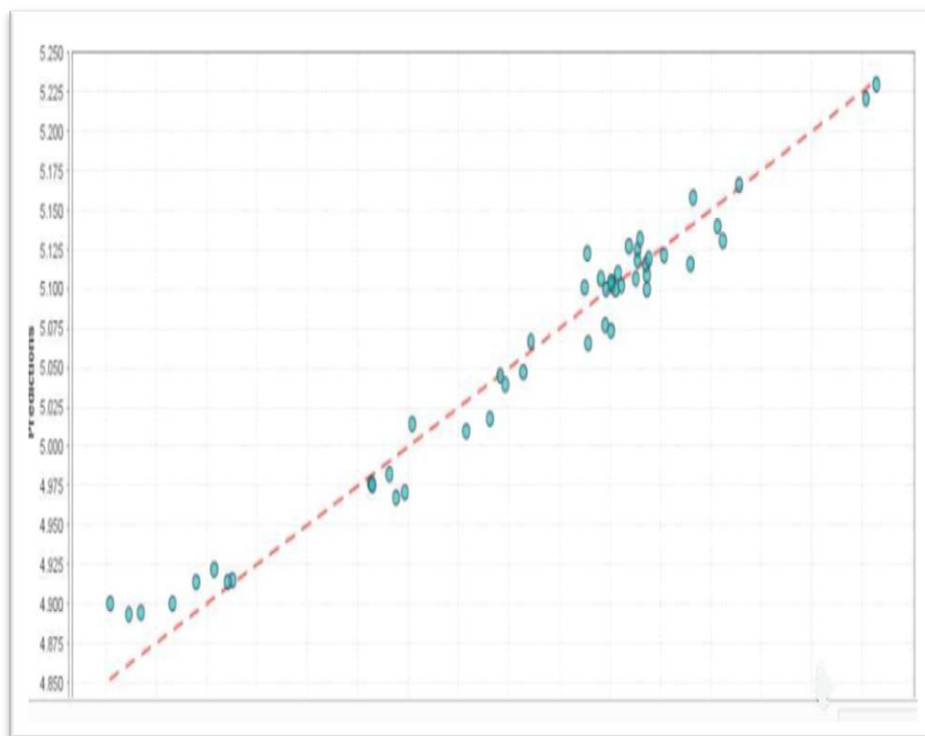
مأخذ: یافته های پژوهش

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 آبی زعفران نگین در استفاده از داده‌های اصلی (نوفه‌دار) و داده‌های نوفه‌زدایی شده برخوردار بودند. مقایسه دقت پیش‌بینی مدل‌های نوفه‌زدایی شده توسط تابع موجک با مدل‌های نوفه‌زدایی نشده (یعنی، مقایسه مدل «الگوریتم ژنتیک-تبدیل موجک- شبکه عصبی عمیق- مونت کارلو» با مدل «الگوریتم ژنتیک- شبکه عصبی عمیق- مونت کارلو» و مقایسه مدل «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت کارلو» با مدل «الگوریتم ژنتیک- شبکه عصبی ساده- مونت کارلو») نشان می‌دهد که به کارگیری نظریه موجک برای نوفه‌زدایی داده‌ها سبب افزایش دقت پیش‌بینی قیمت آبی زعفران نگین شده است. همچنین، مقایسه دقت پیش‌بینی مدل پیشنهادی «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت کارلو» با مدل «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت کارلو» نشان می‌دهد که استفاده از شبکه عصبی عمیق، در مقایسه با شبکه عصبی ساده، قدرت پیش‌بینی قیمت محصولات کشاورزی (قیمت آبی زعفران) را افزایش می‌دهد. در نهایت، مقایسه مدل پیشنهادی مطالعه حاضر «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی عمیق- مونت کارلو» 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 نوفه‌زدایی شده

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- مونت کارلو» 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 مطالعه 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 «رشد 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

به کارگیری مدل داده کاوی هیبریدی.....

انجام سناریوهای داده های قیمت از بین قراردادهای موجود، قراردادهایی که سررسید آنها به هم نزدیکترند، انتخاب می شوند و از داده های انتهایی قیمت قراردادها (مثلاً برای ده روز آخر) میانگین گرفته می شود. این اعداد میانگین گیری شده به عنوان داده های سناریوی قیمت (ستون دوم جدول ۳) انتخاب می شوند. رسم روند این داده ها نیز روند افزایشی یا کاهشی پیش بینی خارج از نمونه را تعیین خواهد کرد. نتایج حاصل از پیش بینی خارج از نمونه با مجموعه داده های جدید در جدول ۴ آمده است. دوره مورد نظر برای پیش بینی خارج از نمونه از دوم تا دهم اردیبهشت ۱۳۹۹ است. سناریوهای داده های جدید قیمت بر اساس توزیع نرمال داده های اولیه کاهشی بوده و از ۸۷ تا ۸۰ هزار ریال در نظر گرفته شده است. مقایسه مقادیر پیش بینی شده خارج از نمونه با داده های واقعی بیانگر کارایی و دقت بالای مدل پیشنهادی برای پیش بینی کوتاه مدت است.

جدول ۴- پیش بینی خارج از نمونه با استفاده از مدل «الگوریتم ژنتیک- تبدیل موجک-

شبکه عصبی عمیق- مونت کارلو» برای قیمت آتی زعفران (ریال)

تاریخ آینده	سناریوهای داده های جدید قیمت (ریال)	مقادیر پیش بینی قیمت خارج از نمونه (ریال)	مقادیر واقعی قیمت (ریال)	اختلاف مقادیر پیش بینی شده با مقادیر واقعی (ریال)
۱۳۹۹/۲/۲	۸۷۰۰۰	۸۷۰۱۴	۸۷۳۴۱	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: یافته های پژوهش

نتیجه گیری و پیشنهادها

هدف پژوهش حاضر طراحی یک مدل داده کاوی هیبریدی برای پیش بینی دقیق قیمت محصولات کشاورزی با بهره گیری از مجموعه مدل های غیرخطی الگوریتم ژنتیک، تبدیل موجک گسسته، شبکه عصبی عمیق و روش مونت کارلو بوده و دارای این ویژگی ها است: ۱- مدل پیشنهادی از نوع هیبریدی دومرحله ای 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 مونت کارلو 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- مونت کارلو»، با استفاده از معیارهای ارزیابی، با سه مدل رقیب «الگوریتم ژنتیک- شبکه عصبی عمیق- مونت کارلو»، «الگوریتم ژنتیک- تبدیل موجک- شبکه عصبی ساده- مونت کارلو» 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 نظریه موجک برای نوفه‌زدایی و نیز استفاده از روش مونت کارلو 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 کارایی 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 مطالعه حاضر، از روش تبدیل موجک برای نوفه‌زدایی داده‌ها و از الگوریتم‌های فراابتکاری به منظور پیدا کردن بیشترین تأثیر وقفه‌های متغیرهای مؤثر برای اجرای مدیریت مخاطره در بورس استفاده شود.
- ۳- استفاده از روش مونت کارلو 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 داده کاوری 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. MIT 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. Reboledo, 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)

